

**МИНИСТЕРСТВО ОБРАЗОВАНИЯ И НАУКИ
РОССИЙСКОЙ ФЕДЕРАЦИИ**

ФЕДЕРАЛЬНОЕ АГЕНТСТВО ПО ОБРАЗОВАНИЮ

**САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИНФОРМАЦИОННЫХ ТЕХНОЛОГИЙ, МЕХАНИКИ И ОПТИКИ**



ПОБЕДИТЕЛЬ КОНКУРСА ИННОВАЦИОННЫХ ОБРАЗОВАТЕЛЬНЫХ ПРОГРАММ ВУЗОВ

НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК

Выпуск 38

**ИССЛЕДОВАНИЯ В ОБЛАСТИ
ОПТИКИ, ПРИБОРОСТРОЕНИЯ
И УПРАВЛЕНИЯ**

Труды молодых ученых



**САНКТ-ПЕТЕРБУРГ
2007**

Выпуск содержит материалы **IV межвузовской конференции молодых ученых.**, организованной 10–13 апреля 2007 года Санкт-Петербургским государственным университетом информационных технологий, механики и оптики в сотрудничестве с Балтийским государственным техническим университетом «Военмех»

Башкирским государственным университетом

Белорусским государственным педагогическим университетом им. Максима Танка

Белорусским государственным технологическим университетом

Белорусским государственным университетом информатики и радиоэлектроники

Дальневосточным государственным университетом

Дальневосточной академией государственной службы

Институтом аналитического приборостроения Российской Академии Наук (РАН)

Институтом Солнечно-Земной Физики СО РАН

Институтом химии высокочистых веществ РАН (г. Нижний Новгород)

Казанским государственным техническим университетом им. А.Н. Туполева

Казанским государственным университетом

Карельским государственным педагогическим университетом

Костромским государственным технологическим университетом

Красноярским государственным техническим университетом

Ленинградским государственным университетом им. А.С. Пушкина

Магнитогорским государственным техническим университетом им. Г.И. Носова

Морской государственной академией им. адмирала Ф.Ф. Ушакова

Московским государственным институтом электронной техники (техническим университетом)

Московским государственным техническим университетом им. Н.Э. Баумана

Московским педагогическим государственным университетом

Муромским институтом Владимирского государственного университета

Петербургским государственным университетом путей сообщения

Пятигорским государственным лингвистическим университетом

Российским государственным гидрометеорологическим университетом

Самарским государственным архитектурно-строительным университетом

Санкт-Петербургским государственным горным институтом им. Г.В. Плеханова (техническим университетом)

Санкт-Петербургским государственным инженерно-экономическим университетом (ИНЖЭКОН)

Санкт-Петербургским государственным политехническим университетом

Санкт-Петербургским государственным университетом

Санкт-Петербургским государственным университетом аэрокосмического приборостроения

Санкт-Петербургским институтом машиностроения (ЛМЗ-ВТУЗ)

Санкт-Петербургским университетом кино и телевидения

Санкт-Петербургской государственной академией физической культуры им. П.Ф. Лесгафта

Санкт-Петербургской государственной педиатрической медицинской академией

Северо-Западной академией государственной службы

Северо-Осетинским государственным университетом им. К.Л. Хетагурова

Тамбовским государственным университетом им. Г.Р. Державина

Татарским государственным гуманитарно-педагогическим университетом

Университетом Aix-Marseille II (Франция)

Университетом Прованса (Франция)

ФГУП "ЦНИИ им. академика А.Н. Крылова"

Энгельским технологическим институтом Саратовского государственного технического университета

В выпуске представлены работы, поддержанные финансированием в рамках:

- аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы (2006–2008 гг.)» (Федеральное агентство по образованию);
 - Федеральной целевой программы развития образования на 2006–2010 гг. (Федеральное агентство по образованию);
 - Российского фонда фундаментальных исследований,
- а также инициативные разработки.

ПРОГРАММНЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Председатель – ректор СПбГУ ИТМО, д.т.н., профессор **В.Н. Васильев**

Сопредседатели – проректор по развитию, д.т.н., профессор **В.О. Никифоров**,
проректор по УО и АР, д.ф.-м.н., профессор **Ю.Л. Колесников**,
проректор по УМР, к.т.н., профессор **А.А. Шехонин**,
декан факультета ППО, д.т.н., профессор **В.Л. Ткалич**

Члены программного комитета – д.т.н., профессор **Ю.А. Гатчин**, д.т.н., профессор **В.М. Мусалимов**, д.т.н., профессор **С.Б. Смирнов**, д.т.н., профессор **В.А. Тарлыков**, д.т.н., профессор **Е.Б. Яковлев**, к.т.н. **Т.В. Точилина**

ОРГАНИЗАЦИОННЫЙ КОМИТЕТ КОНФЕРЕНЦИИ

Председатель – начальник НИЧ **Л.М. Студеникин**

Зам. председателя – к.т.н. **Т.В. Точилина**

Члены организационного комитета – **П.А. Борисов**, **Н.Н. Валентик**, **И.Н. Жданов**, **С.Ю. Кerpелева**, **Н.В. Когай**, **А.В. Козаченко**, **И.М. Кудрявцева**, **Д.В. Лукичѳв**, **А.А. Малинин**, **Л.В. Можжухина**, **Ю.С. Монахов**, **Н.Б. Нечаева**, **М.В. Никитина**, **М.С. Петрищев**, **С.С. Резников**, **В.Н. Фролков**



В 2007 году СПбГУ ИТМО стал победителем конкурса инновационных образовательных программ вузов России на 2007–2008 годы. Реализация инновационной образовательной программы «Инновационная система подготовки специалистов нового поколения в области информационных и оптических технологий» позволит выйти на качественно новый уровень подготовки выпускников и удовлетворить возрастающий спрос на специалистов в информационной, оптической и других высокотехнологичных отраслях экономики.

ISSN 1819-222X

© Санкт-Петербургский государственный университет информационных технологий, механики и оптики, 2007

ИСПОЛЬЗОВАНИЕ ФЕМТОСЕКУНДНОГО ЛАЗЕРНОГО ИЗЛУЧЕНИЯ ДЛЯ СОЗДАНИЯ ВОЛНОВОДНЫХ СТРУКТУР В ХАЛЬКОГЕНИДНЫХ СТЕКЛООБРАЗНЫХ ПОЛУПРОВОДНИКАХ

М.А. Лесик, А.А. Шимко

(Санкт-Петербургский государственный университет)

Научный руководитель – к.ф.-м.н., доцент А.А. Маньшина

Явление изменения свойств оптических материалов под воздействием лазерного излучения может быть использовано для создания оптической памяти, волноводов, устройств экспресс-анализа (lab-on-a-chip), а также элементов интегральной оптики. В наших экспериментах структурные изменения индуцировались в образцах аморфных полупроводников под воздействием фемтосекундного лазерного излучения. Получены зависимости диаметра волноводных структур от условий лазерного воздействия (мощность лазерного излучения, скорость сканирования, число сканирований).

Введение

На сегодняшний момент устройства «оптической памяти» считаются наиболее перспективными при разработке и создании систем хранения информации. В таких устройствах оптическое излучение используется для записи и/или считывания информации. При этом «лазерная запись» информации осуществляется за счет структурных изменений оптических материалов, индуцированных под воздействием лазерного излучения [1–4]. Локальное изменение химической структуры, индуцированное лазерным воздействием, приводит также и к локальному изменению оптических параметров материала (показателя преломления, коэффициента поглощения и т.д.). Выделяют два основных метода создания структур в прозрачных оптических материалах.

Первый метод основан на использовании лазерного излучения УФ-диапазона. При поглощении одного фотона УФ-излучения, соответствующего энергии запрещенной зоны материала (линейный процесс), происходит фотолитический разрыв химических связей и реорганизация структуры материала. Так как коэффициент поглощения УФ-излучения для большинства материалов велик, то структурные изменения индуцируются на поверхности образца. Второй метод основан на использовании фемтосекундного лазерного излучения. Благодаря высокой интенсивности лазерного излучения развиваются нелинейные процессы (многофотонное поглощение, туннельная ионизация и т.д.), приводящие к локальной ионизации и структурным изменениям вещества. Одним из важных отличий такого метода является возможность создания структурных изменений (записи информации) не только на поверхности, но и в объеме материала, что делает возможным создание элементов интегральной оптики с трехмерной архитектурой.

Использование фемтосекундного лазерного излучения открывает новые возможности при создании элементной базы для микроэлектроники, микрооптики, микробиологии и микрохимии [2]. С помощью фемтосекундных лазерных импульсов можно создавать такие оптические элементы, как волноводы, решетки, разветвители, линзы и т.д.

Целью данной работы было создание структурных изменений под воздействием фемтосекундного лазерного излучения и получение волноводных структур в различных образцах халькогенидных стекол (на основе систем As-S и Ga-Ge-S). В работе исследовано влияние геометрии записи, параметров лазерного излучения, скорости сканирования и количества сканирований на свойства полученных волноводных структур.

Геометрии записи

Волноводные структуры, разветвители, а также элементы интегральной оптики могут быть созданы при сканировании сфокусированным лазерным излучением в объеме образца. При этом выделяют 2 геометрии записи:

- (1) перпендикулярная запись – образец перемещается перпендикулярно лазерному лучу, т.е. на рис. 1 в плоскости $xу$;
- (2) параллельная запись – при перемещении образца по оси z .

Степень изменения оптических параметров в значительной степени зависит от механизмов структурных изменений. Так, например, при высокой частоте следования лазерного излучения происходит локальное плавление вещества с последующим затвердеванием, в результате чего меняется плотность и, соответственно, показатель преломления. При высокой мощности лазерного излучения может происходить значительное изменение показателя преломления за счет миграции компонентов материала.

Экспериментальные результаты

В качестве источника фемтосекундного лазерного излучения использовали лазер на Тi-сапфире с накачкой Ar^+ лазера. Параметры лазерного излучения: $\lambda=800$ нм, длительность импульса $\tau = 100$ фс, частота следования $f = 80$ МГц. Лазерное излучение фокусировалось на образец, помещенный на трехмерную подвижку.

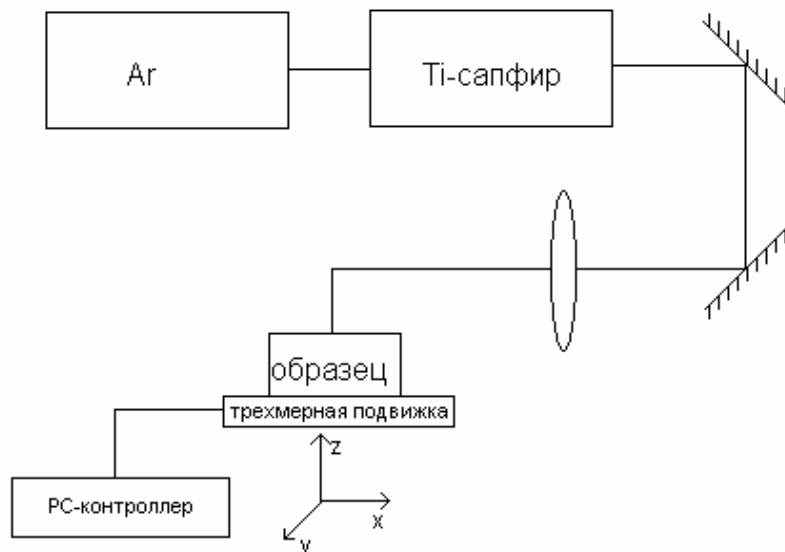


Рис. 1. Блок-схема установки для создания волноводных структур

Целью работы было создание волноводных структур в объемных образцах аморфных полупроводников: матрицы системы $As-S$ и $Ga-Ge-S:Er^{3+}$. Такие материалы обладают высоким показателем преломления и склонностью к фотоиндуцированным изменениям.

Волноводные структуры в объемных образцах As_2S_3 и $Ga-Ge-S:Er^{3+}$ были созданы при различных геометриях записи. На рис. 2, а представлена волноводная структура, созданная со скоростью 1 мм/мин при параллельном перемещении образца относительно лазерного излучения. Ее ширина составляет 12 мкм. При перпендикулярной записи не удалось создать однородную модифицированную область. На рис. 2, б видно, что в результате лазерного воздействия образовалась последовательность сферических структур. Высокая частота повторения лазерных импульсов, достаточно маленькая скорость перемещения образца обеспечивает локальное плавление вокруг фокальной

точки. Сферическая форма полученных структур позволяет предположить тепловой механизм структурных изменений. Ширина полученной структуры 7–14 мкм.

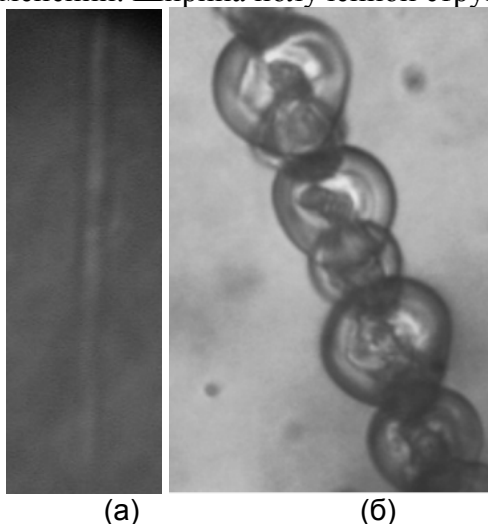


Рис. 2. Волноводная структура в системе Ga-Ge-S при параллельной (а) и перпендикулярной (б) геометрии записи

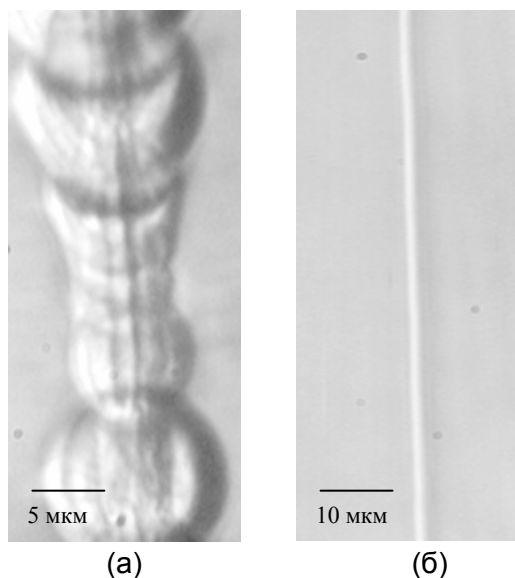


Рис. 3. Волноводная структура в системе As-S при перпендикулярной геометрии записи и скорости сканирования 3 см/мин (а) и 3 см/сек (б)

Затем было исследовано влияние скорости сканирования образца на морфологию структур при одинаковых геометриях записи и мощности излучения 250 мВт лазерного излучения. Из рис. 3а,б видно, что увеличение скорости сканирования дает возможность создавать качественные волноводные структуры меньшим диаметром.

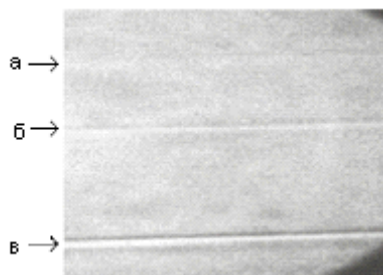


Рис. 4. Волноводная структура в системе As-S при скорости 95 см/сек и количестве сканирований: а) 75, б) 150, в) 300

Затем было исследовано влияния количества сканирований лазерного излучения на морфологию структур при мощности 150 мВт и скорости сканирования 95 см/сек. На рис. 4 видно, что при увеличении числа сканирований заметно увеличивается контрастность полученных структур, а их ширина остается постоянной 5 мкм.

Заключение

Были созданы волноводные структуры в образцах аморфных полупроводников As_2S_3 и Ga-Ge-S:Er^{3+} . В результате работы было исследовано влияние геометрии записи на морфологию полученных структур. Также были созданы волноводные структуры в ХГС As_2S_3 при различных скоростях сканирования. Увеличение скорости перемещения образца делает возможным создание ровных волноводных структур. Была получена зависимость влияния количества сканирований на морфологию структур.

Получены зависимости морфологии волноводных структур от условий лазерного воздействия (мощность лазерного излучения, скорость сканирования, число сканирований)

Литература

1. L. Tong, R.R. Gattas, I. Maxwell, J.B. Ashcom, E. Mazur. Optical loss measurements in femtosecond laser written waveguides in glass. // Optics Communications. 2006. № 259. P.626–630.
2. A. Povolotskiy, A. Shimko, A. Manshina. 2D and 3D laser writing for integrated optical elements creation. // Proceedings of WFOPC2005. Mondello (Palermo) Italy. Eliotici-nese Service Point srl. 2005. №4. P.196–202.
3. Ya. Cheng, K. Sugioka, K. Midorikawa. Microfabrication of 3D hollow structures embedded in glass by femtosecond laser for Lab-on-a-chip applications. // Applied Surface Science. 2005. № 248. P.172–176.
4. H. Hosono, K. Kawamura, S. Matsuishi, M. Hirano. Micro and nano-machining of transparent dielectrics by interfered two short laser pulses. // Nuclear Instruments and Methods in Physics Research. 2002. № 191. P.89–97.

ПРИМЕНЕНИЕ МЕТОДА НАНОИМПРИНТА ДЛЯ ЕДИНИЧНОГО КОПИРОВАНИЯ ПОЛИМЕРНОЙ ФРЕНЕЛЕВСКОЙ И МИКРО-ОПТИКИ

Н.Н. Арёфьева

Научный руководитель – д.ф.-м.н., с.н.с. И.Ю. Денисюк

Исследовано применение метода наноимпринта для единичного копирования полимерной френелевской и микрооптики. Определены условия проведения процесса, обеспечивающие наилучшую точность передачи мелких деталей, в частности, вакуумирование силоксана и УФ-отверждаемого акрилата при затекании в матрицу. Проведена оценка точности передачи формы оптической поверхности при копировании.

Введение

Технология изготовления полимерной френелевской оптики известна и базируется на использовании процесса горячего прессования термопластиков, таких как полиметилметакрилат и поликарбонат. Технология предполагает изготовление первичного оригинала линзы методом токарной обработки ПММА. Затем на оригинал напыляется в вакууме металл, далее металл гальванически наращивается до толщины 0,5–1 мм, и затем металлическая «зеркальная» матрица используется для горячего прессования [1].

В последнее время при использовании светодиодов в железнодорожных светофорах и необходимости обеспечения высокой яркости были разработаны и изготовлены френелевские линзы для получения излучения в конусе 1,5 градусов. При испытании линз оказалось, что одна центральная индикатриса сопровождается появлением шести паразитных индикатрис по шести углам шестиугольной линзы. В результате потери света на паразитных индикатрисах оказались значительны, и это существенно ухудшило энергетiku светофора. Таким образом, при уменьшении ширины индикатрисы до 1,5 градуса разработчики столкнулись с физическими ограничениями процесса горячего прессования.

В настоящей работе проведено исследование, направленное на поиск альтернативных методов изготовления френелевской оптики, свободных от вышеуказанных недостатков и обеспечивающих возможность изготовления единичных копий.

Метод наноимпринта

В работе в качестве основного оптического материала используются УФ-полимеризуемые композиции [2]. УФ-отверждаемые композиции – это новый тип материала, разработанный за рубежом в начале 90-х годов. Композиция обычно состоит из олигомера, мономера и фотоинициатора. Преимуществами УФ-полимеризуемых мономерных композиций являются: быстрое отверждение, обеспечивающее высокую производительность, отсутствие растворителей и выделяющихся летучих органических соединений.

Использованные компоненты УФ-отверждаемой композиции:

1. феноксипропилакрилат (мономер);
2. бисфенол А глицеролат диакрилат (бифункциональный мономер, поперечно-сшивающий агент);
3. диметоксифенилацетофенон (инициатор).

Основной метод изготовления френелевской линзы, использованный в работе – метод наноимпринта, основанный на процессе двукратного копирования оригинала с получением промежуточной «зеркальной» копии на ПДМС гибком штампе и второй копии на УФ-акрилате. Метод наноимпринта литографии предложен в 1995 г. По существу наноимпринт заключается в передаче топологической формы требуемой детали

от образца к копии [3, 4]. Относительная дешевизна метода вместе с возможностью получения пространственного разрешения $\approx 1,5$ нм делают его перспективным для копирования необходимой детали с большой точностью и минимальными затратами. Низкая адгезия отвержденного ПДМС обеспечивает возможность использования оригинала линзы, изготовленного на полиметилметакрилате методом токарной обработки. Поскольку ПДМС легко снимается с оригинала линзы без повреждения ее поверхности, оригинал можно использовать многократно, и получаемые копии будут идентичны друг другу.

Таким образом, для изготовления небольшой или штучной партии деталей целесообразно применить технологию наноимпринта.

Для эксперимента в качестве образца была использована линза Френеля, изготовленная методом литья под давлением поликарбоната по стандартной технологии. Суть эксперимента заключается в создании копии данной линзы с помощью метода наноимпринта и измерении ее оптических характеристик.

Технологический процесс

В качестве полимера для изготовления штампа использовался силикоксан Т-4 [5]. Этот полимер по своим свойствам наиболее подходит для наноимпринта и подобен силикоксанам, используемым в зарубежных работах. Силоксан Т-4 легко отделяется от любой поверхности, не повреждая ее, поэтому форма поверхности передается с большой точностью и процесс копирования может быть повторен многократно.

Изготовление гибкого штампа. На рис. 1 в схематичном виде представлен процесс изготовления гибкого штампа, который заключается в заполнении формы поверхности образца силикоксаном Т-4 с последующим обезгаживанием в вакууме.

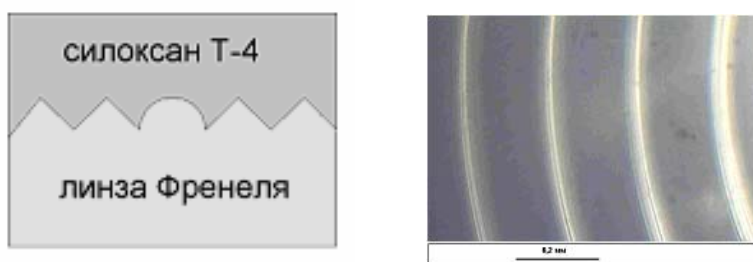


Рис. 1. Заливка силикоксана Т-4 на линзу Френеля (оригинал).
Справа – фотография поверхности оригинала

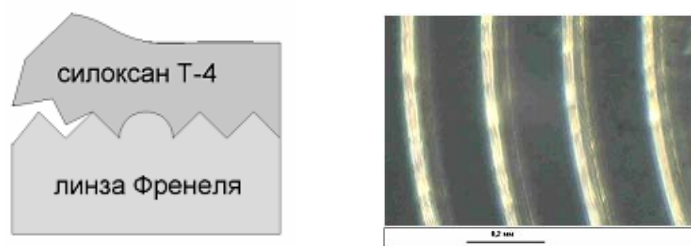


Рис. 2. Снятие гибкого штампа. Справа – фотография поверхности гибкого штампа

Проведенные эксперименты показали необходимость проведения обезгаживания силикоксана, залитого на матрицу. При отсутствии обезгаживания в вакууме в силикоксане остаются пузырьки воздуха, а также, вероятно, и растворенный воздух. В результате мелкие канавки и острые углы не прорабатываются. Напротив, при обезгаживании в вакууме 10^{-3} мм.рт.ст. в течение 30–40 минут жидкий силикоксан кипит, что свидетельствует о выходе воздуха. После этого штамп полимеризуется в течение 18–24 часов и, как показано на рис. 2, отделяется от образца. Рассматривание отвердевшей силикоксановой

копии в микроскоп показывает, что в этом случае обеспечивается затекание силикона во все мелкие элементы оригинала, а форма передается правильно.

Изготовление акриловой копии. На рис. 3 и рис. 4 приведена последовательность операций при получении акриловой копии с гибкого штампа.



Рис. 3. Заполнение гибкого штампа жидкой акриловой композицией и ее полимеризация УФ-излучением

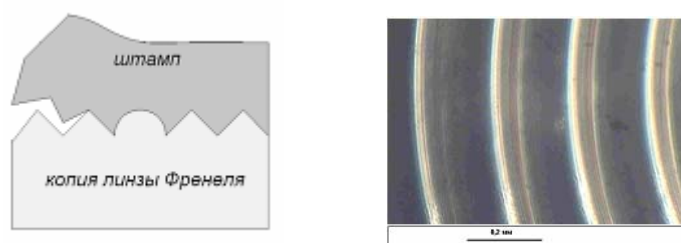


Рис. 4. Снятие гибкого штампа с готовой линзы.
Справа – фотография поверхности копии линзы (акрилат)

После заполнения объема штампа акриловой композицией в течение 45 минут акрилат обезгаживали в вакууме, чтобы уменьшить возникновение пузырей воздуха в копии. Кроме того, силикопан прозрачен для кислорода. В результате в работе обнаружено, что без вакуумирования акрилат вовсе не полимеризуется на силикопане, что может быть объяснено следующим образом. Известно, что УФ-отверждаемые акрилаты ингибируются кислородом воздуха, и на воздухе реакция полимеризации не проходит [6]. Обычно для УФ-полимеризации без использования вакуума или аргона процесс проводят в закрытом объеме. Здесь силикопан мог бы являться одной из поверхностей этого объема. Однако оказалось, что силикопан прозрачен для кислорода, а его емкость по кислороду столь значительна, что УФ-отверждение акрилата, контактирующего с силикопаном Т-4, не начинается даже через 30 минут после начала экспонирования. Только вакуумирование или напуск аргона позволяет начаться процессу фотополимеризации.

Штамп из силикона оптически прозрачен на длине волны более 280 нм, поэтому возможна полимеризация УФ-излучением при экспонировании через него. Реакцию полимеризации обеспечивает фотоинициатор, входящий в состав акриловой композиции, который активизируется под действием энергии фотонов и, расщепляясь, вступает в химическую реакцию с олигомерами и мономерами, сшивая их в длинные трехмерные цепочки полимера. Сравнение спектров пропускания силикопановой композиции, через которую проводится полимеризация, спектра поглощения фотоинициатора и спектра излучения ртутной лампы показывает возможность фотоотверждения с использованием излучения линий 313 и 365 нм. В данной работе фотоотверждение проводилось на длине волны 365 нм. Для этого использовалась лампа OSRAM OSW 125 с УФ-люминофором, т.е. ее люминофор поглощал жесткий ультрафиолет и излучал его на длине волны 365 нм. Мощность УФ света 365 нм составляла 5 Вт. Результаты эксперимента приведены на рис. 5.

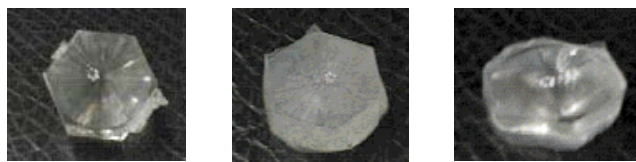


Рис. 5. Фотографии линзы оригинала (ПММА), гибкого штампа (силоксан Т-4) и копии (УФ-акрилат)

Для контроля формы поверхности гибкий штамп был разрезан и разрез сфотографирован. Как видно из таблицы, сравнение полученных углов структуры с рассчитанными на чертеже показало их идентичность с точностью 5–10 минут. Следовательно, данный метод позволяет копировать микрооптику.

N	α_N теор.	α_N практ.	N	α_N теор.	α_N практ.	N	α_N теор.	α_N практ.	N	α_N теор.	α_N практ.
1	1°20'	1°20'	14	37°40'	37°35'	27	60°20'	60°15'	40	72°40'	72°40'
2	4°	4°	15	39°50'	39°50'	28	61°10'	61°10'	41	73°20'	73°20'
3	7°30'	7°35'	16	42°	42°	29	62°45'	62°40'	42	74°	74°
4	10°50'	10°55'	17	44°10'	44°10'	30	63°50'	63°50'	43	74°40'	74°40'
5	13°30'	13°30'	18	46°10'	46°10'	31	64°55'	64°55'	44	75°17'	75°20'
6	16°30'	16°30'	19	48°10'	48°10'	32	66°10'	66°20'	45	75°55'	75°50'
7	19°30'	19°30'	20	50°	50°	33	67°	67°	46	76°33'	76°35'
8	22°30'	22°30'	21	52°10'	52°10'	34	67°50'	67°50'	47	77°12'	77°12'
9	25°30'	25°25'	22	54°	54°	35	68°45'	68°45'	48	78°05'	78°10'
10	28°10'	28°10'	23	55°20'	55°20'	36	69°40'	69°40'	49	79°10'	79°10'
11	30°30'	30°35'	24	56°30'	56°30'	37	70°40'	70°45'	50	80°05'	80°05'
12	32°50'	32°40'	25	57°50'	57°50'	38	71°20'	71°15'			
13	35°10'	35°10'	26	59°10'	59°10'	39	72°05'	72°05'			

Таблица. Сравнение углов гибкого штампа с расчетом

Заключение

С помощью метода наноимпринта была получена копия линзы Френеля. Процесс не требует применения дорогостоящего оборудования и экономически оправдан даже при изготовлении одиночных копий. В связи с отсутствием необходимости нагрева (например, метод литья под давлением) в полимерной копии линзы Френеля не возникают деформации и напряжения, а это, в свою очередь, позволяет изготавливать линзы с большой толщиной, не изменяя их оптических характеристик. В ходе эксперимента была установлена необходимость вакуумирования как силоксана, так и акриловой композиции.

Данная работа выполнялась при поддержке по гранту РФФИ № 05-02-08048.

Литература

1. Окатов А.М., Антонов Э.А., Байгожин А. и др. Справочник технолога – оптика. СПб: Политехника, 2004. 679 с.
2. <http://www.tanzor.ru/polygraphmats/>
3. Gates D. Nanofabrication with molds and stamps. // Materials today. 2005, February. P.44–49.
4. Sotomayor Torres C.M., Zankovych S. Nanoimprint lithography: an alternative nanofabrication approach. // Materials Science and Engineering. 2003. №23. P. 23–31.
5. <http://www.penta-91.ru>
6. Andrejewska E. Photopolymerization kinetics of multifunctional monomers. // Progress in polymer science. 2001. №26. P. 605–665.

ВЛИЯНИЕ ТИПА И КОНЦЕНТРАЦИИ КРАСИТЕЛЯ НА ФОРМИРОВАНИЕ МИКРОЭЛЕМЕНТА НА ТОРЦЕ ОПТИЧЕСКОГО ВОЛОКНА

А.С. Березкина

**Научный руководитель – ассистент кафедры оптики квантоворазмерных систем
М.И. Фокина**

Рассматривается процесс формирования микролинз на торце оптического волокна путем фотополимеризации мономера. Исследуется влияние красителя, введенного в состав полимера, на образование микроэлементов различных форм.

Введение

Актуальным вопросом сейчас является проблема ввода оптического излучения в волокно. При этом необходимо минимизировать потери связи оптоволоконно–источник/приемник. В настоящее время для этого применяются различные согласующие линзовые устройства, диэлектрические конусы, сферическое оплавление световода и т.д. Сферические линзы увеличивают коэффициент связи за счет увеличения апертурного угла, поскольку коэффициент связи пропорционален квадрату числовой апертуры. Эффективность ввода в таком случае может быть увеличена примерно в три раза [1]. Иногда применяются отражательные и гиперболические линзы, дифракционные линзы, фоконы. Однако применяемые линзовые устройства вносят большие потери в систему источник-линза-волокно [2] – так называемые конструктивные потери. К тому же такие системы требуют очень точной оптической юстировки, что не всегда является удобным.

Для того чтобы сделать ввод излучения наиболее эффективным, необходимо, чтобы элемент, фокусирующий оптическое излучение, был идеально согласован с волокном. В последнее время эта проблема решается формированием фокусирующей линзы непосредственно на торце оптоволоконка.

Известен целый ряд методов, позволяющих формировать микролинзы на торцах оптических волокон: непосредственный расплав материала волокна [3], формирование дифракционных микроэлементов на торце волокна с помощью электронного луча из фоторезиста [4], создание фоконов в самом волокне за счет распределения показателя преломления [5], формирование микроэлементов с использованием излучения выходящего непосредственно из оптического волокна из [6, 7]. Использование излучения, выходящего из торца оптического волокна, позволяет получать широкий спектр форм элементов – от цилиндрических (являющихся, по сути, продолжением оптической жилы волокна) [8] до квазисферических и фоконовых [6, 7]. При этом форма элемента в основном определяется составом полимерной композиции и фотоактивными добавками.

В данной работе рассматривается влияние фотоактивных добавок (Родамина В и Кумарина 120) в составе полимерной композиции на форму и скорость развития микроэлемента на торце оптического волокна.

Формирование микроэлементов на торце оптоволоконка

Для формирования микролинз на торце волокна была использована схема установки, изображенная на рис. 1. В волокно вводилось излучение азотного лазера, генерирующего излучение с длиной волны $\lambda = 337$ нм, со средней выходной мощностью 3,2 мВт, длительностью импульса излучения 10 нс и частотой следования импульса 100 Гц.

Кинетика процесса регистрировалась ПЗС-камерой, работающей в режиме видеозаписи с частотой 50 кадров в секунду. Волокно – кварцевое, многомодовое, ступенчатое, диаметром 435 мкм.

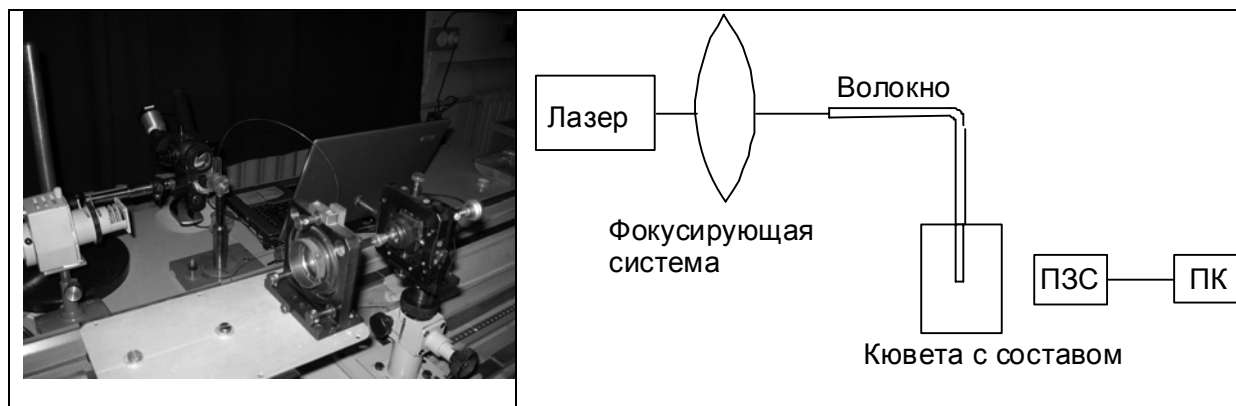


Рис. 1. Фотография и принципиальная схема установки

Свободный торец волокна с выходящим из него светом опускался в УФ-отверждаемую мономерную композицию. Под действием УФ-излучения, генерируемого лазером, происходит фотополимеризация состава, вследствие чего формируется микроэлемент на центральной жиле оптоволокна. Время экспозиции для каждого образца было строго зафиксировано и составляло 10 секунд.

Формирование микроэлемента на торце волокна с использованием чистого полимера (в отсутствии красителей)

При времени облучения порядка 10 секунд происходит формирование оптической полимерной структуры. На рис. 2 показана фотография микроструктуры, полученной данным способом. Диаметр ее совпадает с диаметром световодной жилы волокна – 435 мкм, длина микроструктуры порядка 3 мм.

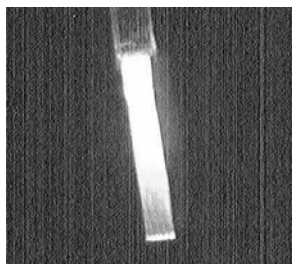


Рис. 2. Микроэлемент, полученный на торце волокна с использованием чистой мономерной композиции

Полимеризующийся материал, имеющий показатель преломления выше окружающего неполимеризованного мономера примерно на 0,4, играет роль своеобразного волновода для света, выходящего из торца оптоволокна. В результате на торце оптоволокна быстро растет цилиндр из полимеризованного материала, по диаметру равный центральной жиле оптоволокна. Данный процесс роста неограничен и, при отсутствии поглощения света в полимеризованном материале, мог бы идти бесконечно с образованием длинного цилиндра.

Зависимость формы микроэлемента на волокне от концентрации красителя в полимере

Эксперимент с родамином В. Введение в мономерную композицию компонента, обеспечивающего абсорбцию света, позволяет получать микроэлементы различной формы, такие как цилиндры, конуса, микролинзы в зависимости от концентрации фотоактивного компонента в полимере. В качестве такого компонента мы взяли краситель родамин В. Для исследования были получены мономерные композиции с содержанием родамина В 0,1, 0,2, 0,3, 0,4, 0,5 и 0,6 весовых процентов.

Композицию экспонировали строго определенное время – 10 секунд. Постоянство времени экспозиции позволило четко проследить зависимость скорости роста элемента от концентрации красителя в составе полимера. Фотографии полученных микроэлементов представлены на рис. 3.

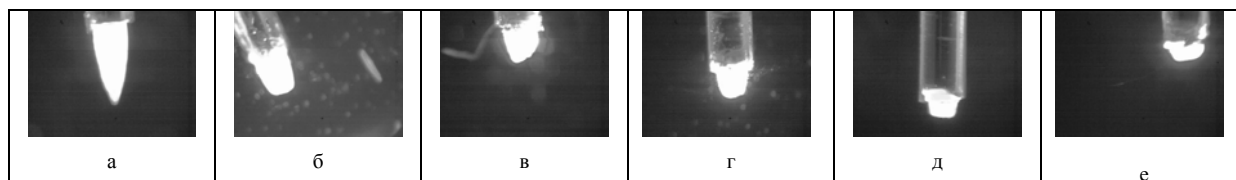


Рис. 3. Прохождение света через микроэлементы полученные при разных концентрациях Родамина В (концентрация родамина В: а – 0,1 wt; б – 0,2wt; в – 0,3 wt; г – 0,4 wt; д – 0,5 wt; е – 0,6 wt)

Для определения более четкой зависимости формы линзы от концентрации красителя были проведены измерения размеров полученных микроэлементов. По полученным данным построен график зависимости размера микроэлемента вдоль оптической оси на оптоволокне от концентрации родамина В (рис. 4).

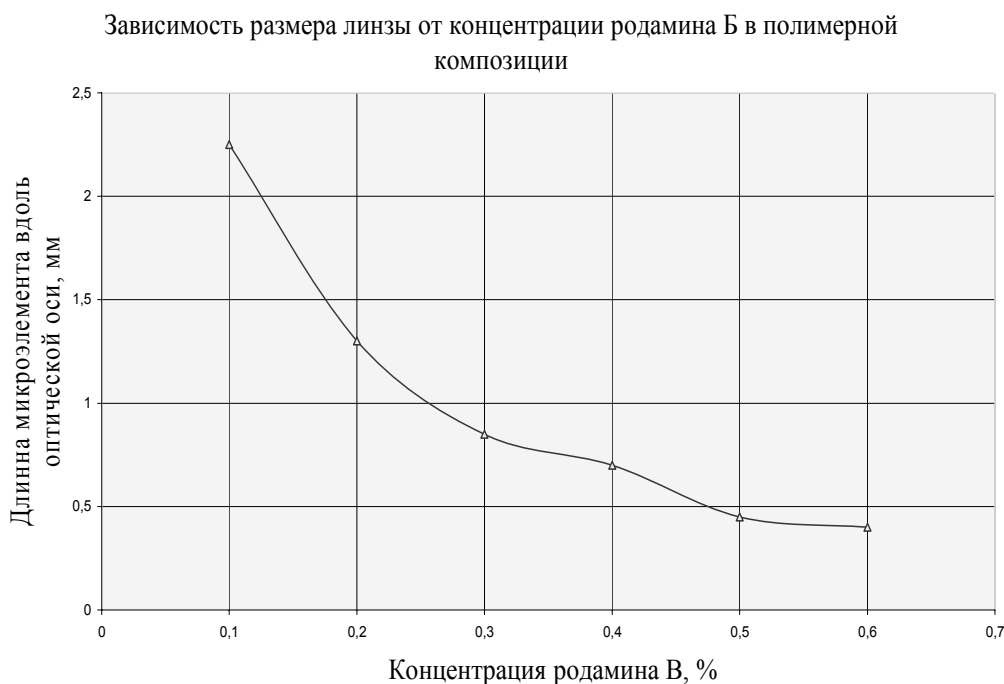


Рис. 4. График зависимости длины микроэлемента вдоль оптической оси от концентрации родамина В в мономерной композиции

Как видно из графика, чем выше концентрация красителя родамина В в составе мономерной композиции, тем меньше размер линзы вдоль оптической оси, сформированной на торце оптоволокна в результате фотополимеризации. В процессе фотополимеризации излучение, выходящее из световода, начинает поглощаться красителем. Таким образом, процесс роста начинает замедляться, поэтому при максимальной в данной

работе концентрации родамина В, мы наблюдаем минимальный размер линзы и специфическую форму. При минимальной концентрации поглощение минимально, следовательно, свет достаточно легко проникает в полимер, и образуется конусообразная микролинза.

Эксперимент с кумарином 120. Также были проведены эксперименты по выращиванию элементов на торце оптоволокна с использованием идентичной мономерной композиции, но с другой фотоактивной добавкой – кумарином 120. Время экспозиции осталось тем же – 10 секунд. Кумарин 120 имеет большее поглощение на длине волны 337 нм, поэтому концентрации кумарина 120, которые использовались в наших исследованиях, на порядок ниже, чем концентрации родамина В.

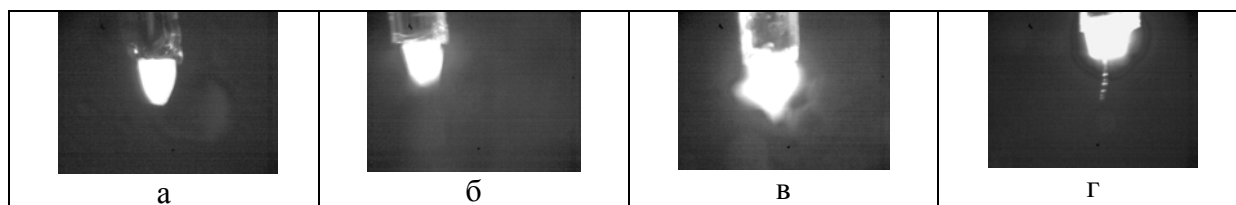


Рис. 5. Микроэлементы, полученные на торце оптоволокна при различных концентрациях красителя кумарина в составе полимера (концентрация кумарина 120: а – 0,03 wt, б – 0,05 wt, в – 0,08 wt, г – 0,1 wt, соответственно)

Как можно заметить, в данном случае происходит формирование микроструктуры, с одной стороны являющейся продолжением оптической жилы волокна, а с другой стороны с явным округлым завершением (рис. 5).

Обсуждение полученных результатов

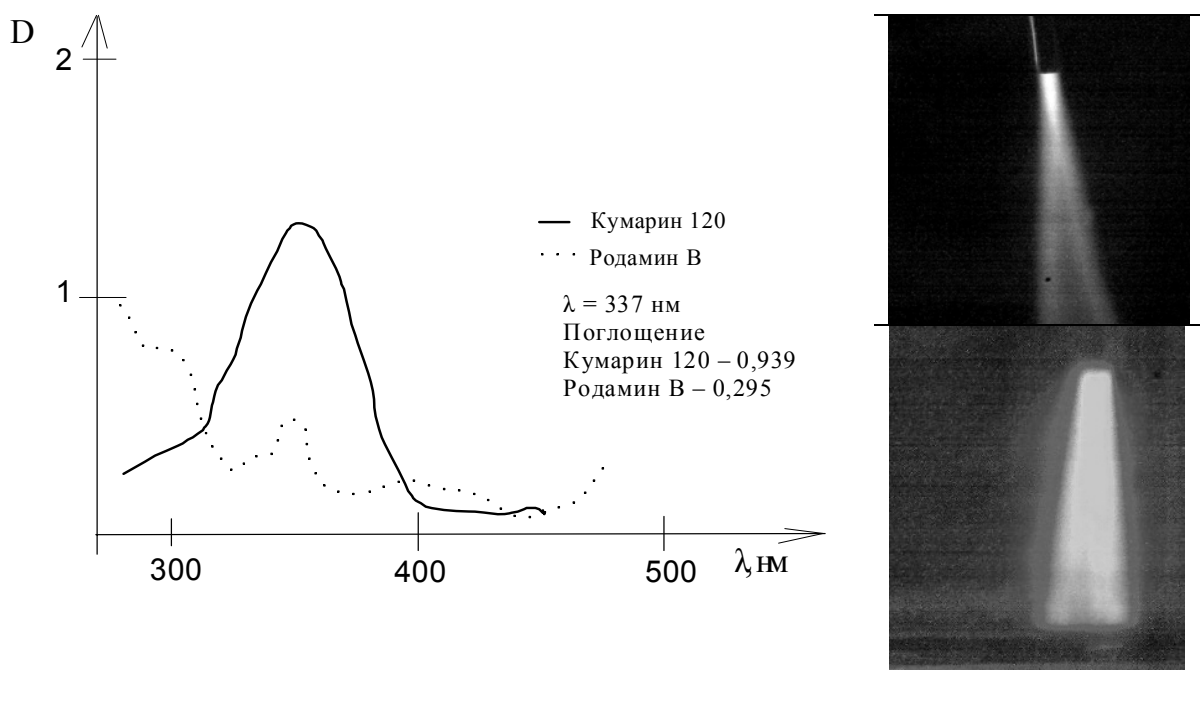


Рис. 6. Спектры поглощения фотоактивных добавок родамина В и кумарина 120 в изопропанол в концентрациях 0,1 wt и 0,05 wt, соответственно, распределение излучения при выходе из волокна в среду с родамином В (справа сверху) и в среду с кумарином 120 (справа внизу)

На рис. 6 приведены спектры поглощения кумарина 120 и родамина В. Из спектров видно, что в первую очередь данные фотоактивные добавки различаются расположением максимумов поглощения. Если основной максимум поглощения родамина В

расположен в районе 500–600 нм, и на рабочей длине волны 337 нм он имеет относительно небольшое, хотя и ощутимое поглощение, то кумарин 120 имеет максимум поглощения как раз в районе 300–400 нм и на рабочей длине волны имеет огромное поглощение – это объясняет необходимость значительного уменьшения его концентрации (на порядок меньше относительно родамина В). Еще одно существенное различие хорошо видно на фотографиях (рис. 6) распределения излучения, выходящего из волокна в среду с родамином В и кумарином 120.

Данное распределение излучения хорошо согласуется с полученными микроэлементами. В случае с родамином В мы видим очертания четкого конуса максимальной интенсивности, который впоследствии повторяется полученным микроэлементом (рис. 2, а). Распределение же интенсивности света, выходящего в раствор кумарина 120, близко к распределению излучения, выходящего из волокна в среду, не имеющую специфического поглощения. И, соответственно, образующиеся микроэлементы близки по форме к микроструктурам, получаемым из полимерной композиции при отсутствии фотоактивной добавки, но значительно меньше их по размеру, что объясняется значительным поглощением кумарина 120 в области рабочей длины волны, т.е. проявляется эффект торможения процесса.

Данная работа выполнена при поддержке по проекту Рособразование РНП.2.1.1.1403 «Исследование процессов формирования микрооптических поверхностей в поле световой волны при фотоотверждении мономерных композиций».

Литература

1. Иволгин В., Коханенко А., Мягков А. Волоконно-оптические линии связи: эффективность ввода оптического излучения в волокно, Электронная версия учебного пособия. Томский государственный университет, кафедра КЭиФ, 2002.
2. Окоси Т. Волоконно-оптические датчики. Л.: Энергоатомиздат, 1991. 108 с.
3. Вейко В.П., Березин Ю.Д. и др. Лазерные технологии формирования волоконно-оптических инструментов. // Известия РАН. Серия физическая. 1997. Т. 61. № 8. С. 1627–1631.
4. F. Schiappelli a, R. Kumar et al. Efficient fiber-to-waveguide coupling by a lens on the end of the optical fiber fabricated by focused ion beam milling. // Microelectronic Engineering. 2004. P.73–74. 397–404.
5. Yi Yang, Jon Lee, Karl Reichard, et al. Fabrication and implementation of a multi-to-single mode converter based on tapered multimode fiber. // Optics Communication. 2005. 249. P.129–137.
6. Hocine M., Bachelot R. et al. End-of-fiber polymer tip: manufacturing and modeling. // Synthetic Metals. 2002. 127. P.313–318.
7. Плеханов А.И., Шелковников В.В. Оптические волокна с концевыми полимерными микролинзами. // Российские нанотехнологии. 2006. №1.
8. K. Yamashita, T. Kuro, K. Oe Fabrication of self-writing wavrguide in photosensitive polyimide resin by controlling photochemical reaction of photosensitizer. // App. Phys. Lett. 2004. V. 85. № 18. P.3962–3964.

ГЕОМЕТРИЧЕСКИЕ И ОПТИЧЕСКИЕ СВОЙСТВА АФОКАЛЬНОЙ ДВУХЗЕРКАЛЬНОЙ СИСТЕМЫ

В.И. Батшев

Научный руководитель – д.т.н., профессор Д.Т. Пуряев
(Московский государственный технический институт им. Н.Э. Баумана)

Установлена связь между математическими выражениями продольных аберраций нормалей поверхностей, образующих афокальную двухзеркальную систему (АДС), исправленную на сферическую аберрацию для осевого пучка лучей (так называемую систему Пуряева, защищенную авторским свидетельством № 1527607). Разработан алгоритм для расчета наклонных меридиональных пучков лучей, идущих через АДС.

1. Введение

АДС широко применяются в лазерной технике и в оптическом приборостроении. Среди возможных сфер применения можно назвать следующие: астрономия, лазерная локация, оптическая связь, интерферометрия и др. Одна из таких систем, состоящая из двух параболических зеркал, известна под названием системы Мерсена [1]. Она полностью свободна от сферической аберрации при любых апертурах зеркал, в области аберраций третьего порядка отсутствуют кома и астигматизм, но кривизна поля принципиально неустранима. Система Пуряева [2], у которой одно из зеркал сферическое, а другое – асферическое, строго эквидистантное мнимому параболоиду, также является АДС с полностью исправленной сферической аберрацией и имеет над системой Мерсена важные технологические преимущества.

Но и система Мерсена, и система Пуряева являются лишь частными случаями глобального решения АДС, полученного профессором Д.Т. Пуряевым [3] в 1988 г. Оно представляет собой соотношения, позволяющие по известному параметрическому уравнению одной из поверхностей системы определить параметрическое уравнение другой поверхности при условии выполнения принципа Ферма для осевого пучка лучей. В качестве параметра в уравнениях поверхностей выступает угол наклона нормали.

Эти соотношения были получены сравнительно недавно, и с тех пор было проведено не так много исследований, направленных на изучение свойств АДС. А изучение свойств – это путь к открытию новых ее возможностей. На это и направлена данная работа; в ней установлено новое общее для всех вариантов АДС ее геометрическое свойство, показывающее связь между математическими выражениями продольных аберраций нормалей поверхностей, образующих АДС. Получена интерпретация глобального решения, дающая новый способ построения поверхности АДС при известном уравнении другой поверхности. Разработан алгоритм для исследования оптических свойств различных вариантов АДС, позволяющий производить расчет наклонных меридиональных пучков и принципиально отличающийся от уже существующих алгоритмов тем, что для расчета хода лучей поверхности могут быть заданы в параметрическом виде.

2. Геометрические свойства афокальной двухзеркальной системы

2.1. Объект исследования

Рассмотрим АДС (рис. 1), состоящую из малого зеркала oA и большого OB . Используем две системы координат: zoU для зеркала oA и ZOY для зеркала OB . Оси oz и OZ совмещены с оптической осью системы. Точки E и C являются центрами кривизны при вершинах большого и малого зеркал, соответственно; φ – параметр (угол наклона нормали к поверхности). Расстояние между зеркалами равно d .

На схеме (рис. 1) изображена так называемая предфокальная АДС; параксиальные радиусы кривизны поверхностей ($r_1 = r$ и $r_2 = R$) имеют одинаковые знаки (оба поло-

жительно), причем $R > r$. Это лишь один из возможных вариантов конструктивного исполнения системы. Приведенные в этом разделе геометрические соотношения и выводы справедливы для этого вида АДС; для других вариантов могут быть получены аналогичные по структуре и сути результаты.

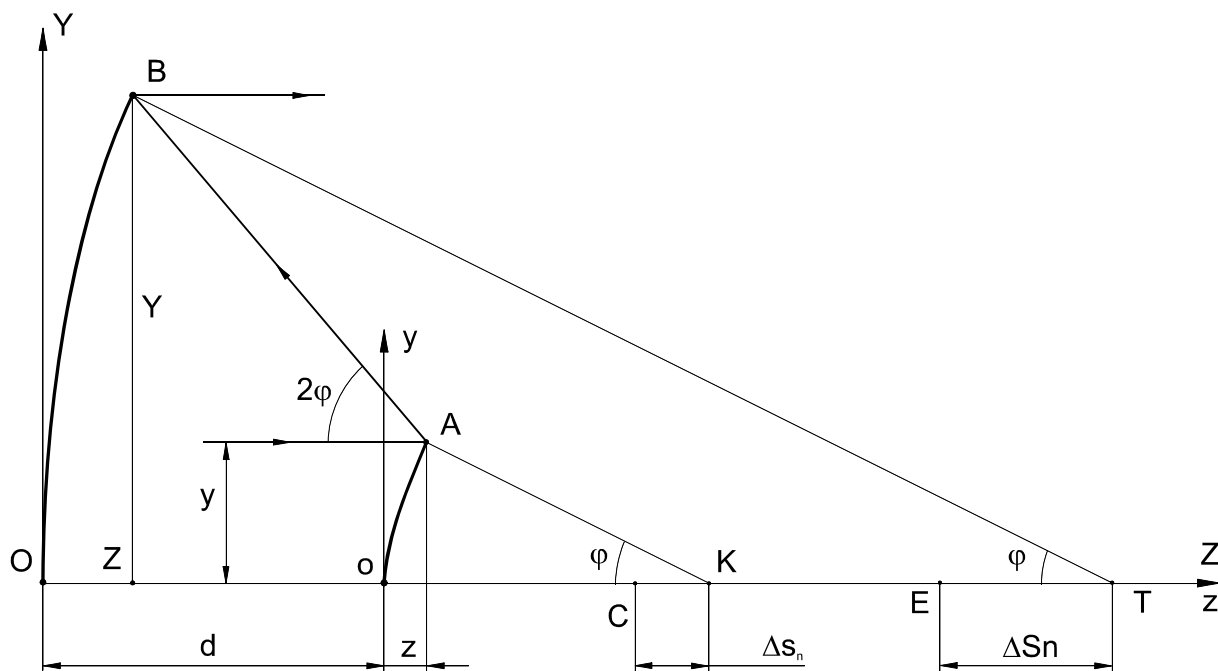


Рис. 1. Предфокальная АДС

В начале рассуждений следует привести два важных соотношения, справедливые для всех вариантов АДС, полностью свободной от сферической aberrации для осевого пучка лучей. Из условия афокальности системы, т.е. совмещения фокусов зеркал, следует параксиальное соотношение, позволяющее рассчитать осевое расстояние между вершинами зеркал:

$$d = \frac{R - r}{2}, \quad (1)$$

где R и r – радиусы кривизны при вершинах большого и малого зеркал, соответственно, d – осевое расстояние между поверхностями.

Глобальное решение АДС, связывающее между собой параметрические уравнения поверхностей, записывается в виде

$$\begin{cases} Z(\varphi) - z(\varphi) = d \cdot \operatorname{tg}^2 \varphi \\ Y(\varphi) - y(\varphi) = 2d \cdot \operatorname{tg} \varphi \end{cases} \quad (2)$$

где (Z, Y) – координаты на поверхности большого зеркала в системе координат ZOY , (z, y) – координаты на поверхности малого зеркала в системе координат $zoу$ (см. рис. 1), φ – параметр (угол наклона нормали к поверхности).

Следующие рассуждения основаны на этих соотношениях и являются продолжением предыдущих исследований АДС.

2.2. Связь между продольными aberrациями нормалей поверхностей АДС

Система Мерсена состоит из двух зеркал параболической формы, следовательно, продольные aberrации нормалей ее поверхностей записываются в виде:

$$\Delta S_n(\varphi) = \frac{R}{2} \cdot \operatorname{tg}^2 \varphi \quad (3)$$

для большого зеркала и

$$\Delta s_n(\varphi) = \frac{r}{2} \cdot \operatorname{tg}^2 \varphi \quad (4)$$

для малого. Если вычесть из уравнения (3) уравнение (4), то получится соотношение:

$$\Delta S_n(\varphi) - \Delta s_n(\varphi) = \frac{R-r}{2} \cdot \operatorname{tg}^2 \varphi,$$

которое с учетом формулы (1) принимает вид

$$\Delta S_n(\varphi) - \Delta s_n(\varphi) = d \cdot \operatorname{tg}^2 \varphi. \quad (5)$$

Известно, что в системе Пурьева, у которой одно из зеркал (например, малое) – сферическое, другое (большое) зеркало представляет собой так называемую эквипараболическую поверхность и имеет продольную aberrацию нормалей, описываемую формулой [2]:

$$\Delta S_n(\varphi) = d \cdot \operatorname{tg}^2 \varphi.$$

Учитывая, что сфера имеет нулевую aberrацию нормалей $\Delta s_n(\varphi) = 0$, можно записать:

$$\Delta S_n(\varphi) - \Delta s_n(\varphi) = d \cdot \operatorname{tg}^2 \varphi. \quad (6)$$

Видно, что выражение (6) идентично выражению (5). Таким образом, в двух различных вариантах АДС наблюдается одна и та же математическая зависимость между продольными aberrациями нормалей зеркал. Логично задаться вопросом: «Только ли для двух вышеуказанных систем справедлива такая зависимость?» Чтобы ответить на этот вопрос, вспомним, что уравнения поверхностей АДС, полностью свободной от сферической aberrации осевого пучка лучей, удовлетворяют глобальному решению (2).

Параметрические уравнения меридиональных профилей поверхностей большого и малого зеркал записываются в виде:

$$\begin{cases} Z(\varphi) = R - F(\varphi) \\ Y(\varphi) = [F(\varphi) + \Delta S_n(\varphi)] \cdot \operatorname{tg} \varphi \end{cases} \quad (7)$$

и

$$\begin{cases} z(\varphi) = r - f(\varphi) \\ y(\varphi) = [f(\varphi) + \Delta s_n(\varphi)] \cdot \operatorname{tg} \varphi \end{cases} \quad (8)$$

где $F(\varphi)$ и $f(\varphi)$ – некоторые функции параметра φ .

Подставив в (2) уравнения (7) и (8), получим:

$$\begin{cases} R - r - [F(\varphi) - f(\varphi)] = d \cdot \operatorname{tg}^2 \varphi \\ F(\varphi) - f(\varphi) + \Delta S_n(\varphi) - \Delta s_n(\varphi) = 2d \end{cases}$$

После простых тригонометрических преобразований, учитывая соотношение (1), получим

$$\Delta S_n(\varphi) - \Delta s_n(\varphi) = d \cdot \operatorname{tg}^2 \varphi. \quad (9)$$

Видно, что выражения (5), (6) и (9) идентичны.

Таким образом, полученное для систем Пурьева и Мерсена соотношение, устанавливающее связь между математическими выражениями продольных aberrаций нормалей поверхностей, является общим для любых вариантов АДС. Это соотношение может быть использовано, например, для синтеза АДС с определенными свойствами, связанными с aberrацией нормалей одной из поверхностей.

2.3. Синтез АДС с непараметрическими поверхностями

Рассмотрим опять систему уравнений (2). Возведем обе части нижнего уравнения в квадрат и разделим его на верхнее. В результате получим соотношение

$$\frac{[Y(\varphi) - y(\varphi)]^2}{[Z(\varphi) - z(\varphi)]} = 4d.$$

Для удобства опустим параметр φ и перепишем это соотношение в виде:

$$(Y - y)^2 = 4 \cdot d \cdot (Z - z). \quad (10)$$

Выражение (10), если переменными считать величины Z и Y , есть уравнение параболы с фокусным расстоянием d , вершина которой смещена из начала координат в точку с координатами (z, y) . На основании этого можно сделать предположение, что кривая, описываемая уравнениями $Z = Z(\varphi)$ и $Y = Y(\varphi)$, есть огибающая семейства вышеупомянутых парабол. Чтобы доказать это предположение, найдем по известной методике [4] уравнение огибающей семейства парабол (10).

Пусть известно параметрическое уравнение меридионального профиля малого зеркала:

$$\begin{cases} z = z(\varphi) \\ y = y(\varphi) \end{cases} \quad (11)$$

Перепишем уравнение (10) в виде

$$[y(\varphi) - Y]^2 = -4d \cdot [z(\varphi) - Z]. \quad (12)$$

Продифференцируем уравнение (12) по параметру φ :

$$2[y(\varphi) - Y] \cdot \frac{dy(\varphi)}{d\varphi} = -4d \cdot \frac{dz(\varphi)}{d\varphi}. \quad (13)$$

Известно, что угол наклона нормали φ к кривой (11) в соответствии с правилом знаков определяется из соотношения:

$$\operatorname{tg} \varphi = \frac{dz}{dy}. \quad (14)$$

Выражение (13) с учетом (14) преобразуется к виду

$$Y - y(\varphi) = 2d \cdot \operatorname{tg} \varphi. \quad (15)$$

Из уравнений (12) и (15) следует, что

$$Z - z(\varphi) = d \cdot \operatorname{tg}^2 \varphi. \quad (16)$$

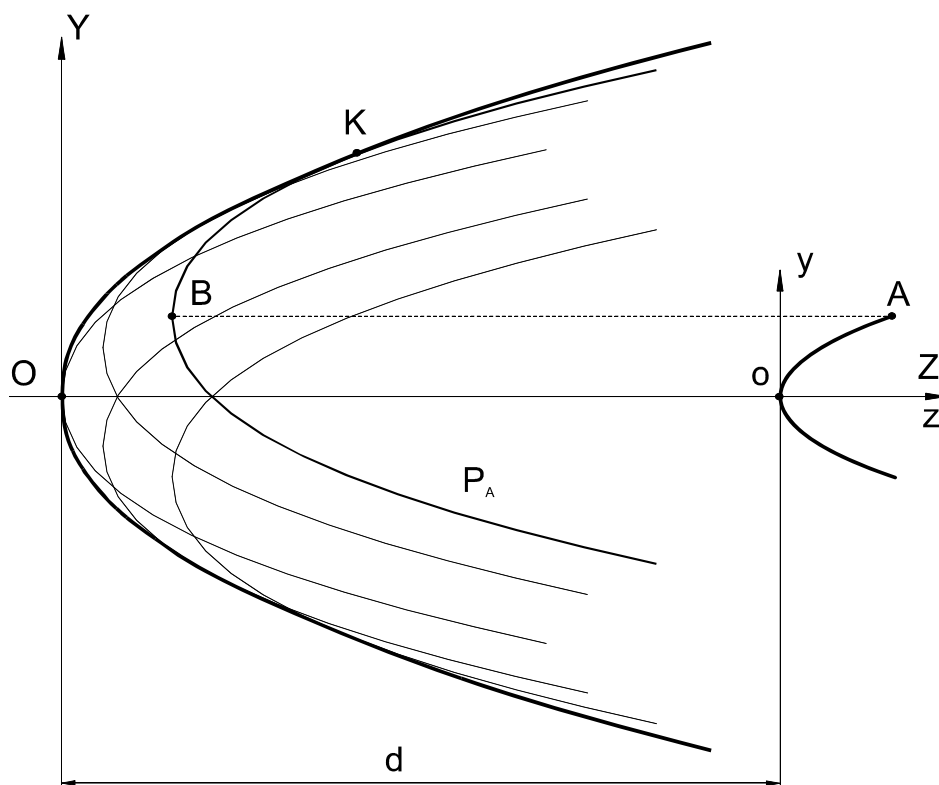


Рис. 2. К иллюстрации свойства АДС

Выражения (15) и (16), представляющие собой параметрическое уравнение огибающей семейства парабол (10), совпадают с выражениями глобального решения АДС, что доказывает правильность сделанного предположения.

Таким образом, получено новое свойство АДС.

Меридиональный профиль каждой из поверхностей АДС есть огибающая семейства парабол, фокусное расстояние которых равно осевому расстоянию между поверхностями, причем фокусы парабол лежат на меридиональном профиле другой (исходной) поверхности, а их оси параллельны оптической оси системы.

Так, на рис. 2 профиль поверхности большого зеркала ОК есть огибающая семейства парабол, одна из которых – кривая P_A – имеет вершину в точке В, а фокус – в точке А, лежащей на профиле исходного малого зеркала ОА. Фокусное расстояние параболы равно $BA = d$, а ее ось параллельна оси OZ , т.е. оптической оси АДС. Кривая P_A касается профиля большого зеркала в единственной точке – точке К.

Это свойство позволяет синтезировать АДС, если известно уравнение одной из поверхностей. Важно, что это уравнение может быть задано в любом виде, не обязательно параметрическом.

3. Алгоритм расчета АДС

В настоящее время ни в одной из известных программ расчета оптических систем, таких как *Oral*, *Prizma*, *Zemax*, *Code V*, нет возможности задать поверхность в параметрическом виде. Это существенно затрудняет анализ работы АДС в наклонных пучках лучей. Есть возможность аппроксимировать параметрически заданную поверхность асферической поверхностью стандартного вида и рассчитать систему в одной из известных программ. При этом произойдет некоторая потеря точности, которую возможно оценить; в некоторых случаях ей даже можно пренебречь. Но этот путь весьма трудоемок и не всегда приводит к удовлетворительным результатам. Поэтому приведенный ниже алгоритм расчета АДС может быть полезным инструментом для разработчика. Этот алгоритм реализован в программе *Mathcad* и позволяет производить расчет наклонных меридиональных пучков лучей через АДС, поверхности которой заданы параметрически.

Анализ системы производится в меридиональном сечении, рассчитываемая АДС задается уравнениями меридиональных профилей ее поверхностей (двумерных кривых) в единой системе координат, как показано на рис. 3. Здесь ОА и ВD – первая и вторая по ходу лучей поверхности системы. Начало координат – точка О – совпадает с вершиной первой поверхности ОА, а ось OZ – с оптической осью.

Кроме уравнений поверхностей, пользователем задается угловое поле в пространстве предметов ω , высота луча на входном зрачке Y_m , положение входного зрачка МР и плоскости анализа QN.

Положение входного зрачка МР задается относительно вершины первой поверхности отрезком $PO = -Sp$. Положение плоскости анализа NQ задается отрезком $DQ = Sa$ относительно вершины последней поверхности.

Результатом расчета луча через АДС являются высота Y_n в плоскости анализа и угол ω' с оптической осью.

На рис. 3 изображена так называемая предфокальная АДС с главным вогнутым и вторичным выпуклым зеркалами, причем $|r_2| < |r_1|$. Это лишь один из возможных вариантов конструкции системы. Алгоритм предназначен для расчета любых исполнений АДС. Расчет основан на двукратном применении закона отражения луча от поверхности. Вначале из исходных данных определяется уравнение луча, входящего в АДС. Затем находятся координаты точки встречи луча с первой поверхностью и угол падения;

эти значения дают возможность получить уравнение луча после отражения от первой поверхности. Аналогичные действия проделываются и применительно ко второй поверхности. Результатом этих действий является уравнение луча в системе координат ZOY (рис. 3) после прохождения им АДС, а из уравнения получают выходные данные.

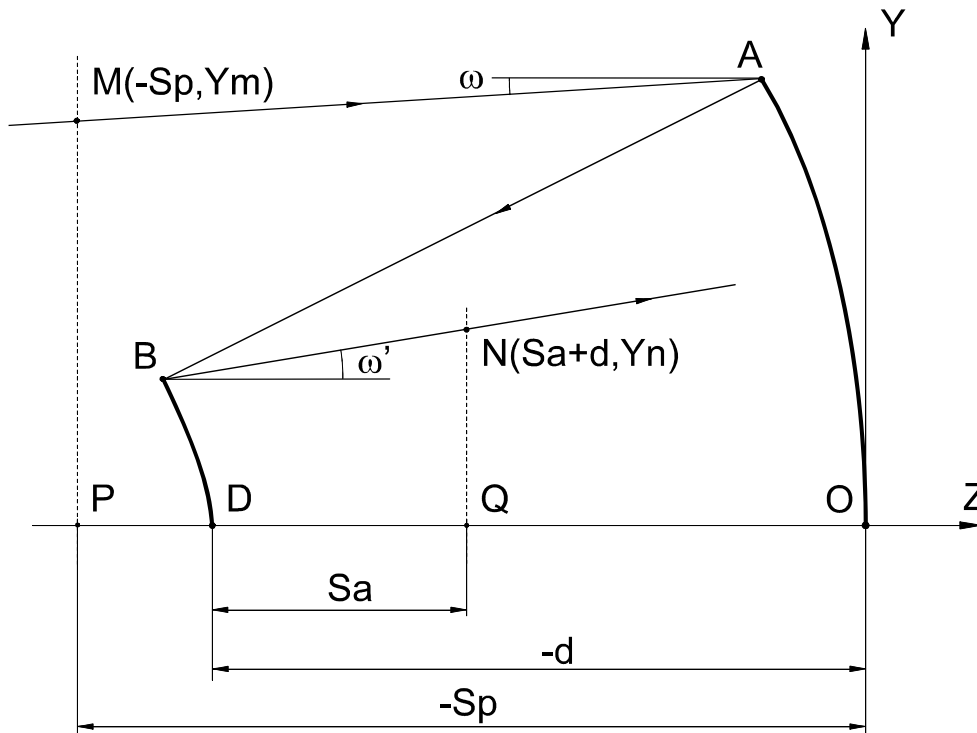


Рис. 3. Предфокальная АДС

Определение уравнения прямой в декартовой системе координат по известным координатам одной ее точки и известному углу, который она образует с одной из осей координат – задача тривиальная. Сложнее определить точку встречи луча с меридиональным профилем параметрически заданной поверхности и угол падения луча на эту поверхность. Эта процедура разделяется на два этапа.

Пусть меридиональный профиль рассматриваемой поверхности – кривая, описываемая уравнениями (7). Практически задача сводится к определению значения параметра Φ , соответствующего точке пересечения. А по известному Φ можно определить и координаты искомой точки, и угол падения. На первом этапе определятся приближенное значение параметра $\Phi_{\text{пр}}$, соответствующее точке пересечения луча с параболой, фокусное расстояние которой равно $r/2$, а вершина совпадает с вершиной рассматриваемой параметрической кривой. На рис. 4, а луч, идущий вдоль линии РА, пересекает кривую ОА в точке А; нормаль к поверхности, восстановленная в точке А (отрезок НА), образует с оптической осью OZ угол Φ . Кривая ОР есть вышеупомянутая парабола; она пересекается с лучом в точке Р. Нормаль МР к параболе ОР образует с осью OZ угол $\Phi_{\text{пр}}$.

Если парабола ОР описывается уравнением $y^2 = 2r \cdot z$, а прямая РА – уравнением $y(z) = k \cdot z + b$, то координата z точки их пересечения определяется по формуле:

$$z_p = \frac{r - k \cdot b \pm \sqrt{r^2 - 2r \cdot k \cdot b}}{k^2},$$

где знак «+» соответствует отрицательному значению r , а минус – положительному. Известно, что параметрически параболу можно описать уравнениями

$$\begin{cases} z(\varphi) = r/2 \cdot \operatorname{tg}^2 \varphi \\ y(\varphi) = r \cdot \operatorname{tg} \varphi \end{cases}$$

Так как точка Р принадлежит параболе, и ей соответствует значение параметра $\Phi_{\text{ПР}}$, то можно записать

$$z_P = r/2 \cdot \operatorname{tg}^2 \Phi_{\text{ПР}},$$

откуда и определяется величина $\Phi_{\text{ПР}}$.

Второй этап – уточнение параметра $\Phi_{\text{ПР}}$ до тех пор, пока его отличие от точного значения Φ не станет малым (каково именно должно быть это отличие, указано ниже).

Для того чтобы понять процедуру уточнения приближенного параметра, обратимся к рис. 4, б.

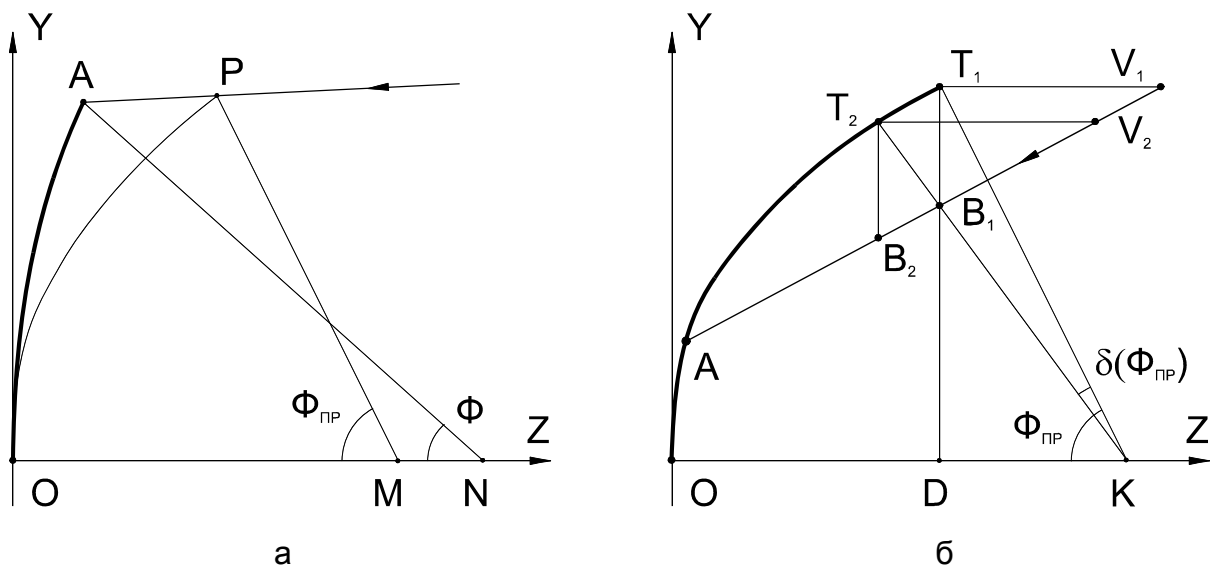


Рис. 4. а – к вопросу об определении предварительного значения параметра (парабола ОР имеет тот же параксиальный радиус, что и кривая ОА); б – процедура уточнения параметра

Луч, идущий по направлению V_1A , встречает меридиональную кривую OT_1 в точке А. Предварительно определенному значению параметра $\Phi_{\text{ПР}}$ на кривой соответствует точка T_1 . Точка V_1 принадлежит лучу и имеет ту же ординату, что и точка T_1 . Точка B_1 лежит на пересечении прямых V_1A и T_1D , прямая T_1D перпендикулярна оси OZ . Нормаль к меридиональной кривой, проходящая через точку T_1 пересекает ось OZ в точке K . Так как координаты точек T_1 , B_1 , K и D известны, то известен и угол $\angle T_1KB_1$, величина которого зависит от значения $\Phi_{\text{ПР}}$ и который обозначен через $\delta(\Phi_{\text{ПР}})$. Прямая KB_1 пересекает меридиональную кривую в точке T_2 . Если пренебречь разностью между продольными абберациями нормалей к точкам T_1 и T_2 , то прямую T_2K можно считать нормалью к кривой, а величину $[\Phi_{\text{ПР}} - \delta(\Phi_{\text{ПР}})]$ – новым, более точным значением параметра.

Таким образом, по известному значению $\Phi_{\text{ПР}}$ определяется $\delta(\Phi_{\text{ПР}})$, а разность $\Phi_{\text{ПР}} - \delta(\Phi_{\text{ПР}})$ дает новое значение параметра. Повторяя описанные действия, мы все более приближаем параметр к точному его значению Φ .

После некоторого числа таких итераций параметр φ будет иметь значение Φ_i , ему будут соответствовать точки T_i и V_i . Итерационный цикл прекращается, когда длина отрезка V_iT_i станет меньше некоторого порогового значения ΔL . Чтобы погрешность определения параметра Φ не оказывала заметного влияния на результаты расчета, величина ΔL должна быть существенно (на порядок) меньше погрешности формы зеркала.

ла. Например, если поверхность зеркала выполнена и проконтролирована с погрешностью, не превышающей $0,05 \cdot \lambda$, что для длины волны $\lambda = 0,6328$ мкм составляет $0,03$ мкм, то величина ΔL должна быть не более $0,003$ мкм.

Вышеописанный алгоритм уточнения параметра отражается в формуле:

$$\Phi_{i+1} = \operatorname{arctg} \left[\frac{y(Z(\Phi_i))}{Y(\Phi_i)} \operatorname{tg} \Phi_i \right],$$

где Φ_i и Φ_{i+1} – значения параметра после соответствующей итерации, $Y(\varphi)$ и $Z(\varphi)$ – уравнения, описывающие меридиональную кривую, $y(z) = k \cdot z + b$ – уравнение луча.

Вычисления производятся до тех пор, пока не выполнится условие

$$\left| \frac{Y(\Phi_i) - b}{k} - Z(\Phi_i) \right| \leq \Delta L.$$

Алгоритм расчета АДС реализован в программе Mathcad, что дает возможность для его модификации. Например, можно рассчитать фокусирующую систему с параметрическими поверхностями или АДС, не исправленную на сферическую aberrацию; можно при необходимости задать большее количество поверхностей, но при этом потребуется модификация некоторых формул. Такие возможности делают алгоритм полезным инструментом для расчета оптических систем с параметрически заданными поверхностями.

4. Заключение

Глобальное решение АДС в своем первоначальном виде, полученном профессором Д.Т. Пуряевым, представляет собой математическую связь между параметрическими уравнениями поверхностей АДС, свободной от сферической aberrации осевого пучка лучей. Показано, что оно также определяет связь между математическими выражениями продольных aberrаций нормалей этих поверхностей и позволяет синтезировать АДС с поверхностями любого (не обязательно параметрического) вида.

Эти результаты расширяют возможности разработчика по синтезу АДС с необходимыми геометрическими свойствами.

Для анализа оптических характеристик системы, состоящей из параметрически заданных отражающих поверхностей, разработан специализированный алгоритм расчета наклонных меридиональных пучков.

Литература

1. Михельсон Н.Н. Оптические телескопы. Теория и конструкция. М.: Наука, 1976. 512 с.
2. Пуряев Д.Т. Зеркальная телескопическая система. Авторское свидетельство СССР № 1527607. Бюллетень изобретений, № 45, 1988.
3. Puryayev D.T. Afokal two-mirror system. // Opt. Engin. 1993. V.32. №6. P.1325–1327.
4. Лузин Н.Н. Дифференциальное исчисление. М.: Советская наука, 1958. 475 с.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ВЕРТИКАЛЬНОГО ГРАДИЕНТА ТЕМПЕРАТУРЫ ВОЗДУШНОГО ТРАКТА НА ПОГРЕШНОСТЬ ИЗМЕРЕНИЙ В АВТОРЕФЛЕКСИОННОЙ ОПТИКО-ЭЛЕКТРОННОЙ СИСТЕМЕ

К.Г. Араканцев

Научный руководитель – к.т.н., с.н.с. А.Н. Тимофеев

Обоснована принципиальная возможность построения авторефлексионной оптико-электронной системы (ОЭС) контроля смещений объекта, работающей в условиях стабильных вертикальных градиентов температур воздушного тракта и реализующей двухволновый метод измерений.

Введение

Проблемы пространственного расположения составляющих элементов крупногабаритных конструкций при монтаже, эксплуатации и ремонте, а также позиционирования относительно протяженной базы, особенно при инженерно-геодезических измерениях в машиностроении [1–3] и управлении исполнительными органами машин при производстве земляных работ [4], связаны с решением задач по обеспечению высокоточного, непрерывного, многоточечного контроля смещений частей объекта. В связи с этим разработка автоматизированных систем контроля смещений, призванных обеспечить решение перечисленных задач, а также снизить затраты труда на подготовительные и доделочные операции, сэкономить ресурсы, повысить объективность контроля и сократить ручной труд, является актуальной и практически важной задачей.

В настоящее время для высокоточного измерения линейных поперечных смещений объектов важную роль играют ОЭС. При контроле смещений частей объектов элементы наблюдения зачастую рассредоточены в пространстве и находятся на удалении от устройства контроля и управления, а излучение оптической головки распространяется в приземном слое открытой атмосферы. В этих условиях, как показывает теоретический анализ, наибольшее влияние на точность измерений среди внешних источников погрешностей оказывает рефракция воздушного тракта, зависящая от длины волны излучения, температуры, давления и влажности воздуха. Из этих параметров наибольшее воздействие оказывает изменение температуры. Многоволновые методы компенсации влияния атмосферы позволяют повысить точность оптических измерений [5].

На кафедре ОЭПиС СПбГУ ИТМО разработана авторефлексионная ОЭС, предназначенная для измерения смещений при центрировке турбоагрегатов. Линейные смещения контрольного элемента (КЭ) (трипельпризма), закрепляемого при помощи стандартного приспособления в контролируемой расточке, измеряются в двух взаимно перпендикулярных направлениях. Применение специальных решений в оптической и электронных схемах позволяет производить измерения на дистанциях от 0,5 до 20 м. Измерительная информация обрабатывается программными средствами с помощью Notebook'a в автоматическом режиме. Диапазон измерения смещений по вертикали ± 6 мм, по горизонтали ± 4 мм с погрешностью не более 0,09 мм (погрешность измерений приведена для серии измерений, имеющей 30 отсчетов, проведенной в лабораторных условиях). В рабочих условиях величина вертикального градиента температуры воздушного тракта может достигать 10 К/м. Реализация в такой системе двухволнового метода измерений позволила бы снизить влияние регулярной рефракции на точность измерений.

Обоснование принципиальной возможности применения двухволнового метода измерений в авторефлексионной ОЭС контроля линейных смещений объекта

Рассмотрим наиболее общий случай, когда температурный градиент действует на одном из участков воздушного тракта (рис. 1). Затем полученный результат можно бу-

дет распространить на случай действия градиента температуры на всем протяжении воздушного тракта. При наличии в воздушном тракте температурного градиента возникает градиент показателя преломления, что приводит к искривлению лучей и ошибке в определении координат контролируемого объекта. Двухволновый метод измерений позволяет ослабить влияние градиента температуры при обработке результатов измерений. Этот метод заключается в измерениях координат объекта с использованием излучения двух длин волн (основной и вспомогательной). На основе полученных данных, можно внести поправки в измерения смещения объекта для рабочей длины волны [6]. Для определения алгоритма нахождения поправок была исследована упрощенная измерительная схема: контролируемый элемент перемещается только в вертикальной плоскости; светодиоды расположены в вертикальной плоскости; в качестве контролируемого элемента используется призма Б-180; в системе действует стабильный вертикальный температурный градиент.

Пусть первоначально градиент температуры в воздушном тракте отсутствует, тогда вертикальное смещение объекта с оптической оси объектива должно рассчитываться по формуле (1) с использованием координат изображений, полученных на основной длине волны:

$$Y_1 = \frac{b(y'_{21} + y'_{11})}{4(y'_{21} - y'_{11})}, \quad (1)$$

где y'_{11} , y'_{21} – вертикальные координаты изображений первого и второго светодиодов, измеренные на основной длине волны без градиента температуры в воздушном тракте; b – расстояние между светодиодами в вертикальной плоскости, мм.

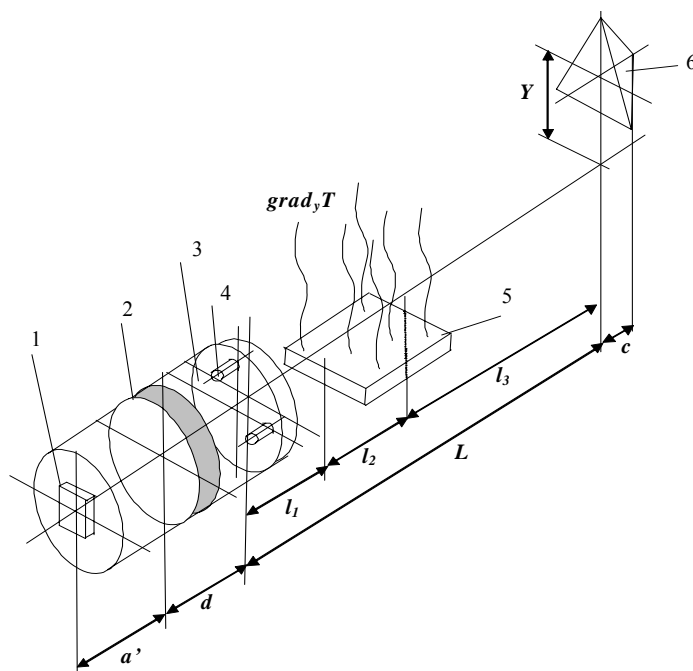


Рис. 1. Авторефлексионная ОЭС с вертикальным градиентом температуры, действующим на участке l_2 воздушного тракта

Здесь 1 – ПЗС-матрица, 2 – объектив, 3 – защитное стекло, 4 – двухцветные светодиоды, 5 – источник вертикального градиента температуры, 6 – контрольный элемент (трипельпризма).

Предположим теперь, что на участке l_2 дистанции L в вертикальном направлении действует постоянный стабильный и равномерный градиент температуры $grad_y T$. Из соотношений геометрической оптики, а также выражений, описывающих распростра-

нение луча в среде с градиентом показателя преломления (вызванным градиентом температуры воздушного тракта), была получена формула, позволяющая рассчитать величину вертикального смещения КЭ с учетом действия регулярной рефракции

$$Y_{1g} = \frac{b(y'_{21g} + y'_{11g})}{4(y'_{21g} - y'_{11g})} + \frac{\tan(\alpha_1)}{2} (L_1 - l_1 + l_3 + 2c - \Delta_{p1}). \quad (2)$$

Здесь y'_{11g}, y'_{21g} – вертикальные координаты изображений первого и второго светодиодов, измеренные на основной длине волны при действии на участке l_2 воздушного тракта стабильного вертикального градиента температуры; α_1 – угол по отношению к оптической оси объектива, на который на участке l_2 отклонится луч основной длины волны, вошедший на этот участок воздушного тракта параллельно оптической оси объектива (рис. 2); c – размер трипельпризмы вдоль оптической оси объектива, мм; Δ_{p1} – удлинение, вносимое стеклом трипельпризмы, мм ($\Delta_{p1} = 2c(n_1 - 1)/n_1$); n_1 – показатель преломления стекла призмы для излучения основной длины волны; $L_1 = \frac{a'b}{2(y'_{11g} - y'_{21g})} + \frac{1}{2}(\Delta_{p1} - 2c - d)$ – расстояние, измеренное на основной длине волны, от плоскости, в которой размещены излучающие площадки светодиодов, до КЭ, мм; a' – задний отрезок объектива, мм; d – расстояние от передней главной плоскости объектива до плоскости, в которой размещены излучающие площадки светодиодов, мм.

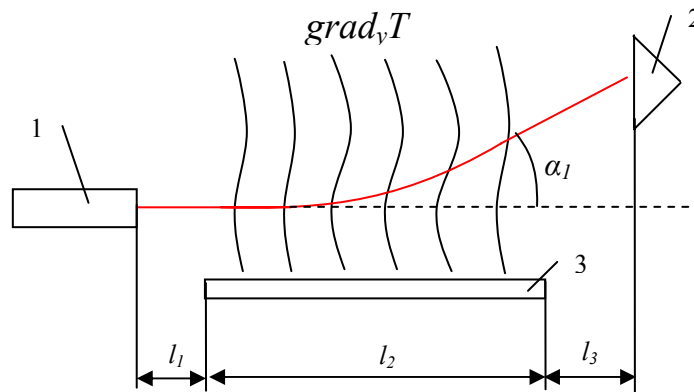


Рис. 2. К выводу формулы (2) для определения величины вертикального смещения (Y) КЭ в условиях действия на участке l_2 воздушного тракта стабильного вертикального градиента температуры

Здесь 1 – излучающая головка системы, 2 – контрольный элемент, 3 – источник вертикального градиента температуры.

Из сравнения формул (1) и (2) видно, что величина ошибки в определении Y из-за действия в системе $grad_y T$ составит

$$\Delta_1 = \frac{\tan(\alpha_1)}{2} (L_1 - l_1 + l_3 + 2c - \Delta_{p1}). \quad (3)$$

Чтобы произвести вычисление вертикального смещения КЭ по формуле (2), необходимо знать величину $\tan(\alpha_1)$. Из теоретических положений, изложенных в [7], легко получить, что $\tan(\alpha_1) = -\frac{n_1 - 1}{T} grad_y T \cdot l_2$, где n_1 – средний показатель преломления атмосферы для излучения основной длиной волны; T – средняя температура атмосферы, К; $grad_y T$ – величина стабильного вертикального градиента температуры, действующего на участке l_2 воздушного тракта. Для определения величины $grad_y T$ может быть использован двухволновый метод измерений. Для этого после измерения Y на основной длине волны, производится измерение Y на вспомогательной длине волны. Теоретически измеренное значение Y на вспомогательной длине волны выражается формулой

$$Y_{2g} = \frac{b(y'_{22g} + y'_{12g})}{4(y'_{22g} - y'_{12g})} + \frac{\tan(\alpha_2)}{2} (L_2 - l_1 + l_3 + 2c - \Delta_{p2}). \quad (4)$$

Здесь y'_{12g} , y'_{22g} – вертикальные координаты изображений первого и второго светодиодов, измеренные на вспомогательной длине волны при действии на участке l_2 воздушного тракта стабильного вертикального градиента температуры; α_2 – угол по отношению к оптической оси объектива, на который на участке l_2 отклонится луч вспомогательной длины волны, вошедший на этот участок воздушного тракта параллельно оптической оси объектива; Δ_{p2} – удлинение, вносимое стеклом трипельпризмы, мм ($\Delta_{p2} = 2c(n_2 - 1)/n_2$); n_2 – показатель преломления стекла призмы для излучения вспомогательной длины волны; $L_2 = \frac{a'b}{2(y'_{12g} - y'_{22g})} + \frac{1}{2}(\Delta_{p2} - 2c - d)$ – расстояние, измеренное

на вспомогательной длине волны, от плоскости, в которой размещены излучающие площадки светодиодов, до КЭ, мм;. В этом случае ошибка в определении Y составит

$$\Delta_2 = \frac{\tan(\alpha_2)}{2} (L_2 - l_1 + l_3 + 2c - \Delta_{p2}). \quad (5)$$

Поскольку величина $\text{grad}_y T$ изначально неизвестна, при измерениях на обеих длинах волн расчет Y производится по формуле (1). Разность полученных значений будет равна разности ошибок Δ_1 и Δ_2 :

$$\Delta Y_{12} = \Delta_2 - \Delta_1. \quad (6)$$

Измерив разность ΔY_{12} , с помощью формул (3), (5), (6) можно рассчитать величину температурного градиента $\text{grad}_y T$ и по формуле (2) рассчитать скорректированное с учетом действия регулярной рефракции значение вертикального смещения КЭ.

На рис. 3 показан макет авторефлексионной ОЭС, реализующей двухволновый метод измерений [8].

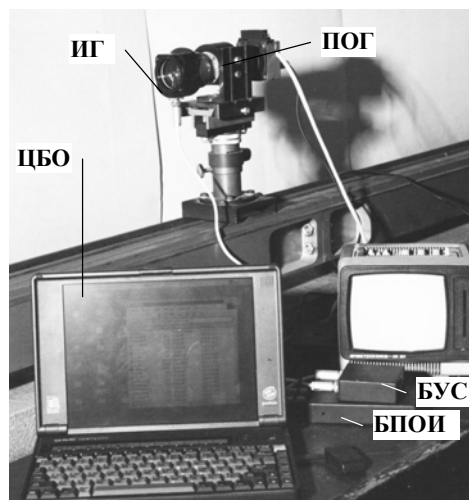


Рис. 3. Общий вид системы (без КЭ) на основе ЦБО – HP Omnibook

Здесь ЦБО – центральный блок обработки информации, БПОИ – блок предварительной обработки информации, БУС – блок управления светодиодами, ИГ – излучающая головка (с двумя двухцветными светодиодами), ПОГ – приемная оптическая головка (на основе ПЗС-матрицы).

Анализ формул (3), (5), (6) показал, что для регистрации малых величин вертикального градиента температуры в воздушном тракте системы необходимо оптимальным образом выбрать параметры системы. А именно:

- использовать объектив с максимально возможной величиной заднего отрезка a' (в данной системе $a' = 250$ мм);

- использовать ПЗС-матрицу высокого разрешения и большого формата (SONY ICX255AL с разрешением 752×582);
- максимизировать разность показателей преломления n_1 и n_2 ; при этом выбранные длины волн, на которых излучают светодиоды, должны быть согласованы со спектральной чувствительностью ПЗС-матрицы (для рассматриваемой системы указанным требованиям удовлетворяют светодиоды, излучающие на длинах волн $\lambda_1=940$ нм, $\lambda_2=430$ нм);
- выбрать максимально возможную дистанцию измерений ($L=20$ м);
- создавать вертикальный градиент температуры в непосредственной близости от излучающей головки и, по возможности, обеспечить как можно большую зону действия градиента вдоль оптической оси.

Предложена следующая методика двухволнового метода измерений:

- измерение вертикального смещения КЭ Y_1 на основной длине волны;
- измерение вертикального смещения КЭ Y_2 на вспомогательной длине волны;
- вычисление хроматической разности ΔY_{12} ;
- вычисление величины вертикального градиента температуры воздушного тракта $\text{grad}_y T$;
- внесение поправки в значение вертикального смещения КЭ Y_1 , измеренное на основной длине волны.

Анализ погрешностей измерения линейных смещений объекта авторефлексионной ОЭС при действии в воздушном тракте системы стабильного и равномерного вертикального градиента температуры

Значения показателей преломления воздушного тракта для основной и вспомогательной длины волны можно найти из выражения для индекса рефракции, величина которого в зависимости от температуры воздуха (T , К), атмосферного давления (P , мбар), парциального давления водяного пара ($P_{\text{ен}}$, мбар) и длины волны излучения (λ , мкм) достаточно точно выражается формулой [7]

$$N = \frac{77.6P}{T} + \frac{0.584P}{T\lambda^2} - 0.06P_{\text{ен}}, \quad (7)$$

при этом $n = N \cdot 10^{-6} + 1$.

Значения T , P , $P_{\text{ен}}$ из-за неоднородности приземного слоя воздуха в зоне распространения излучения непостоянны. Их изменения относительно значений в начале трассы носят неопределенный характер (в смысле знания их конкретных величин), поэтому численное значение показателя преломления, полученное из выражения для N , будет приближенным.

Из графиков, приведенных на рис. 4, видно, что для разных длин волн изменение индекса рефракции воздушного тракта при изменении температуры от 250 К до 300 К остается постоянным. При этом с увеличением длины волны излучения зависимость индекса рефракции от длины волны (при фиксированном значении T) становится менее сильной. С целью повышения точности измерения ошибки Δ_l необходимо, чтобы разность ΔY_{12} была максимально возможной. Анализ формул (3), (5), (6) показал, что для получения максимальной величины дисперсионной разности необходимо максимизировать разность показателей преломления воздушного тракта для излучения основной и вспомогательной длин волн.

Для этого основная и вспомогательная длины волн должны быть выбраны в длинноволновой и коротковолновой частях спектра соответственно. Эти значения должны быть согласованы со спектральной чувствительностью приемника излучения (ПЗС-матрица).

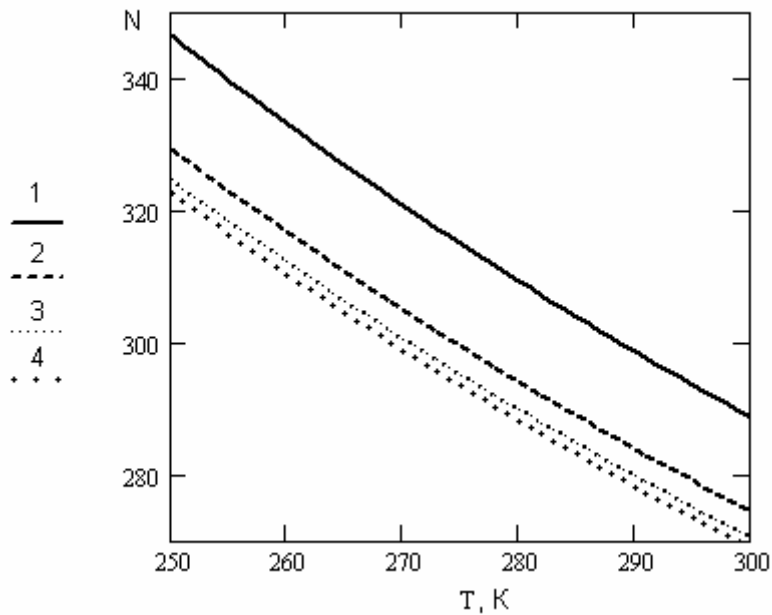


Рис. 4. Зависимости индекса рефракции воздушного тракта от температуры для излучения разных длин волн: 1 – $\lambda=0.3$ мкм, 2 – $\lambda=0.5$ мкм, 3 – $\lambda=0.7$ мкм, 4 – $\lambda=0.9$ мкм

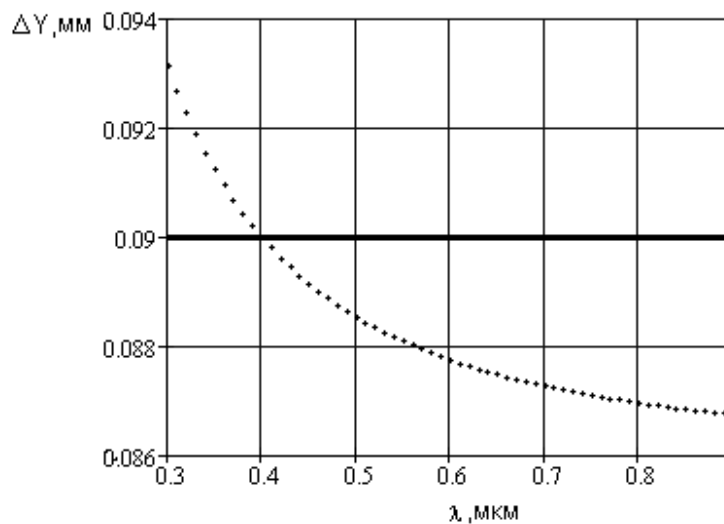


Рис. 5. Зависимость ошибки в определении вертикального смещения КЭ от длины волны для вертикального градиента температуры воздушного тракта 0.4 К/м

График на рис. 5 построен в соответствии с формулой (3), определяющей величину ошибки при измерении вертикального смещения КЭ при действии в системе вертикального температурного градиента. Величина ошибки стремится к некоторому постоянному значению с увеличением длины волны излучения. Среднеквадратическая погрешность измерения вертикального смещения КЭ в рассматриваемой системе составляет 0.09 мм. Из графика на рис. 5 видно, что не имеет смысла использовать в данной системе двухволновый метод измерений, если в воздушном тракте действует вертикальный градиент температуры меньший, чем 0.4 К/м, поскольку в этом случае погрешность от рефракции воздушного тракта меньше, чем погрешность, обусловленная случайными факторами. В данном случае предполагалось, что вертикальный градиент температуры действует на всем протяжении воздушного тракта. Однако в общем случае зона действия градиента температуры может занимать лишь часть трассы (рис. 6).

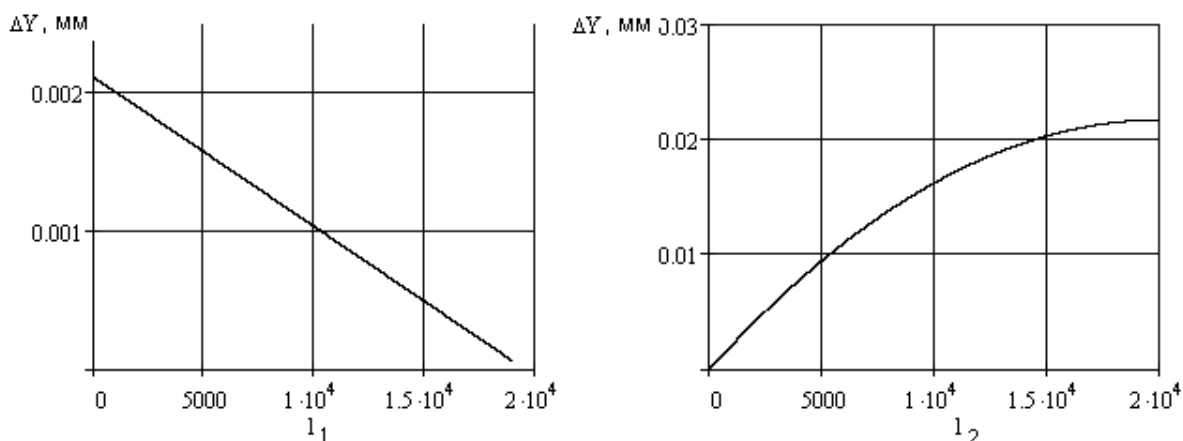


Рис. 6. Зависимости погрешности измерения вертикального смещения КЭ от положения (l_1) и протяженности (l_2) участка действия вертикального градиента температуры

Графики на рис. 6 поясняют положение о том, что для обеспечения максимальной чувствительности авторефлексионной измерительной системы к вертикальному градиенту температуры воздушного тракта при испытаниях системы в лабораторных условиях необходимо создать вертикальный градиент температуры в непосредственной близости от излучающей головки и, по возможности, обеспечить как можно большую зону действия градиента вдоль оптической оси излучающей головки.

Основные результаты исследования

- Обоснована возможность и целесообразность применения в авторефлексионных ОЭС двухволнового метода для уменьшения погрешности измерений от стабильного и равномерного вертикального градиента температуры воздушного тракта.
- Разработана методика измерений двухволновым методом.
- Разработан макет авторефлексионной ОЭС контроля поперечных линейных смещений объекта, в котором заложен двухволновый метод измерений.
- Оптимальным образом подобраны параметры системы для наиболее эффективного использования двухволнового метода при измерениях линейных смещений объектов.

Заключение

В работе доказана принципиальная возможность построения авторефлексионной ОЭС контроля смещений, работающей в условиях стабильных вертикальных градиентов температур. Обоснован выбор параметров макета авторефлексионной ОЭС, реализующей двухволновый метод измерений. В дальнейшем планируется модификация экспериментальной установки и проведение углубленных экспериментальных исследований трехволнового метода измерений, который позволит повысить точность измерений положения объекта в условиях действия в системе регулярной рефракции по сравнению с двухволновым методом.

Литература

1. Высокоточные угловые измерения. / Под ред. Ю.Г. Якушенкова. М.: Машиностроение, 1987. 480 с.

2. Панков Э.Д., Коняхин И.А. Трехкоординатные оптические и оптико-электронные угломеры. М.: Недра, 1991. 251 с.
3. Афанасьев В.А., Усов В.С. Оптические приборы и методы контроля прямолинейности в инженерной геодезии. М.: Недра, 1973.
4. Неумывакин Ю.К. Автоматизация геодезических измерений в мелиоративном строительстве. М.: Недра, 1984.
5. Араканцев К.Г. Исследование двухспектрального метода ослабления влияния вертикального градиента температуры воздушного тракта в оптико-электронной системе контроля смещений. // Оптика-2005. Труды четвертой международной конференции молодых ученых и специалистов "Оптика-2005". Санкт-Петербург, 17-21 октября 2005. / Под ред. проф. В. Г. Беспалова, проф. С.А. Козлова. СПб: СПбГУ ИТМО, 2005. 398 с.
6. Прилепин М.Т., Голубев А.Н. Оптические квантовые генераторы в геодезических измерениях. М.: Недра, 1972. 168 с.
7. Сытник В.С. Лазерные геодезические приборы в строительстве. М.: Стройиздат, 1988.
8. Крайлюк А.Д., Краснящих А.В., Мусяков В.Л., Тимофеев А.Н., Ярышев С.Н. Оптико-электронная система контроля положения центра корпусных деталей турбоагрегатов относительно оптической оси. // Изв. вузов. Приборостроение. 2003. Т. 46, №8. С. 61–63.

АНАЛИЗ ПРОСТРАНСТВЕННОГО РАСПРЕДЕЛЕНИЯ ЭНЕРГИИ В ПЛАНАРНОЙ ОПТИЧЕСКОЙ РАВНОСИГНАЛЬНОЙ ЗОНЕ

Е.М. Богатинский

Научный руководитель – к.т.н., с.н.с. А.Н. Тимофеев

Задача поставленной работы заключалась в исследовании распределения облученности, создаваемой несколькими источниками излучения, в произвольной точке пространства в пределах зоны действия для оптико-электронной системы управления.

Введение

В связи с тем, что облученность в каждой точке пространства от задатчика базовой плоскости распределена неравномерно [1], возникает необходимость исследовать распределение энергии в пространстве, охваченном областью работы предлагаемой системы. Требуется, чтобы оптическая равносигнальная зона (ОРСЗ) была как можно ближе к плоскости и ее форма, главным образом, определяет распределение энергии.

Поэтому при выполнении проекта предлагается посвятить основные исследования вопросу о влиянии расфокусировки и аберраций на пространственное распределение энергии в планарной ОРСЗ, что в конечном итоге определяет главную погрешность позиционирования [2].

Основная часть

Методика расчета пространственного распределения энергии в оптико-электронной системе с круговой ОРСЗ приведена в статье [1]. Объектив сфокусирован на максимальную дистанцию $L_{\max} = 100$ метров, найдем величину облученности в точке $M(x, y, L)$ на различных дистанциях L при $y = 0$ (равносигнальная зона); $x = var$. Мы ищем распределение энергии в горизонтальной плоскости, т.е. изменение облученности на оси в зависимости от смещения x при заданной дистанции L (рис. 1). Согласно представленным зависимостям, можно сделать вывод, что величина провала энергии до значения нуля (просвета) на больших дистанциях контроля существенна [3].

Проведем расчет энергии для оптической системы, имеющей аберрации. С помощью программы OPAL была рассчитана поперечная сферическая аберрация представленного объектива, которая была представлена графически. На основании вышеизложенного согласно алгоритму [1], был произведен расчет пространственного распределения энергии в произвольной точке с учетом поперечной аберрации оптической системы. Алгоритм вычисления весьма трудоемкий и долгий. Вычисления производились поэлементно. Однако не исключена и граничная погрешность, обусловленная тем, что элементы имеют размеры. В некоторых случаях элемент не полностью попадал на выходной луч, хотя и был включен в общий результат. Результаты расчета (при $y = 0$) представлены на рис. 2. Данные зависимости показывают, что наличие аберрации идет в нашей системе на пользу, из-за чего величина области нечувствительности сокращается (эффект замытия).

Площади под графиками без учета и с учетом аберраций на одинаковых дистанциях равны, это означает, что энергетических потерь у нас нет и можно считать предложенную методику правильной.

Дальнейшая цель наших исследований состоит в рассмотрении градиента энергии в вертикальной плоскости в зависимости от смещения y при заданной дистанции L . Только в этом случае мы еще будем задавать смещение по оси x , т.е. мы будем «нарезать» графики на рис. 2 перпендикулярными плоскостями и расчет производить сразу с учетом аберрации оптической системы прожектора. Результаты представлены на рис. 3.

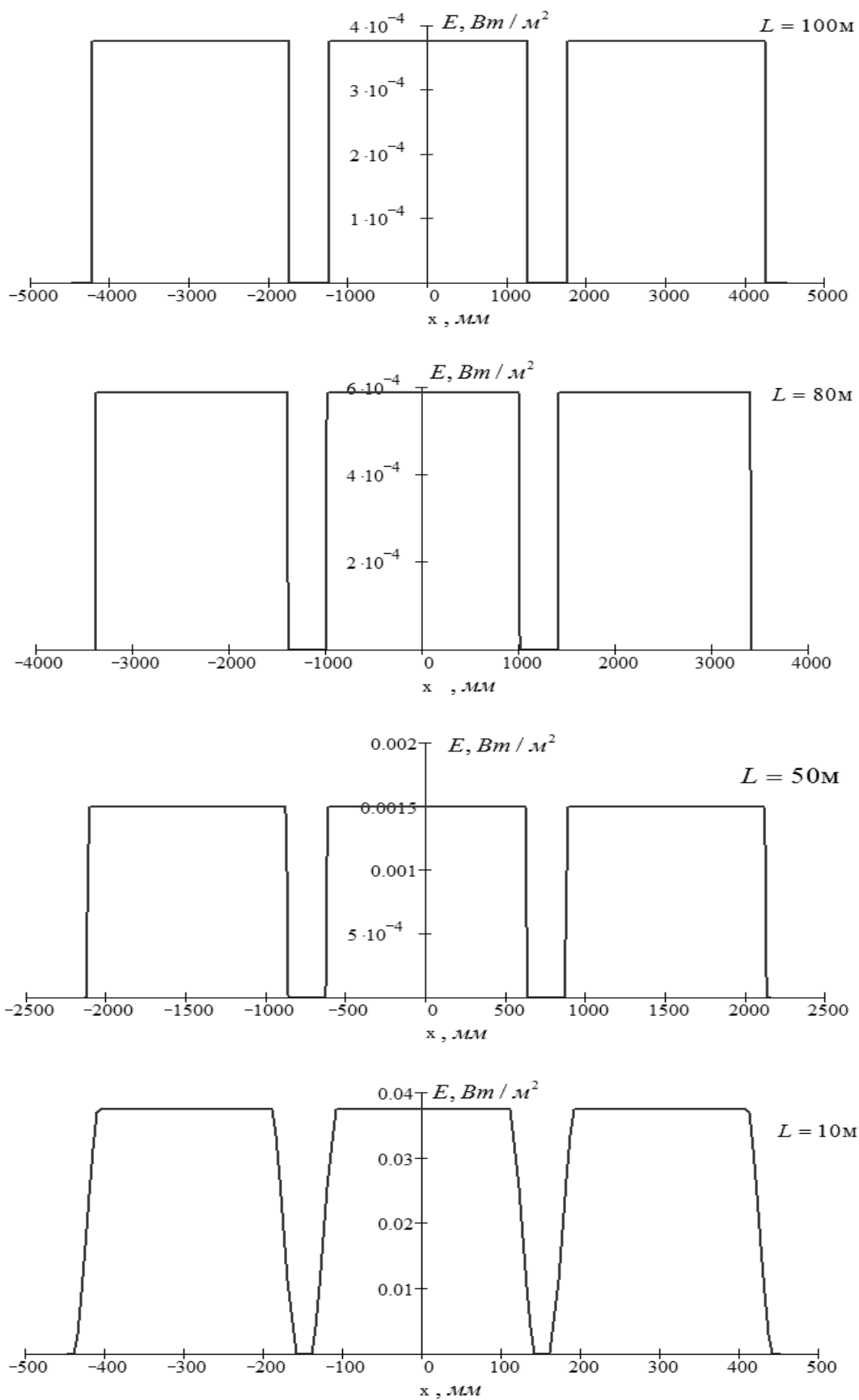


Рис. 1. Пространственное распределение энергии ($\text{Вт}/\text{м}^2$) от 3-х источников в горизонтальной плоскости для оптической системы без наличия aberrаций при различных дистанциях

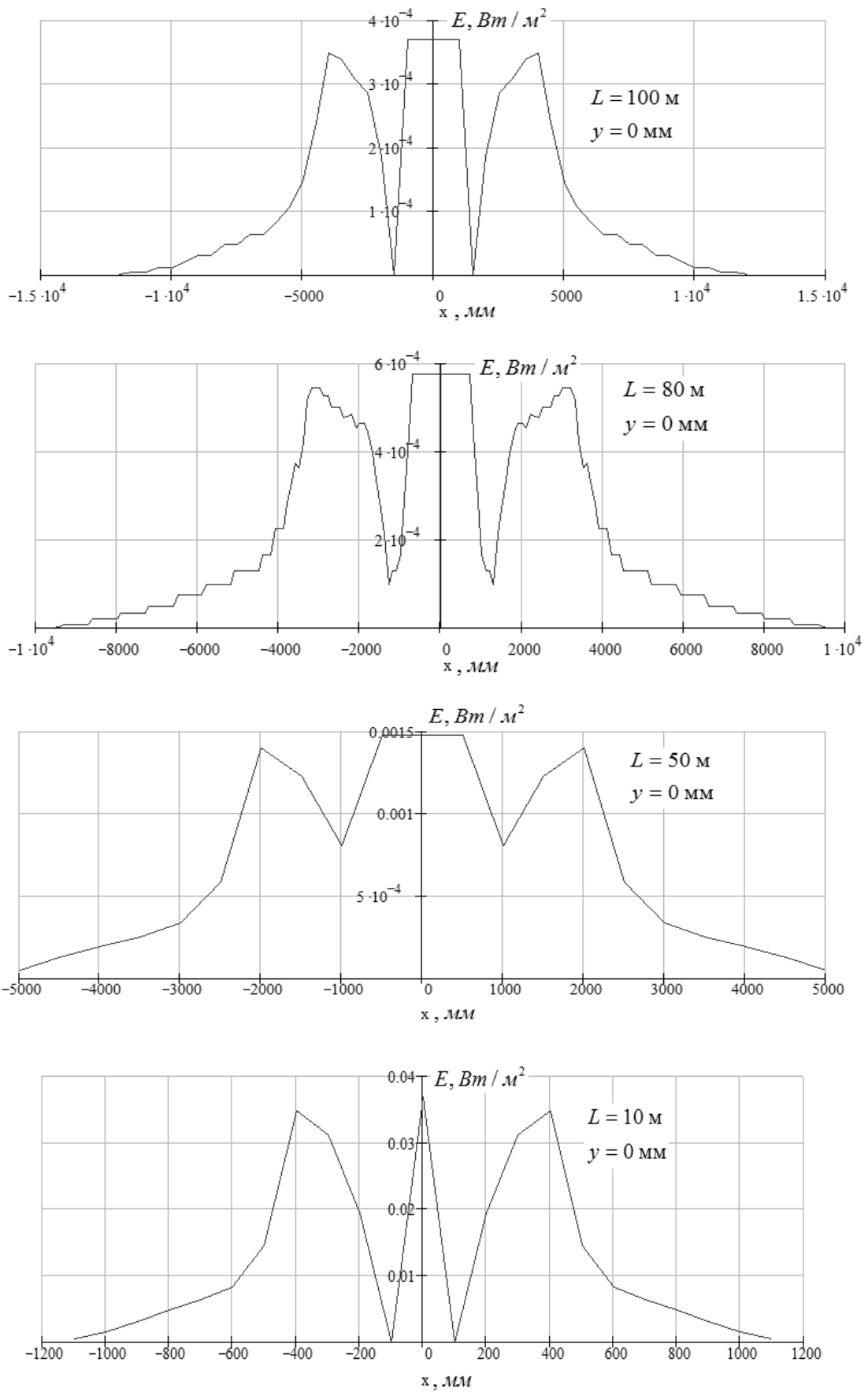
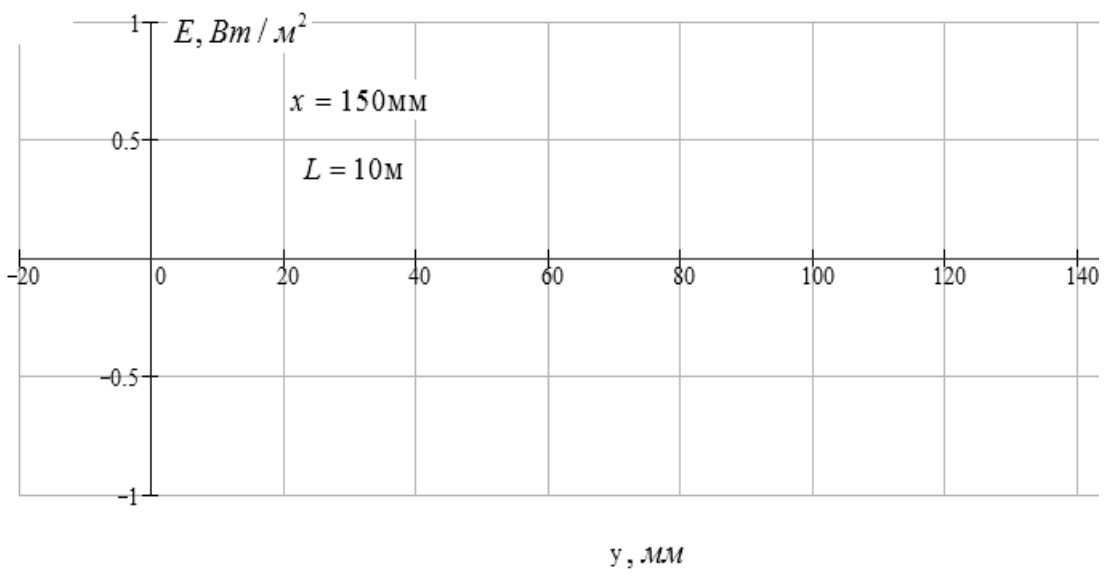
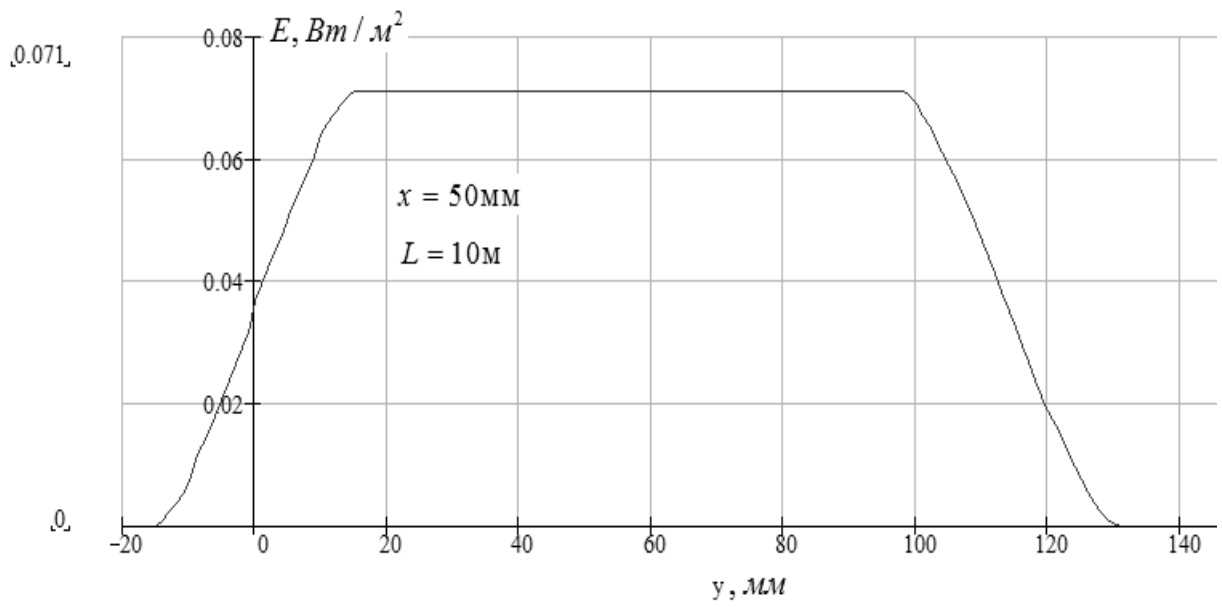
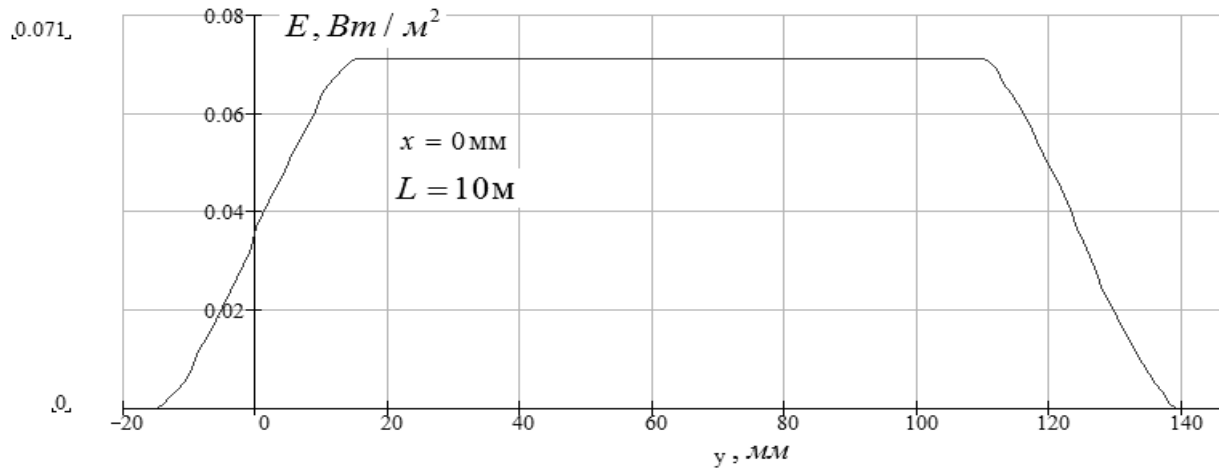


Рис. 2. Пространственное распределение энергии ($\text{Вт}/\text{м}^2$) от 3-х источников в горизонтальной плоскости для оптической системы при наличии поперечной сферической aberrации для различных дистанций



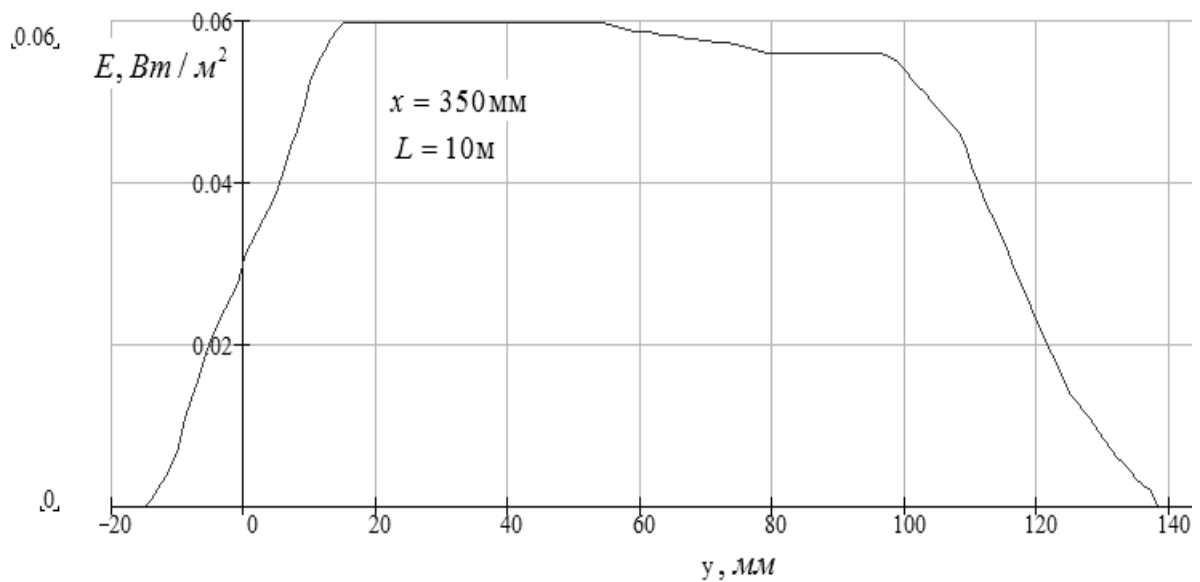
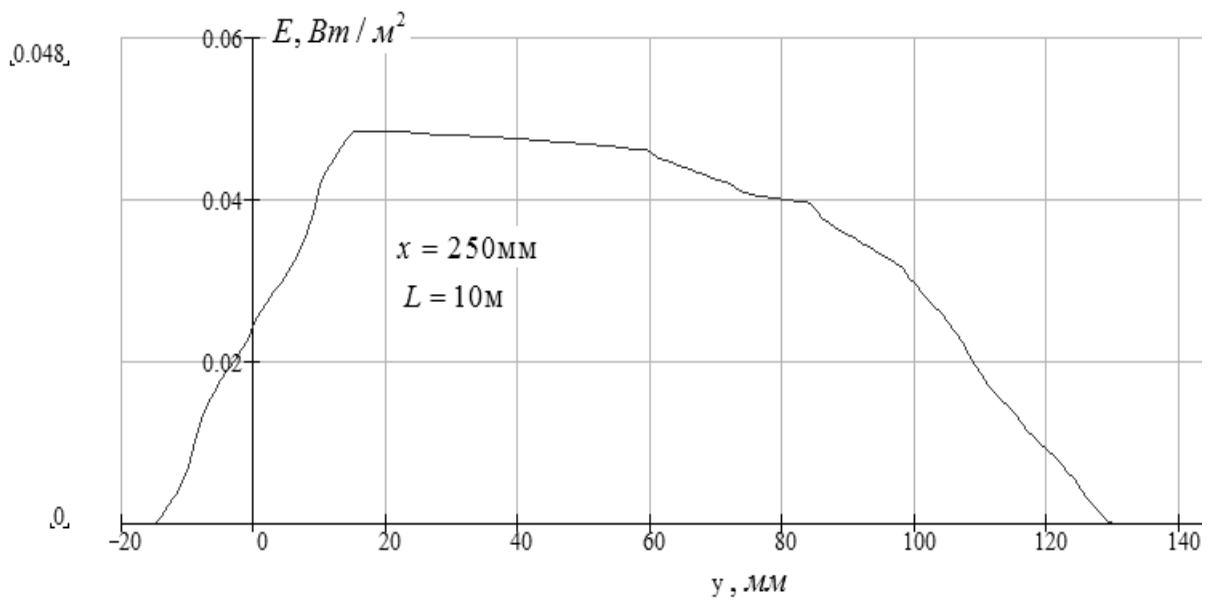


Рис. 3. Распределение энергии от источников в вертикальной плоскости на дистанции 10 м в зависимости от диапазона y и величины смещения x

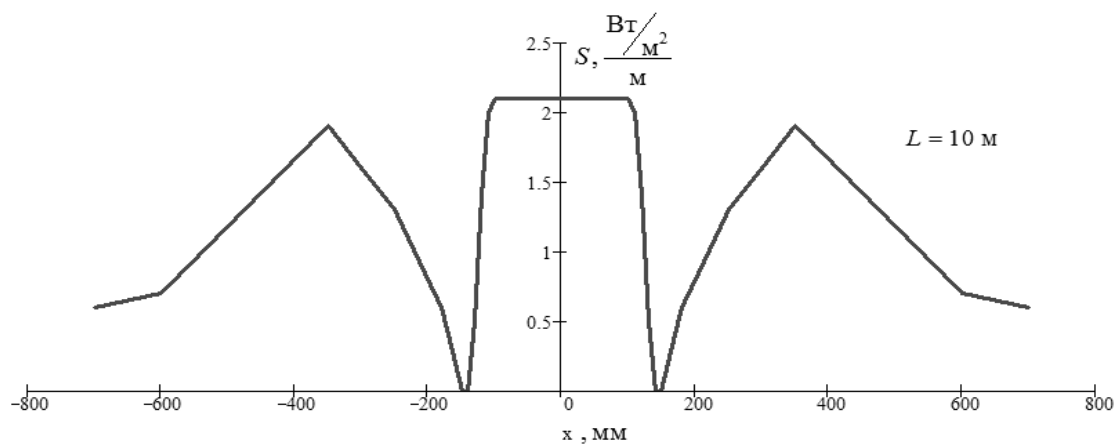


Рис. 4. Распределение чувствительности в вертикальной плоскости в зависимости от величины смещения x на дистанции 10 м

Для нахождения чувствительности системы [4] необходимо найти тангенс наклона восходящего линейного участка графиков, представленных на рис. 3. В результате получим зависимость чувствительности от величины смещения (рис. 4).

В дальнейших исследованиях планируется создать макет предлагаемой схемы и оценить влияние характера aberrаций с помощью физической модели на точность позиционирования, а также провести дополнительные исследования стабильности пространственного положения планарной ОРСЗ.

Заключение

- С помощью пакета прикладных программ MathCAD составлен алгоритм расчета облученности по результатам предложенной методики [1].
- Рассчитана чувствительность оптико-электронной системы с ОРСЗ.
- Представленные в работе графические зависимости позволяют найти и оценить наиболее предпочтительные зоны управления.

Литература

1. Богатинский Е.М. Формирование и анализ пространственного распределения энергии в планарной оптической равносигнальной зоне // VII Международная конференция «Прикладная оптика – 2006» Т. 3. СПб: Оптическое общество им. Д.С. Рождественского, 2006. С. 297–301.
2. Богатинский Е.М. Выбор параметров построения оптической схемы прожектора, реализующего круговую оптическую равносигнальную зону. / Девятая Санкт-Петербургская Ассамблея молодых ученых и специалистов. Аннотации работ по грантам Санкт-Петербургского конкурса 2004 г. для молодых ученых и специалистов. СПб.: СПбГУ, 2004. С. 48–49.
3. Барсуков О.А., Тимофеев А.Н. Особенности формирования оптической равносигнальной плоскости. // Оптико-электронные приборы и системы: Сб. науч. статей. Вып. 99./ Под ред. Э.Д. Панкова. СПб: СПбГИТМО (ТУ), 1999.
4. Джабиев А.Н., Мусьяков В.Л., Панков Э.Д., Тимофеев А.Н. Оптико-электронные приборы и системы с оптической равносигнальной зоной. Монография / Под общей редакцией Э.Д. Панкова. СПб: ИТМО, 1998.

ОБ ОСНОВНЫХ ПОГРЕШНОСТЯХ КОНТРОЛЯ СООСНОСТИ С ПОМОЩЬЮ АВТОРЕФЛЕКСИОННОЙ ОПТИКО-ЭЛЕКТРОННОЙ СИСТЕМЫ

А.М. Алеев, А.Г. Анисимов, А.В. Пантюшин
Научный руководитель – к.н.т., с.н.с. А.Н. Тимофеев

Работа посвящена исследованию погрешностей, возникающих в авторефлексионной оптико-электронной измерительной системе с матричными фотоприемными устройствами, вследствие искривления и изменения референтного направления при перефокусировании. Проводится исследование влияния децентрирования компонентов на точностные характеристики. Теоретически полученные зависимости иллюстрируются на примере оптико-электронной системы контроля соосности.

Введение

Соблюдение высокой точности пространственного расположения составляющих элементов крупногабаритных конструкций относительно систем прямолинейных осей является актуальной задачей для многих отраслей промышленности [1]. С развитием автоматизации на производстве возникла потребность автоматизации такого контроля. В связи с этим необходим пространственный контроль, который часто осуществляется с помощью оптико-электронных измерительных систем (ОЭИС), в том числе и с фоточувствительными приборами с зарядовой связью (ФПЗС) [2] или другими матричными фотоприемниками. Основным требованием при разработке подобных устройств является обеспечение необходимой точности измерений, которое зависит как от внешних факторов, влияющих на ОЭИС, так и от внутренних параметров самой оптической системы. В литературе существуют указания на то, что в результате децентрирования оптических компонентов измерительных систем появляется погрешность измерения [3], однако их применимость, зачастую, ограничивается визуальными измерительными системами.

Авторефлексионные схемы устройств с матричными фотоприемниками применяются в современном приборостроении относительно недавно и еще недостаточно изучены. Так, до настоящего времени неполно исследовано влияние децентрирования компонентов системы относительно оптической оси ОЭИС на величину погрешности измерения. Данная работа призвана восполнить указанный пробел в литературе.

Проводилось исследование погрешностей, возникающих в системе контроля положения элементов крупногабаритных конструкций. В системе применена авторефлексионная схема на основе оптико-электронных измерительных преобразователей, осуществляющих предварительную обработку видеосигнала и передачу измерительной информации на ЭВМ.

1. Анализ погрешностей авторефлексионной схемы

Схема исследуемой авторефлексионной системы представлена на рис. 1. На ФПЗС 1, находящийся в приемном блоке 2, через объектив 3 и светофильтр излучающей головки 6, проецируются изображения двух пар полупроводниковых излучающих диодов (ПВД) 4 и 5, закрепленных на оптической головке. В качестве отражателя предлагается использовать трипельпризму 7, а в качестве объектива – телеобъектив с подвижным отрицательным компонентом.

Необходимость использования двух пар ПВД, объясняется тем, что для обеспечения широкого диапазона измеряемых величин, в частности, дистанции, без ущерба для точностных характеристик системы, оказывается невозможно найти оптимальное значение базового расстояния между двумя ПВД. В связи с этим и возникает необходимость установки в системе нескольких пар полупроводниковых излучающих диодов .

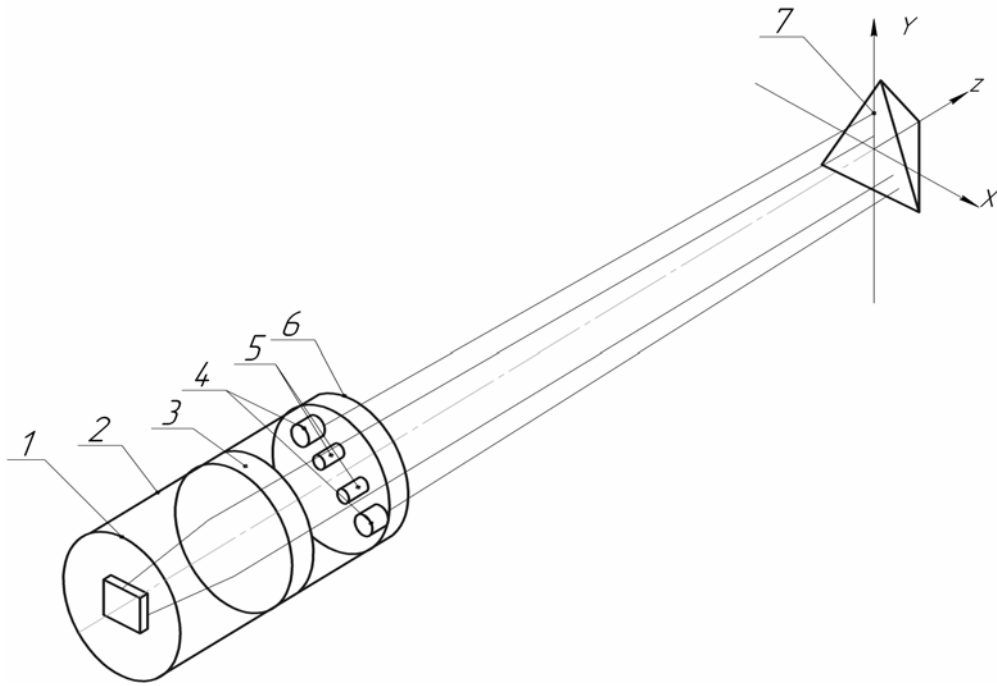


Рис. 1. Схема авторефлексионная

Следует отметить, что использование в системе трипельпризмы в качестве отражателя приводит при смещении ее в направлении нормали к оптической оси к двойному смещению изображения по сравнению с тем, которое происходит в отсутствие трипельпризмы при сдвиге контролируемого объекта. Соответственно, происходит и увеличение погрешности измерения в два раза по сравнению с тем случаем, когда вместо авторефлексионной схемы реализуется коллимационная, т.е. не применяются отражатель и пучки лучей проходят прямо от источника к приемнику.

Как известно, референтное направление есть прямая, проходящая через центр матрицы фотоприемника и заднюю узловую точку телеобъектива [4]. Теоретически, при идеальном центрировании всех компонентов измерительной системы, референтное направление представляет собой прямую, совпадающую с оптической осью системы. Но в любой реальной оптической системе существуют различные погрешности, в частности – децентрирование ее компонентов, в результате которого происходит изменение (с сохранением прямолинейности) либо искривление референтного направления.

Неточное совмещение центра матрицы фотоприемника с оптической осью системы вызывает изменение референтного направления при перефокусировании на различные расстояния, что, в свою очередь, влечет за собой появление систематической погрешности измерения смещений δh_l для каждой дистанции. Используя положения геометрической оптики [5], при исследовании хода лучей, идущих в системе вдоль ее оптической оси, была получена следующая формула для величины погрешности δh_l :

$$\delta h_l = \frac{\delta c_1 f'}{(f' - A'_s)}, \quad (1)$$

где $f' = \frac{f'_1 f'_2}{f'_1 + f'_2 - d_s}$ – эквивалентное фокусное расстояние телеобъектива, с фокусировкой отрицательным компонентом; расстояние между положительным и отрицательным компонентами телеобъектива определяются выражением [6]

$$d_s = \frac{1}{2} \left(a - L - \sqrt{(L - a)(L - 4f'_2 - a)} \right),$$

величины a и A'_s определяются соотношениями [6]

$$a = \frac{s \cdot f_1'}{s + f_1'}, \quad A_s' = L - d_s / f_2'(L - d_s),$$

где s – расстояние до объекта (дистанция фокусирования), δc_1 – погрешность совмещения центра матрицы фотоприемника с оптической осью, f_1', f_2' – фокусные расстояния положительного и отрицательного компонентов телеобъектива, соответственно, L – длина телеобъектива.

Изменение референтного направления происходит также при смещении с оптической оси центра положительного компонента телеобъектива системы, что приводит к появлению систематической погрешности измерения δh_2 . Задавая некоторую величину δc_2 смещения положительного компонента телеобъектива с оптической оси и проведя исследование хода лучей, идущих вдоль оптической оси системы (т.е. так же, как и в случае определения δh_1), для величины δh_2 , можно получить формулу

$$\delta h_2 = \frac{\delta c_2 s}{f_1}. \quad (2)$$

В отличие от двух предыдущих случаев, при децентрировании относительно оптической оси отрицательного компонента телеобъектива (т.е. фокусирующей линзы) происходит не просто изменение, но искривление референтного направления. Соответственно, полученная тем же методом, что δh_1 и δh_2 , зависимость возникающей при этом погрешности измерения δh_3 от дистанции фокусирования имеет нелинейный характер:

$$\delta h_3 = \frac{\delta c_3 (L - d_s)}{f_2'}, \quad (3)$$

где δc_3 – величина смещения фокусирующей линзы с оптической оси системы.

В рассмотренной схеме также могут присутствовать погрешности установки центра базового расстояния полупроводниковых излучающих диодов, а также погрешности вызванные поворотами оси симметрии трипельпризмы относительно оптической оси.

Ошибки, связанные с поворотом трипельпризмы, в приближении геометрической оптики можно записать в виде:

$$\delta h_4 = \frac{n-1}{n} d_{\Pi} \sin \alpha \cdot \left(1 + r \cdot n / d_{\text{ВП}} (n-1) \right), \quad (4)$$

где $d_{\Pi} = k \cdot a$, a – наибольшая длина пучка, могущего пройти через развертку призмы [5], k – коэффициент призмы для трипельпризмы, $k=2$, α – угол между осью симметрии призмы и оптической осью, r – расстояние от вершины трипельпризмы до точки вращения (отрицательное направление вглубь стекла), $d_{\text{ВП}}$ – высота трипельпризмы. Точка, при вращении вокруг которой трипельпризма нечувствительна к повороту, располагается от вершины призмы на расстоянии [7]

$$d_1 = -d_{\Pi} (n-1)/n.$$

При расчете суммарной погрешности, вносимой трипельпризмой, не учитываются погрешности изготовления трипельпризмы. В данном случае это оправдано, так как в правилах эксплуатации высокоточных оптико-электронных систем, построенных по авторефлексионной схеме, не допускается замена трипельпризмы на аналогичную, а также ее поворот вокруг своей оси [8]. Также не учитывается то, что в реальных системах трипельпризма может быть смещена в своем корпусе относительно оптической оси и вдоль нее.

Штативы 5 и 7 рис. 3 устанавливались в нулевое положение, после чего штатив 5 перемещался по оптической скамье, прямолинейность движения штатива контролировалась борштангой.

Вышесказанное приводит тому, что выражение 4 носит приблизительный характер и может быть использовано только для оценки величины погрешности, вносимой

трипельпризмой. Для более строгого рассмотрения поведения трипельпризмы использовалась математическая модель ее поведения [7].

Смещение центра базового расстояния между ПИД δc_5 с оптической оси приводит к появлению погрешности

$$\delta h_5 = -\delta c_5. \quad (5)$$

Кроме рассмотренных выше погрешностей, существуют также и погрешности, вызванные погрешностью определения энергетического центра излучения. Теория этой погрешности достаточно сложна и является предметом отдельного изучения [9, 10]. Однако существует эмпирически подобранное выражение, дающее хорошую сходимость с экспериментальными данными:

$$\delta h_6 = \frac{s \cdot m \cdot \delta c_6}{f'}, \quad (6)$$

где f' – эквивалентное фокусное расстояние телеобъектива, s – расстояние до объекта, m – линейный размер пикселя матрицы фотоприемника, δc_6 – погрешность определения энергетического центра.

2. Исследование суммарной погрешности авторефлексионной схемы

Суммарную погрешность системы, построенной по авторефлексионной схеме можно вычислить как

$$\delta h = \sqrt{\delta h_1^2 + \delta h_2^2 + \delta h_3^2 + \delta h_4^2 + \delta h_5^2 + \delta h_6^2},$$

где, $\delta h_1 - \delta h_6$ находятся из выражений (1) – (6), соответственно.

При погрешностях на технологическом уровне [11], равных $\delta c_1=0,05$ мм $\delta c_2=0,01$ мм $\delta c_3=0,01$ мм, $\alpha=5'$, $d_{II}=130$ мм, $\delta c_5=0,01$ мм, $\delta c_6=0,05$ мм, при данных телеобъектива теодолита Т2 $f_1'=119,87$ мм, $f_2'=-47,97$ мм, $L=147,11$ мм и размере пикселя 9,6 мкм суммарная погрешность на расстоянии фокусировки 2400 мм составит 0,292 мм, при этом она будет возрастать с увеличением дистанции фокусировки (рис. 2).

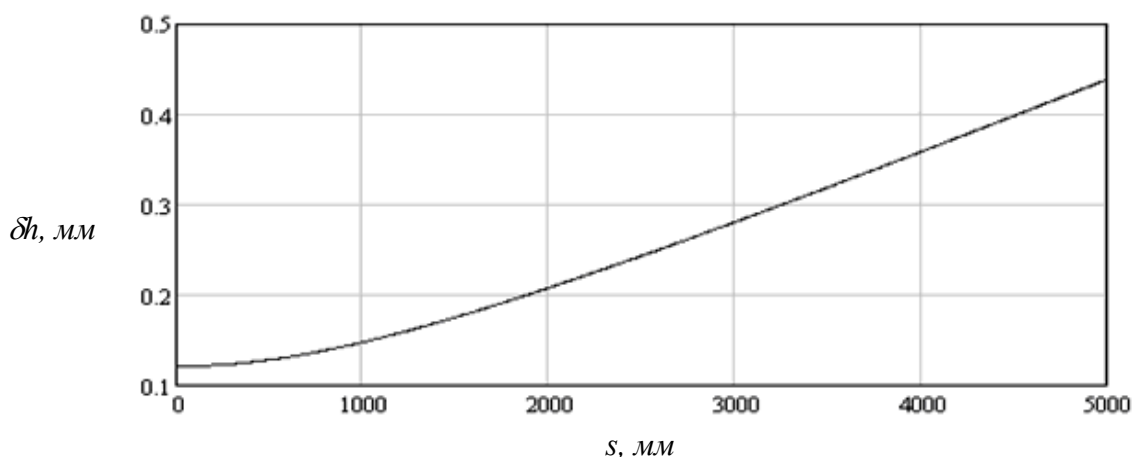


Рис. 2. Теоретическая зависимость погрешности от дистанции

Для экспериментальной проверки влияния рассмотренных погрешностей на результаты измерений с помощью оптико-электронной системы контроля соосности (ОЭСКС) использовался стенд, включающий в себя оптическую скамью типа ОСК-2 6 (рис. 3) со специальными штативами 5 и 7, в которые были закреплены исследуемые части системы. ОЭСКС. В состав ОЭСКС, построенной по авторефлексионной схеме, входят ЭВМ 1 связанного с ней блока управления 2, базовый блок и трипельпризмы 4.

Штативы 5 и 7 рис. 3 устанавливались в нулевое положение, после чего штатив 5 перемещался по оптической скамье, прямолинейность движения штатива контролировалась борштангой.

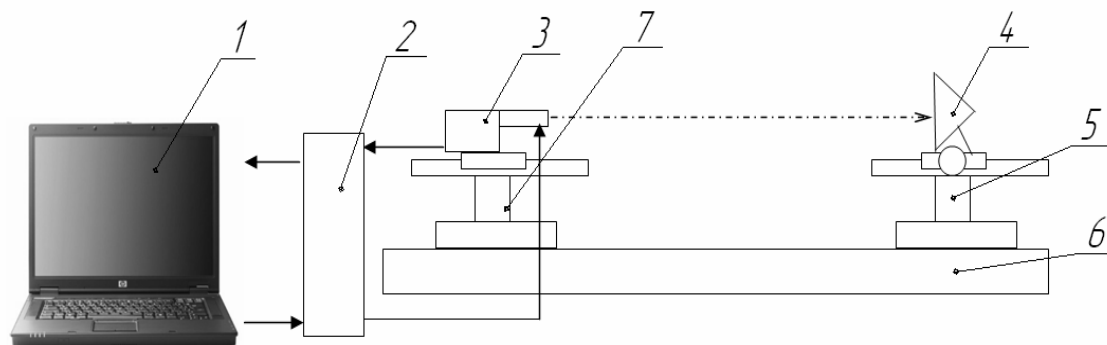


Рис. 3. Принципиальная схема экспериментального стенда

3. Основные результаты исследования

В результате эксперимента получено, что искривление оптической оси на расстоянии 2240 мм составило 0,212 мм, что говорит об отсутствии грубых ошибок в юстировке, однако необходимо более точное позиционирования. Однако такая погрешность оказалась значительно большей, чем допустимо условиями эксплуатации ОЭСКС, поэтому предлагается провести дополнительную доводку системы: в частности, исключить из работы трипельпризму, вынеся ПИД в плоскость измерения, внести дифференцированные поправки от дистанции в алгоритм расчета смещения.

Заключение

В работе были рассмотрены погрешности, связанные с установкой матрицы фотоприемника, смещением с оптической оси центра положительного и отрицательного компонентов телеобъектива, поворотом трипельпризмы, смещением центра базового расстояния ПИД с оптической оси и погрешностью определения энергетического центра изображения ПИД.

В дальнейшем планируется установка коэффициентов влияния отдельных погрешностей на суммарную с учетом технологического уровня, оптимизация оптической схемы с учетом полученных результатов, доводка ОЭСКС и улучшение точностных характеристик системы с использованием дифференцированных поправок.

Литература

1. Коротаев В.В., Тимофеев А.Н., Иванов А.Г. Проблемы разработки оптико-электронных систем для контроля деформаций крупногабаритных объектов // Оптический журнал. 2000. Т. 67. № 4. С. 43–46.
2. Витол Э.А., Иванов А.Г., Прокофьев А.В. Повышение точности распределенной оптико-электронной системы для измерения линейных смещений. / В кн.: Сб. трудов второй международной конференции молодых ученых и специалистов «Оптика-2001» 16–19 октября 2001 г. СПб: СПбГИТМО (ТУ), 2001. С. 228.
3. Кузнецов П.Н., Васютинский И.Ю., Ямбаев Х.К. Геодезическое инструментоведение: / Учебник для вузов. М.: Недра, 1984. 364 с.
4. Погарев Г.В. Юстировка оптических приборов. Л.: Машиностроение, Ленинградское отделение, 1982. 237 с.
5. Чуриловский В.Н. Теория оптических приборов. М.-Л.: Машиностроение, 1966. 564 с.

6. Захаров А.И. Новые теодолиты и оптические дальномеры. М.: Недра, 1978.
7. Пантюшин А.В., Шомрина М.А., Анисимов А.Г., Яковлев А.С. Исследование влияния погрешности установки отражателя на точность системы контроля соосности. / В кн.: Сб. трудов седьмой международной конференции «Прикладная оптика-2006» 16-20 октября 2006 г. СПб.: СПбГИТМО (ТУ), 2006.
8. Аникст Д.А., Голубовский О.М., Петрова Г.В., Фельдман Г.А. Оптические системы геодезических приборов. М. Недра, 1981. 240 с.
9. Соломатин В.А., Якушенков Ю.Г. Сравнение некоторых способов определения координат изображений, осуществляемых с помощью многоэлементных приемников излучения. // Известия вузов. Приборостроение. 1986. № 9. С. 62–69.
10. Парвалюсов Ю.Б., Радионов С.А., Салдатов В.А. и др., Проектирование оптико-электронных приборов. / Под ред. Якушенкова Ю.Г. М. Логос, 2000. 488 с.
11. Панов В.А., Кругер М.Я., Кулагин В.В. и др. Справочник конструктора оптико-механических приборов. / Под общ. ред. В.А. Панова. 3-е изд., перераб. и доп. Л.: Машиностроение, Ленингр. отд-ние, 1980, 742 с.

ИЗМЕРЕНИЕ ПОКАЗАТЕЛЯ ПРЕЛОМЛЕНИЯ ОДНОСЛОЙНОГО ПРОСВЕТЛЯЮЩЕГО ПОКРЫТИЯ

А.А. Немкова

Научный руководитель – д.т.н., профессор Э.С. Путилин

С помощью эллипсометрического и фотометрического методов выполнены измерения показателя преломления пленки на основе тетраэтоксисилана. Показано, что методом эллипсометрии возможно не только определить показатель преломления, но и выявить неоднородность оптического профиля.

Введение

Просветляющие покрытия находят широкое применение в областях, где требуется увеличение пропускания или снижение потерь на отражение от поверхности материала. В солнечных элементах нанесение просветляющего покрытия позволяет улучшить эффективность преобразования солнечного излучения. Использование просветляющих покрытий в лазерных системах увеличивает мощность выходящего излучения, обеспечивая при этом высокую лазерную прочность элементов.

Для получения однослойного просветляющего покрытия показатель преломления материала пленки должен быть предельно малым. Нулевое отражение достигается при показателе преломления пленки, равном квадратному корню из показателя преломления подложки. Пленкообразующих материалов, удовлетворяющих данному требованию, для подложек, прозрачных в видимой области спектра, не существует. Следовательно, задача получения просветляющего покрытия не может быть успешно решена без точной информации о его показателе преломления. Наиболее чувствительные методы определения оптических постоянных тонких слоев на поверхности стекла и других прозрачных сред – эллипсометрия и спектрофотометрия – основаны на анализе света, отраженного от поверхности с пленкой.

Целью данной работы является измерение показателя преломления просветляющей пленки двумя методами и сравнение полученных результатов.

1. Методы измерения показателя преломления тонких пленок

1.1. Эллипсометрия

Сущность эллипсометрического метода измерений состоит в исследовании изменения состояния поляризации пучка света в результате его отражения от изучаемого объекта. Эллипсометрические параметры Δ и ψ связаны основным уравнением

$$\operatorname{tg} \psi e^{i\Delta} = \frac{R_p}{R_s}, \quad (1)$$

где R_p и R_s – полные комплексные амплитудные коэффициенты отражения для p - и s -поляризаций [1]. Таким образом, $\operatorname{tg} \psi$ является соотношением амплитуд, а Δ – разностью фаз компонент отраженного света. Поскольку измеряется отношение двух величин, а не их абсолютное значение, то на результат не влияет изменение интенсивности источника света и поглощение воздуха.

Эллипсометрия не является прямым методом измерения. Для получения значений оптических констант пленки (n и k) необходимо использование модели, включающей начальные и граничные значения искомых параметров. Поиск осуществляется путем минимизации функции качества F по методу наименьших квадратов.

В многоугловой эллипсометрии используется монохроматический источник света. Обычно это лазер, длина волны излучения которого лежит в видимой области спектра. В этом случае измерение эллипсометрических параметров осуществляется при нескольких углах падения света. Спектральная эллипсометрия [2] позволяет измерять оп-

тические константы на любой длине волны в ультрафиолетовой, видимой и инфракрасной областях спектра. Кроме того, за счет большого количества независимых данных возможно измерение толщин и структуры каждого слоя многослойной системы, степень их кристалличности, наличие поверхностного примесного слоя.

Одним из способов получения пленки с предельно низким показателем преломления является образование внутри ее пористой структуры. Пористость может быть обусловлена неплотной упаковкой, газопоглощением и газовыделением, химическими реакциями, микроусадкой и др. Адсорбционно-эллипсометрический метод [3], заключающийся в регистрации объема адсорбирующегося вещества, позволяет измерять пористость пленок.

1.2 Спектрофотометрия

Спектрофотометрический метод определения оптических характеристик прозрачных пленок на поверхности прозрачной подложки основан на измерении спектральных значений коэффициента отражения R_λ для ряда длин волн выбранного участка спектра. Для просветляющих пленок измерение коэффициента отражения дает более точный результат, чем измерение коэффициента пропускания, кроме того, на результат не влияет более толстая подложка.

Показатель преломления определяется по величине экстремума на спектральной кривой коэффициента отражения. В случае просветляющей пленки измеряется минимальный коэффициент отражения, определяемый выражением

$$R_{\min} = \left(\frac{n_1^2 - n_0 n_2}{n_1^2 + n_0 n_2} \right)^2, \quad (2)$$

где n_0 – показатель преломления воздуха, n_1 – показатель преломления пленки, n_2 – показатель преломления подложки. Разница значений R_{\min} на кривой спектрального отражения позволяет судить о дисперсии пленкообразующего вещества.

Для многослойных покрытий определение оптических постоянных с помощью спектрофотометрического метода требует специальных способов расчета с использованием моделей покрытий.

2. Результаты измерений

Образцом для измерений служило однослойное покрытие, полученное золь-гель методом из 10 %-го раствора тетраэтоксисилана на стеклянной подложке с показателем преломления $n = 1,514$. Для измерения показателя преломления использовался эллипсометр ЛЭФ-3М, а в качестве источника излучения – гелий-неоновый лазер с длиной волны $\lambda = 632,8$ нм. Сначала для измерений была выбрана модель однородной пленки, которая предполагает постоянство показателя преломления по толщине пленки. Измеренное значение составило 1,43.

Пленки, полученные из раствора тетраэтоксисилана, обладают пористой структурой и зачастую неравномерным распределением частиц SiO_2 по толщине. Для более детальной характеристики оптического профиля использовалась экспоненциальная модель с изгибом у подложки. В данном случае профиль аппроксимируется системой из десяти однородных пленок, для каждой из которых определяются оптические константы. Результат измерения исследуемой пленки представлен на рис. 1. Показатель преломления пленки на границе с воздухом имеет минимальное значение, постепенно увеличиваясь по мере приближения к подложке.

Спектральный коэффициент отражения определялся на микроскопическом спектрофотометре МСФУ-ЭВМ в интервале длин волн от 500 – 800 нм (рис. 2). Минимальное значение коэффициента отражения составило 2,31 %. Соответствующий показатель преломления пленки, рассчитанный на основе формулы (2), имеет значение

1,434, которое в пределах экспериментальной погрешности согласуется с результатом эллипсометрических измерений.

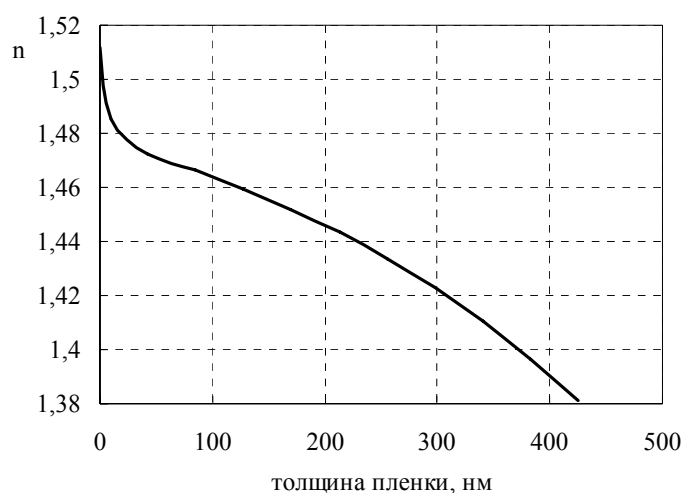


Рис. 1. Оптический профиль просветляющего покрытия

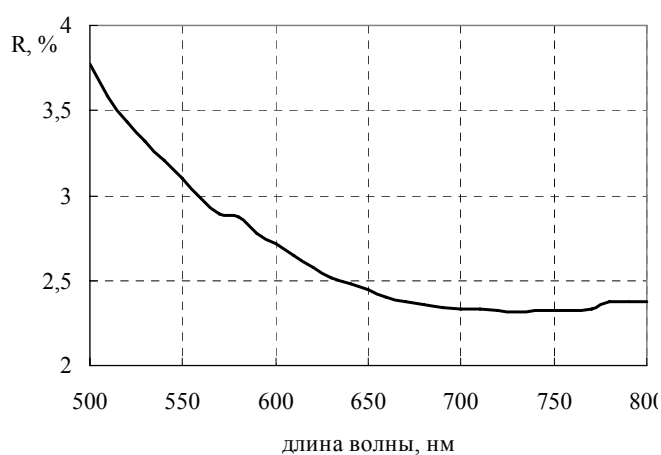


Рис. 2. Спектральный коэффициент отражения покрытия

Заключение

Методы эллипсометрии и спектрофотометрии позволяют бесконтактным неразрушающим способом определять показатель преломления тонких пленок. Результаты эллипсометрии зависят от используемой модели, значений начальных и граничных параметров. В то же время с помощью эллипсометрического метода возможно получение информации о неоднородном оптическом профиле. Спектрофотометрия дает однозначный результат показателя преломления пленки. Поскольку измеренные значения находятся в соответствии друг с другом, то достаточно использования лишь одного из методов. Для уточнения результата измерения и увеличения его достоверности предпочтительнее использовать оба метода определения показателя преломления.

Литература

1. Азам Р., Башара Н. Эллипсометрия и поляризованный свет. М.: Мир, 1981. 252 с.
2. Vedam K. Spectroscopic ellipsometry: a historical overview. // Thin Solid Films, 1998. Vol. 313–314. P.1–9.
3. Толмачев В.А. Адсорбционно-эллипсометрический метод исследования оптического профиля, толщины и пористости тонких пленок. // Оптический журнал. 1999. Т. 66. №7. С. 20–34.

АНАЛИЗ СХЕМ ПОСТРОЕНИЯ СИСТЕМ ИЗМЕРЕНИЯ ПАРАМЕТРОВ УГЛОВОЙ ПРОСТРАНСТВЕННОЙ ОРИЕНТАЦИИ НА ОСНОВЕ ЯВЛЕНИЯ АНАМОРФОЗЫ

А.Д. Мерсон

Научный руководитель – д.т.н., профессор И.А. Коняхин

Углоизмерительные системы широко применяются для контроля взаимного положения объектов при решении измерительных задач в промышленности и научно-технической деятельности. Перспективным является построение измерительной системы на основе явления анаморфозы. Недостатком существующей системы является невозможность измерения коллимационных углов. Автором предложена улучшенная реализация, позволяющая измерять все три угловые координаты контролируемого объекта.

Введение

Измерение угловых смещений объектов требуется при решении многих прикладных задач. В приборостроении – это контроль геометрических параметров изделий и их пространственного положения, в строительстве и машиностроении – контроль деформаций отдельных элементов и конструкции в целом, в робототехнике и навигации – определение положения контролируемых объектов в некой фиксированной системе координат. Наиболее эффективными являются бесконтактные оптико-электронные измерительные средства.

Угловая пространственная ориентация контролируемого объекта определяется тремя координатами: углом скручивания (угол поворота относительно линии визирования объекта) и двумя коллимационными углами. Перспективным является построение измерительной системы на основе явления анаморфозы, использование которого позволяет увеличить диапазон измерения угла скручивания [1]. Известные устройства такого типа являются однокоординатными и не позволяют измерять коллимационные углы. Автором предлагается модификация анаморфозной измерительной системы, позволяющая определять все три угловые координаты.

Система измерения угла скручивания на основе явления анаморфозы

Однокоординатная анаморфозная измерительная система [1] состоит из автоколлиматора, установленного на жесткой базе, и отражательного контрольного элемента (КЭ), расположенного на контролируемом объекте (рис. 1).

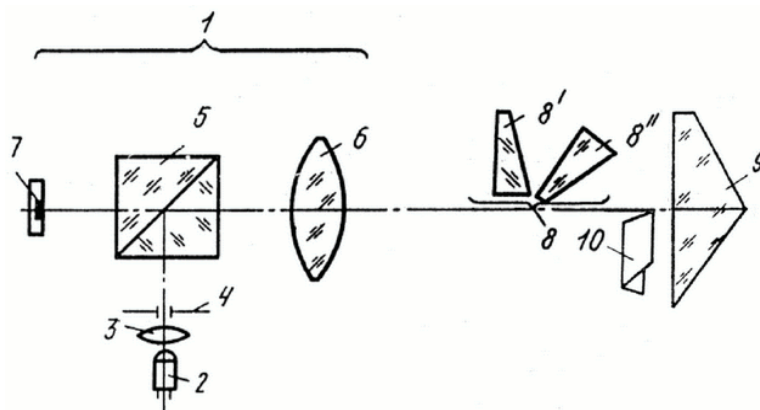


Рис. 1. Оптическая схема системы измерения угла скручивания, основанной на явлении анаморфозы

Автоколлиматор 1 включает в себя источник излучения 2, марку 4, установленную в фокальной плоскости объектива 6, и светоделитель 5. Приемный канал состоит

из ПЗС-матрицы 7 и цифровой системы обработки изображения (на рисунке не показана). КЭ состоит из анаморфозной системы 8, ретрорефлектора (трипель-призмы) 9 и дополнительного отражателя (ДО) 10.

Параллельный пучок лучей, формируемый объективом 6 автоколлиматора, проходит анаморфозную телескопическую систему 8, затем ретрорефлектором 9 направляется на дополнительный отражатель 10, представляющий собой двугранное зеркало. Двугранное зеркало отражает пучок, одновременно разворачивая его на 90 градусов в плоскости, перпендикулярной оптической оси, после чего он вторично проходит ретрорефлектор и анаморфозную систему. Такой разворот с повторным прохождением оптических элементов позволяет увеличить коэффициент анаморфозы и, следовательно, повысить чувствительность системы [1].

Объектив 6 автоколлиматора 1 формирует изображение марки 4, трансформированное анаморфозной системой 3, на анализаторе – ПЗС-матрице 7.

Поворот контролируемого объекта на угол скручивания вследствие соответствующего разворота главного сечения анаморфозной телескопической системы приводит к изменению формы и размера изображения марки на ПЗС-матрице, что и позволяет определить угол скручивания.

Для измерения коллимационных углов предлагается использовать свойства дополнительного отражателя 10. Проанализируем различные варианты его конфигурации.

Сравнительный анализ ДО в виде двугранного зеркала и косоугольного отражателя

Орт \mathbf{B} пучка после отражателя находим из соотношения [2]:

$$\mathbf{B} = M_r M_d M_r^{-1} \mathbf{A}, \quad (1)$$

где \mathbf{A} – орт падающего на отражатель пучка; M_d – матрица действия дополнительного отражателя в исходном положении; M_r и M_r^{-1} – матрицы прямого и обратного преобразований координат, соответственно, описывающие поворот отражателя из исходного положения в текущее. Для малых углов матрица преобразования координат в общем виде записывается следующим образом:

$$M_r = \begin{pmatrix} 1 - \frac{1}{2}(\Theta_2^2 + \Theta_3^2) & -\Theta_3 + (1-w)\Theta_1\Theta_2 & \Theta_2 + (1-p)\Theta_1\Theta_3 \\ \Theta_3 + w\Theta_1\Theta_2 & 1 - \frac{1}{2}(\Theta_1^2 + \Theta_3^2) & -\Theta_1 + (1-v)\Theta_2\Theta_3 \\ -\Theta_2 + p\Theta_1\Theta_3 & \Theta_1 + v\Theta_2\Theta_3 & 1 - \frac{1}{2}(\Theta_1^2 + \Theta_2^2) \end{pmatrix}, \quad (2)$$

где $\Theta_1, \Theta_2, \Theta_3$ – углы трех последовательных поворотов отражателя относительно осей связанной с ним системы координат; p, v, w – параметры, значения которых определяются конкретной последовательностью поворотов и приведены в таблице [2].

Последовательность поворотов	Значения параметров		
	p	v	w
$\Theta_1 - \Theta_2 - \Theta_3$	1	1	1
$\Theta_3 - \Theta_1 - \Theta_2$	0	0	1
$\Theta_2 - \Theta_3 - \Theta_1$	0	1	0
$\Theta_3 - \Theta_2 - \Theta_1$	0	0	0
$\Theta_2 - \Theta_1 - \Theta_3$	1	1	0
$\Theta_1 - \Theta_3 - \Theta_2$	1	0	1

Таблица. Значения параметров матрицы преобразования координат

Матрица действия двугранного зеркала имеет вид

$$M_d = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (3)$$

Орт падающего пучка при осевом ходе лучей

$$\mathbf{A} = \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix}. \quad (4)$$

Подставив (2), (3), (4) в (1) и выполнив матричное умножение, получим:

$$\mathbf{B} = \begin{pmatrix} \Theta_2 - \Theta_1 - (1+\nu)\Theta_2\Theta_3 + (1-p)\Theta_1\Theta_3 \\ \Theta_2 - \Theta_1 + (1-\nu)\Theta_2\Theta_3 - (1+p)\Theta_1\Theta_3 \\ 1 - (\Theta_1 - \Theta_2)^2 \end{pmatrix}. \quad (5)$$

Из анализа выражения (5) следует, что координаты орта \mathbf{B} не позволяют однозначно определить коллимационные углы Θ_1, Θ_2 . Таким образом, при использовании в качестве ДО двугранного зеркала измерение коллимационных угловых смещений контролируемого объекта невозможно.

Предлагается использовать вместо двугранного зеркала косоугольный уголкового отражатель (рис. 2).

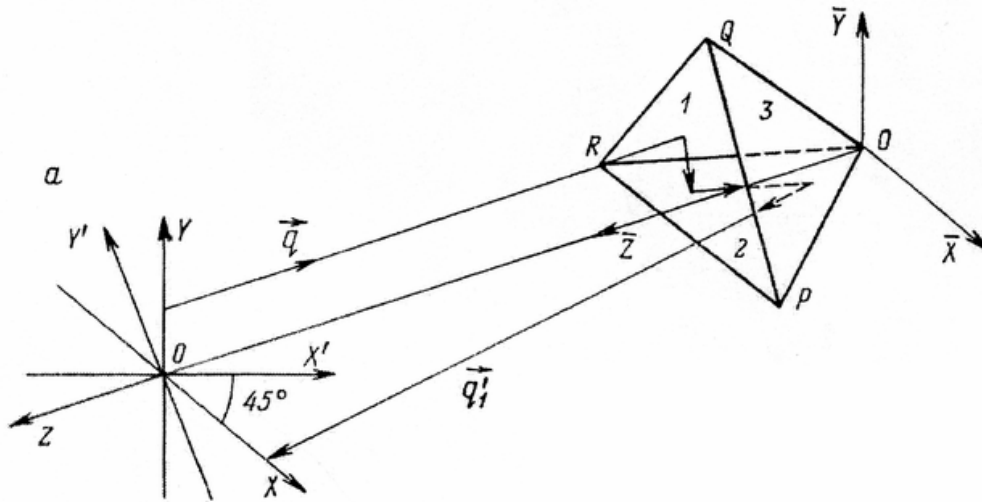


Рис. 2. Косоугольный уголкового отражатель

Двугранные углы между гранями с номерами 1, 2 и 2, 3 равны 60° , а между гранями 1, 3 – 120° . Отражатель ориентируется относительно системы координат $O\bar{X}\bar{Y}\bar{Z}$ так, что фронтальная грань PQR лежит в плоскости $\bar{X}O\bar{Y}$, а ребро двугранного угла, имеющего величину 120° , – в плоскости $\bar{Y}O\bar{Z}$ [2].

Матрица действия отражателя имеет вид:

$$M_d = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (6)$$

Подставив аналогично предыдущим расчетам (6), (2), (4) в (1), получим:

$$\mathbf{B} = \begin{pmatrix} \Theta_2 + \Theta_1 - (1-\nu)\Theta_2\Theta_3 + (1-p)\Theta_1\Theta_3 \\ \Theta_2 - \Theta_1 + (1-\nu)\Theta_2\Theta_3 + (1+p)\Theta_1\Theta_3 \\ 1 - \Theta_1^2 - \Theta_2^2 \end{pmatrix}. \quad (7)$$

Очевидно, что влияние угла скручивания Θ_3 на коллимационные углы отсутствует при значениях параметров $p=1, v=1$, что соответствует последовательностям поворотов отражателя $\Theta_1-\Theta_2-\Theta_3$ и $\Theta_2-\Theta_1-\Theta_3$, заданных в строках 1, 5 таблицы, соответственно. Взаимное влияние коллимационных углов на координаты пучка устраняется поворотом на 45° системы координат, связанной с анализатором ($OX'Y'Z$ на рис. 2).

Координаты орта пучка после поворота анализатора имеют вид:

$$\mathbf{B} = \begin{pmatrix} \sqrt{2}\Theta_1 \\ \sqrt{2}\Theta_2 \\ 1 - \Theta_1^2 - \Theta_2^2 \end{pmatrix}. \quad (8)$$

Выражение (8) позволяет однозначным образом определить коллимационные углы Θ_1, Θ_2 . Таким образом, использование предложенного отражателя вместо двугранного зеркала позволяет реализовать трехкоординатную углоизмерительную систему.

4. Заключение

Одним из перспективных методов измерения пространственного смещения объектов является метод на основе явления анаморфозы. Существующая реализация метода позволяет определить только угол скручивания, в то время как в большинстве практических применений необходимо также измерять коллимационные углы. Автором предложено в качестве дополнительного отражателя вместо двугранного зеркала использовать уголкового косоугольный отражатель. Такая модификация дает возможность измерять коллимационные углы, что позволяет реализовать трехкоординатную углоизмерительную систему.

Литература

1. Зубенко Д.Ю. Исследование отражателя с анаморфозным элементом для автоколлимационного угломера. // Изв. вузов. Приборостроение. 1993. №3. С.76–80.
2. Бондаренко И.Д. Принципы построения фотоэлектрических автоколлиматоров. Мн.: Изд-во «Университетское», 1984. 190 с.
3. Джабиев А.Н., Коняхин И.А., Панков Э.Д. Автоколлимационные углоизмерительные средства мониторинга деформаций. СПб: СПбГИТМО (ТУ), 2000. 197 с.

АНАЛИЗ РЕГУЛЯРНОГО ВИНЬЕТИРОВАНИЯ В АВТОРЕФЛЕКСИОННЫХ УГЛОИЗМЕРИТЕЛЬНЫХ СИСТЕМАХ

А.М. Ворона

Научный руководитель – д.т.н., профессор И.А. Коняхин

В работе показано, что в авторефлексионной углоизмерительной системе во всем диапазоне измерений наблюдается влияние виньетирования. Такое виньетирование носит регулярный характер и может быть описано теоретически и затем скомпенсировано на этапе измерений. Произведен теоретический анализ виньетирования, рассмотрена зависимость функции виньетирования от строения оптического пучка и решена задача для простого случая сечения пучка.

Введение

В процессе точного монтажа крупногабаритных объектов, например, частей фюзеляжа самолета, стеновых панелей, мостовых опор, монтажная операция состоит из двух этапов – предварительной установки монтируемого блока и последующего этапа точной установки блока на место. На этапе предварительной установки монтируемого блока необходим прибор, обладающий широким диапазоном измерения, при этом его точность может быть хуже, чем требуемая для итоговой установки блока. На втором этапе необходим прибор с меньшим диапазоном, точность которого должна обеспечивать требуемое качество монтажа блока. В настоящее время на разных этапах монтажа используются различные приборы [1–3]. Принципиально возможно использование единой измерительной системы на основе автоколлимационного метода. Однако недостатком таких систем является снижение точности измерения при увеличении диапазона. Это объясняется влиянием так называемой погрешности измерения вследствие виньетирования оптического пучка.

Влияние виньетирования проявляется во всем диапазоне измерений (кроме нулевого положения), носит регулярный характер и, следовательно, может быть описано теоретически и затем скомпенсировано на этапе измерений.

Проанализируем процесс виньетирования и определим зависимость функции виньетирования от строения оптического пучка.

Описание процесса виньетирования

Виньетирование оптических пучков, формирующих изображение марки в плоскости анализа, имеет место при любом строении и положении контролируемого объекта. Но при «нулевом» положении контролируемого объекта виньетирование носит симметричный относительно изображения марки характер и не влияет на результат измерения. При повороте контролируемого элемента виньетирование становится несимметричным. Наиболее виньетированный и наименее виньетированный пучок находятся на линии, совпадающей с направлением поворота контрольного элемента (перпендикулярном оси поворота). Такая асимметричность приводит к незначительному искажению формы изображения марки и может быть названа «регулярным» виньетированием.

При достижении значения угла поворота некоторого порогового значения Θ_{\max} наблюдается полное виньетирование внеосевых пучков. Значение порогового угла поворота определяется как

$$\Theta_{\max} = \frac{D_{об}}{4L},$$

где $D_{об}$ – диаметр объектива, L – расстояние до контрольного элемента (КЭ). Такое виньетирование приводит к значительному искажению формы изображения марки и поэтому может быть названо «радикальным» виньетированием. Такое виньетирование рассмотрено в работах [4, 5].

Данная работа посвящена рассмотрению процесса регулярного виньетирования.

Общая теория регулярного виньетирования

Удобно рассматривать функцию виньетирования, описывающую относительную величину виньетирования оптического пучка в зависимости от угла поворота КЭ Θ и угла между осевым и рассматриваемым пучками β (рис. 1). Функция виньетирования $\Psi(\beta, \Theta)$ может быть определена как отношение эффективных площадей зрачка для внеосевых оптических пучков в направлении и против направления поворота КЭ (см. рис. 2).

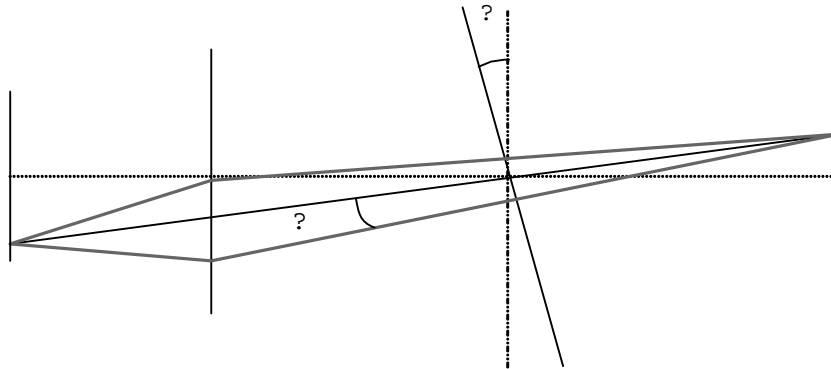


Рис. 1. Ход лучей в авторефлексионной схеме

В качестве меры искажения изображения марки в целом используется коэффициент виньетирования

$$d(\Theta) = \frac{\int (1 - \Psi(\beta, \Theta)) d\beta}{\int \Psi(\beta, \Theta) d\beta}.$$

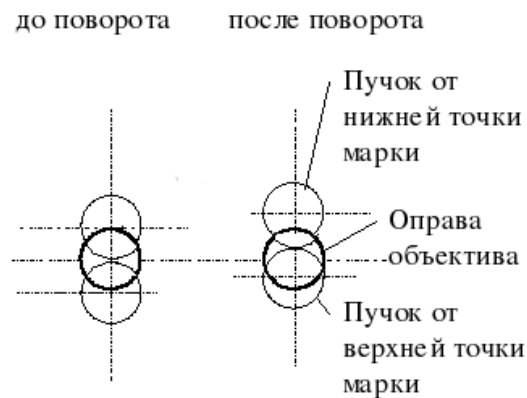


Рис. 2. Иллюстрация процесса регулярного виньетирования

Заметим, что вследствие симметричного строения оптического пучка при отсутствии виньетирования коэффициент виньетирования равен нулю, тогда как при максимальном регулярном виньетировании он равен единице.

Так как при малых углах поворота КЭ (единицы градусов) зависимость светового диаметра от угла поворота для каждого элементарно оптического пучка имеет линейный вид и является нечетной функцией, то функция виньетирования тривиально связана с функцией, описывающей распределение энергии в пучке. Так как в авторефлексионной схеме плоскость анализа удалена от источника на расстояние, большее дистанции формирования пучка, то распределение энергии в пучке будет описываться ФРТ системы. Таким образом, функция виньетирования тривиально связана с ФРТ системы.

Искажение формы изображения марки в плоскости анализа описывается нормированной функцией виньетирования $\psi(x, \varepsilon)$ (где x – линейная координата в плоскости анализа в направлении поворота КЭ), что проиллюстрировано рис. 3.

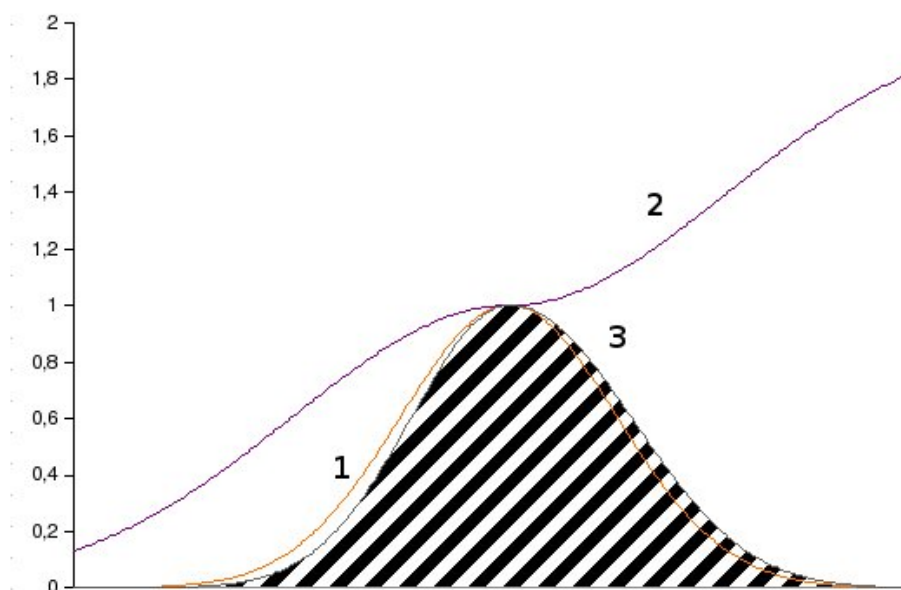


Рис. 3. Искажение формы изображения марки в плоскости анализа.
1 – форма изображения при нулевом положении; 2 – нормированная функция виньетирования при повороте КЭ; 3 – форма изображения при отклонении КЭ

Исходя из симметричности процесса регулярного виньетирования, физического смысла и того, что $x_{\max} = \Theta_{\max} F = \frac{D_{\text{об}} F}{4L}$, на нормированную функцию виньетирования могут быть наложены следующие граничные условия:

1. $\psi(x = 0, \varepsilon) \equiv 1$
 $\psi(x, \varepsilon = 0) \equiv 1$ (условие нормировки)
2. $\psi(x, \varepsilon) - 1 = -(\psi(-x, \varepsilon) - 1)$ (условие симметричности)
3. $\psi(x = -\infty, \varepsilon) = 0$
 $\psi(x = \infty, \varepsilon) = 2$ (условие сохранения энергии)
4. $\psi\left(x = \frac{D_{\text{об}} \cdot F}{4L}, \varepsilon\right) = 1 - \varepsilon$ (граничное условие регулярного виньетирования),

где F – фокусное расстояние объектива.

Из геометрических соображений можно увидеть полиномиальную зависимость $\varepsilon(\Theta) \sim P(\sin(\Theta))$.

Таким образом, зная ФРТ системы (в первом приближении можно использовать функцию $E_0(x, y)$ – распределение освещенности в плоскости анализа при отсутствии поворота КЭ), метод определения координат центра изображения по распределению освещенности изображения $E(x, y)$ и основные характеристики авторефлексионной измерительной системы, инженер-проектировщик может рассчитать и учесть влияние регулярного виньетирования.

Моделирование системы на основе линейной ФРТ

Рассмотрим первое приближение – линейная ФРТ вида $E_0(x) = 1 - |x|$. В этом случае функция $\psi(x, \varepsilon)$ линейна относительно x и, с учетом граничных условий, имеет вид

$$\psi(x, \varepsilon) = 1 - 4 \frac{L\varepsilon}{D_{об} \cdot F} x.$$

Обычным способом определения координат центра изображения является центровзвешенное суммирование

$$x_{изм} = \frac{\int x \cdot E(x) dx}{\int E(x) dx}.$$

Интегрируя это выражение с подстановкой $E(x) = E_0(x) \cdot \psi(x, \varepsilon)$, получаем погрешность измерения вследствие виньетирования

$$\sigma_B = \frac{2L\varepsilon}{3D_{об} \cdot F} = \frac{2L}{3D_{об} \cdot F} \sin^2 \Theta.$$

Экспериментальное исследование

Для проверки данных, полученных путем теоретических выкладок, а также для выработки практических рекомендаций был произведен эксперимент. Экспериментальная установка была построена в соответствии с авторефлексионным методом измерения (рис. 4). Структурная схема включает источник излучения (1), изображение (1') источника зеркалом (2), объектив (3), матричный фотоприемник (4) и блок компьютерной обработки кадра (5).

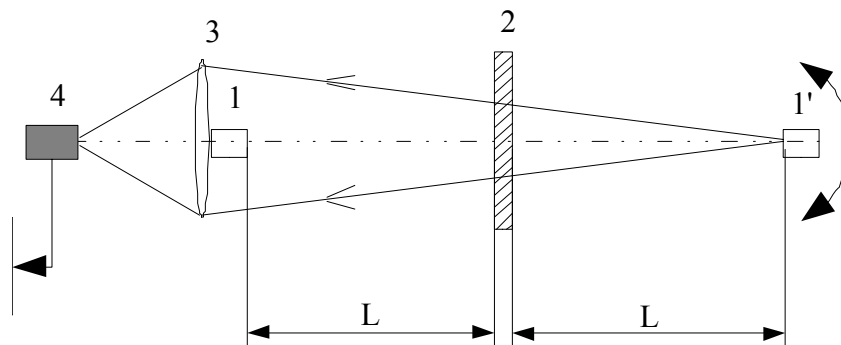


Рис. 4. Структурная схема авторефлексионной системы

Особенностью экспериментальной системы является использование специально разработанного адаптивного программного обеспечения. Используемое программное средство позволяет экспериментатору задавать параметры пороговой и медианной фильтрации, алгоритм первичного поиска изображения и параметры алгоритма взвешенного суммирования вплоть до написания дополнительных этапов обработки, не предусмотренных авторами. Результаты эксперимента представлены на рис. 5.

Экспериментальные исследования показали, что статическая характеристика исследуемой системы линейна в диапазоне измерения ± 10 угловых минут. Оценка среднеквадратического значения случайной составляющей погрешности измерения не превосходит 0.35 угл.сек. Систематическая погрешность измерения определяется виньетированием пучка, формирующего изображение светодиода и при отсутствии компенсации составляет величину до 25 угловых секунд для крайних точек указанного диапазона измерения (рис. 6).

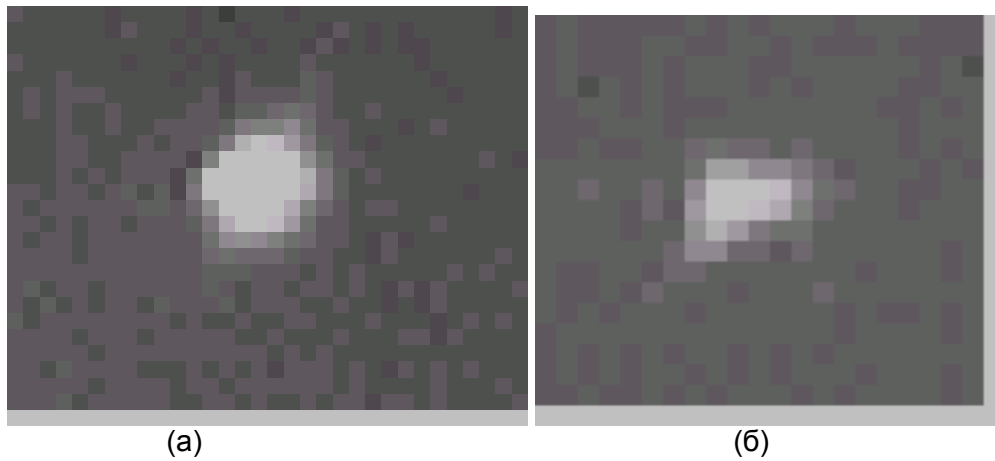


Рис. 5. Вид изображения в плоскости анализа: полное (а) и виньетированное (б) изображения светодиода

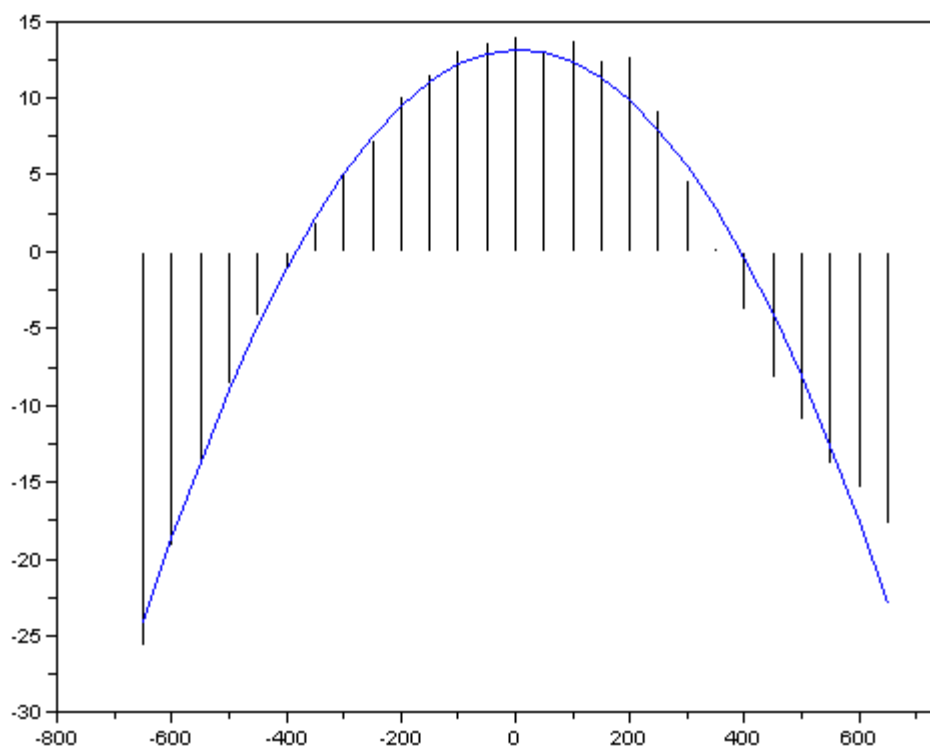


Рис. 6. Значения ошибки измерения угла поворота объекта и ее квадратичная аппроксимация

Среднее значение ошибки измерения угла поворота объекта при учете влияния виньетирования составляет 1.34 угл. сек. Используя для компенсации погрешности полученные теоретическим путем функции, можно дополнительно уменьшить ошибку измерения угла поворота объекта до уровня случайной составляющей погрешности.

Заключение

1. Вид внесенной виньетированием погрешности позволяет отнести ее к систематическим погрешностям.
2. Внесенная виньетированием погрешность может быть скомпенсирована на этапе компьютерной обработки кадра.

3. Полученное теоретическое описание хорошо согласуется с имеющимися экспериментальными данными, хотя еще требует проведения дополнительных уточняющих экспериментов.

Литература

1. Высокоточные угловые измерения. / Под ред. Ю.Г. Якушенкова. М.: Машиностроение, 1987. 480 с.
2. Голубовский Ю.М. Фотоэлектрические автоколлиматоры. // Оптико-механическая промышленность. 1970. № 5.
3. Мусяков В.Л., Панков Э.Д. К вопросу об использовании уголкового отражателя для измерения скручивания. / В сб.: Оптико-электронные приборы в контрольно-измерительной технике. Под ред. проф. Л.Ф. Порфирьева. Труды ЛИТМО, вып. 81, 1975.
4. Коняхин И.А., Панков Э.Д. Трехкоординатные оптические и оптико-электронные угломеры. Справочник. М.: Недра, 1991. 224 с.
5. Ворона А.М., Лю Лэй. Экспериментальное исследование погрешности авторефлексионных измерений вследствие виньетирования. / III межвузовская конференция молодых ученых. СПб: СПбГУ ИТМО, 2006.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ТЕХНОЛОГИЧЕСКИХ ПАРАМЕТРОВ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ ПРИЗМЕННОЙ АНАМОРФОТНОЙ СИСТЕМЫ

Ю.К. Пруненко

Научный руководитель – д.т.н., профессор Л.Н. Андреев

В статье приведены результаты исследования влияния формы рабочих поверхностей, клиновидности, наклонов рабочих поверхностей и призмённых блоков на качество изображения ахроматической афокальной анаморфотной призмённой системы лазерного осветителя с коэффициентом анаморфозы 0,3.

Введение

Призмённые анаморфотные системы несколько отличаются от привычных для нас осесимметричных оптических систем. Это не осесимметричные системы, которые позволяют получать изображения с различным масштабом в двух взаимно перпендикулярных направлениях. Призмённые анаморфотные системы нашли применение в полиграфии, кинематографии, голографии и в других областях [1]. В настоящее время анаморфотные системы, в том числе и призмённые, стали применять для исправления расходимости излучения лазерных диодов [2].

В процессе изготовления любой спроектированной оптической системы происходит изменение конструктивных параметров, децентрировок и наклонов компонентов системы относительно расчетных, что ухудшает качество изображения. Для того чтобы реальная оптическая система имела заданные в техническом задании характеристики, необходимо провести расчет допусков, т.е. знать допустимые смещения и наклоны компонентов системы и точность изготовления поверхностей.

Призмённые системы относятся к классу пространственных систем, и расчет допусков является сложной многофакторной задачей. Вследствие этого необходимо исследовать влияние параметров на качество изображения призмённого анаморфота. Результатом такого исследования будет назначение допусков на изготовление призмённой системы с коэффициентом анаморфозы 0,3 [2].

Исследование влияния технологических параметров

На качество изображения призмённой анаморфотной системы, состоящей из преломляющих призм, влияют такие технологические параметры, как форма рабочих поверхностей, точность преломляющих углов при вершине и клиновидность призм. Наклоны призмённых блоков приводят к изменению коэффициента анаморфозы системы и вносят хроматизм положения [3].

Призмённая анаморфотная система с коэффициентом анаморфозы 0,3, показанная на рисунке, имеет дифракционное качество изображения [2]. Вследствие изменения волновой аберрации при оценке качества изображения должно быть не более четверти длины волны. Изменение размера светового пятна при смене лазерных диодов не должно превышать 5 %, но у призмённого анаморфота изменение уже составляет 3,5 %, что определяется хроматическими аберрациями [2]. Поэтому вторым критерием оценки качества является величина хроматических аберраций, которая не должна превышать 1 секунды (допустимы 1,5 % изменения размера светового пятна при смене лазерных диодов). Коэффициент же анаморфозы может изменяться в пределах 5 % от величины анаморфозы, что составляет 0,02.

Исследование оптической призмённой системы проводилось в программе OPAL. При изменении толщин призм и воздушного промежутка между призмёнными блоками

коэффициент анаморфозы и aberrации системы не изменяются. Был назначен допуск на воздушный промежуток ± 5 мм.

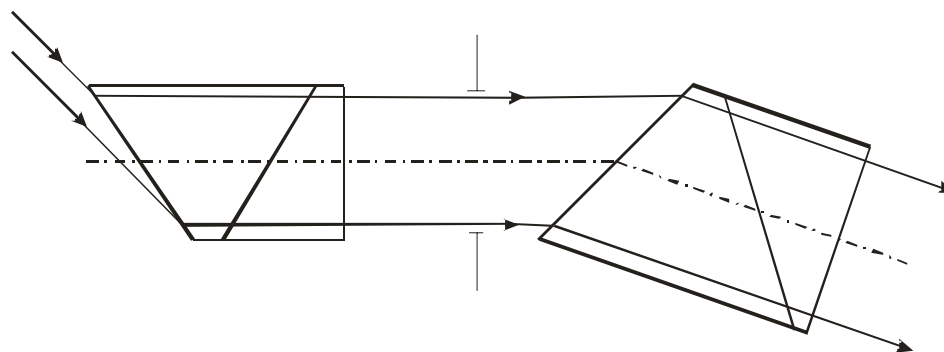


Рис. Оптическая схема призмной афокальной анаморфотной насадки в плоскости YZ

Требования к форме плоских поверхностей (допуск на стрелку кривизны – N) призм определялись в зависимости от допустимого изменения волновой aberrации при задании радиусов кривизны каждой плоской поверхности. Допуск на стрелку кривизны N работающих поверхностей оптических деталей измеряют числом интерференционных колец или полос, укладываемых на наименьшем размере детали при контроле пробным стеклом. Изменения волновой aberrации системы в зависимости от формы плоских поверхностей призм приведены в табл. 1.

№ пов.	Кол-во колец N	Радиус кривизны, мм	Изменение волновой aberrации, дл. волн	
			ΔW_{01}	ΔW_{02}
1	1	219780	-0,544	-1,713
		-219780	0,544	1,709
2	1	219780	0,007	-0,093
		-219780	-0,007	0,093
3	1	179487	0,965	-1,251
		-179487	-0,965	1,252
4	1	219780	-0,807	0,500
		-219780	0,805	-0,500
5	1	219780	0,028	-0,155
		-219780	-0,028	0,155
6	1	219780	0,191	0,259
		-219780	-0,191	-0,259

Таблица 1. Влияние формы поверхностей на волновую aberrацию системы

Из табл. 1 видно, что наиболее сильное влияние оказывают поверхности, граничащие с воздухом. На эти поверхности необходимо назначить более жесткий допуск. На поверхности, по которым склеиваются призмные блоки, можно назначить более свободный допуск.

Точность преломляющих углов при вершине призм определялась из влияния наклона каждой работающей поверхности на aberrации всей системы. Наклоняя каждую поверхность призмы на угол $\Delta \nu$ в одну и другую сторону в плоскости YZ, оценивалось изменение коэффициента анаморфозы и aberrаций системы. Значения влияния наклонов поверхностей на aberrации системы приведены в табл. 2.

№ пов.	Наклон пов-ти. Δv , сек	ΔW , дл.волн	$\Delta \sigma' y$, сек	ΔA
1	+30	0,091	0,2	0,0002
	-30	-0,09	-0,2	-0,0002
2	+30	0,169	-0,5	0
	-30	-0,168	0,5	0
3	+30	0,034	-0,1	0
	-30	-0,033	0,1	0
4	+15	-0,331	0,9	0
	-15	0,330	-0,9	0
5	+30	0,220	0,6	0
	-30	-0,219	-0,6	0
6	+30	0,219	-0,6	0
	-30	-0,218	0,6	0

Таблица 2. Влияние наклонов поверхностей в плоскости YZ на абберации системы

Допуск на преломляющие углы назначается из допустимого поперечного хроматизма системы, равного 1 секунде. Допуск на каждую призму равен 1 минуте.

В призмах возникает также такая ошибка, как *клиновидность* θ , которая вызывает отклонения луча от расчетного направления и вносит хроматизм положения. Клиновидность – это отклонение поверхностей от плоскостности. В плоскости XZ задавалось значение клиновидности θ сначала каждой призмы по отдельности, а потом и всех призм системы. Влияние изменения наклона поверхностей на абберации призмной системы представлено в табл. 3.

№ пов.	Наклон пов-ти, сек.	ΔW , дл.волн	$\Delta \sigma' x$, сек
1	10	-0,146	0,4
	20	-0,292	0,8
2	30	0,174	-0,5
	40	0,233	-0,6
3	20	0,145	-0,4
	30	0,217	-0,6
4	10	-0,177	0,5
	20	-0,353	1,0
5	30	0,175	-0,5
6	20	0,144	-0,4
	30	0,216	-0,6

Таблица 3. Влияние наклона поверхностей в плоскости XZ на абберации системы

В рассчитанной призмной системе величина хроматизма положения не должна превышать 1 секунды. При сложении углов наклона двух поверхностей и соответствующих им аббераций получаем значения клиновидности каждой призмы системы и ее влияние на качество системы. При клиновидности первой и третьей призм, равной 1 минуте, изменение поперечная абберация всей системы $\Delta \sigma' x$ равна 0,7 и 1,0 секундам, соответственно, а при $\theta = 50''$ второй и четвертой призм $\Delta \sigma' x$ равны 0,9'' и -0,9'', соответственно. В итоге в системе с учетом клиновидности всех призм получаем величину

хроматизма положения $0,9''$. Поэтому назначаем допуск на клиновидность первой и третьей призмы – 1 минута, а второй и четвертой – 50 секунд.

Угол наклона блока	ΔW , дл.волн	$\Delta\sigma'y$, сек.	ΔA
Первый блок			
30''	-0,187	0,5	-0,0002
10''	-0,113	0,3	-0,0001
5''	-0,094	0,3	0
-10''	-0,038	0,1	0
-30''	0,036	-0,1	0
-1'	0,149	-0,4	0
Второй блок			
30''	0,150	-0,4	0
20''	0,075	-0,2	0
10''	0,000	0,0	0
-10''	-0,150	0,4	0
-20''	-0,225	0,6	0
-30''	-0,300	0,8	0

Таблица 4. Наклоны призмных блоков в плоскости YZ

Угол наклона блока	ΔW , дл.волн	Поперечные aberrации		ΔA
		$\Delta\sigma'x$, сек	$\Delta\sigma'y$, сек	
Первый блок				
30''	-0,205	0,6	0,2	0
10''	-0,075	0,2	0,2	0
5''	-0,075	0,1	0,2	0
1''	-0,075	0,0	0,2	0
-1''	-0,075	0,0	0,2	0
-5''	-0,075	-0,1	0,2	0
-10''	-0,075	-0,2	0,2	0
-30''	0,204	-0,6	0,2	0
Второй блок				
30''	-0,254	0,7	0,0	0
10''	-0,085	0,2	0,0	0
5''	-0,042	0,1	0,0	0
1''	-0,009	0,0	0,0	0
-1''	-0,009	0,0	0,0	0
-5''	0,042	-0,1	0,0	0
-10''	0,085	-0,2	0,0	0
-30''	0,254	-0,7	0,0	0

Таблица 5. Наклоны призмных блоков в плоскости XZ

Так как наклоны призмных блоков влияют на увеличение и хроматизм призмной системы, то определим допуски на их наклоны в двух взаимно перпендикулярных плоскостях. В табл. 4 и 5 приведены изменения aberrаций и коэффициента анаморфозы для нулевого пучка в зависимости от наклона призмных блоков в двух плоскостях, соответственно. Наклоны призмных блоков в плоскости YZ вызывают изменение увеличения и aberrаций в системе. Поэтому допуск на наклон первого блока назначаем от +10 до -30 секунд, а на наклон второго блока – от +30 до -10 секунд. При

наклоне блоков в плоскости XZ появляется хроматизм положения, а увеличение системы остается постоянным. Допуск на наклоны призмённых блоков в этой плоскости составляет ± 30 секунд.

Заключение

В работе проведено исследование влияния формы рабочих поверхностей, клиновидности, наклонов рабочих поверхностей и призмённых блоков на качество изображения ахроматической афокальной анаморфотной призмённой системы лазерного осветителя с коэффициентом анаморфозы 0,3.

После оценки полученных данных было выявлено, что изменение толщин призм и воздушного промежутка не влияет на качество системы. Наиболее влияющими являются поверхности, граничащие с воздухом. На эти поверхности необходимо назначить более жесткий допуск. На поверхности, по которым склеиваются призмённые блоки, можно назначить более свободный допуск. Точность изготовления преломляющих углов каждой призмы равна 1 минуте. Допуск на клиновидность первой и третьей призмы составляет 1 минуту, а второй и четвертой – 50 секунд. Наклоны призмённых блоков в плоскости YZ вызывают изменение увеличения и аберраций в системе. Поэтому допуск на наклон первого блока назначаем от +10 до -30 секунд, а на наклон второго блока – от +30 до -10 секунд. При наклоне блоков в плоскости XZ появляется хроматизм положения, а увеличение системы остается постоянным. Допуск на наклоны призмённых блоков в этой плоскости составляет ± 30 секунд.

Литература

1. Бегунов Б.Н. Трансформирование оптических изображений. М.: Искусство, 1965. 232 с.
2. Пруненко Ю.К. Проектирование анаморфотной системы осветителя // Прикладная оптика-2006. СПб: Оптическое общество им. Д.С. Рождественского. 2006. Т. 3. С. 155–157.
3. Чуриловский В.Н., Халилулин К.А. Теория и расчет призмённых систем. Л.: Машиностроение, 1979. 272 с.

СВОЙСТВА ОПТИЧЕСКИХ СИСТЕМ, СОДЕРЖАЩИХ ДЕЦЕНТРИРОВАННЫЕ ПЛАНОВИДНЫЕ ЭЛЕМЕНТЫ В ОБЛАСТИ АБЕРРАЦИЙ ТРЕТЬЕГО ПОРЯДКА

С.А. Чупраков

(Институт солнечно-земной физики СО РАН)

Научный руководитель – д.т.н., профессор М.Н. Сокольский

Конечная децентрировка центрально-симметричных planoидных элементов возникает в оптических системах, являющихся чисто зеркальными аналогами систем, содержащих преломляющие корректоры Шмидта. Нулевая и близкая к нулевой оптическая сила planoидных поверхностей позволяет применить при расчете таких систем теорию aberrаций третьего порядка. Показано, что соответствующая модификация зрачковых координат при основном коррекционном параметре planoидной поверхности в выражениях для меридиональной и сагиттальной составляющих позволяет количественно и качественно оценить геометрические aberrации. Приведены примеры расчетов систем с зеркальным planoидным элементом во входном зрачке и вблизи фокальной плоскости. Рассмотренный метод позволяет рассчитать величину основного коррекционного параметра planoидной поверхности для минимизации специфических aberrаций в центре поля зрения в системах с децентрированным planoидным зеркалом во входном зрачке, а также практически полностью исправить aberrации децентрировки в центре поля зрения с помощью второго planoидного зеркала.

Существующие внеосевые оптические системы можно разделить на несколько разновидностей по степени децентрировки, типам применяемых оптических поверхностей и геометрической компоновке. Эти ключевые свойства обуславливают характер aberrаций и методику расчета децентрированных систем. Оптические системы с конечными децентрировками элементов отличаются от систем с малыми децентрировками и систем с децентрировкой входного зрачка значительной величиной и особым характером aberrаций. Расчет таких систем производится с помощью коэффициентов aberrаций четных порядков – второго, а в некоторых случаях и четвертого.

Особым случаем среди систем с большой децентрировкой являются системы, в которых децентрированы оптические поверхности малой оптической силы. Примером являются коррекционные поверхности Шмидта – преломляющие и отражающие, часто называемые planoидными поверхностями или planoидами. Случай конечной децентрировки (как правило – разворота planoидной поверхности вокруг вершины) возникает для зеркальных planoидов, установленных в произвольном месте оптической системы.

Учитывая малую оптическую силу planoидной поверхности, наиболее удобным методом таких исследований является использование коэффициентов aberrаций третьего порядка. Следуя общепринятой терминологии [2], коэффициенты третьего порядка поверхности могут быть выражены через параметры P , W , π и b . Planoидная поверхность отличается тем, что может иметь произвольные значения параметра P_p – «коррекционной силы».

$$P_p \neq 0, W_p = 0, \pi_p = 0, b_p = 0. \quad (1)$$

Существует также подход к заданию planoидной поверхности с помощью конечного радиуса кривизны при вершине r_{0p} и коэффициента асферичности b_p . Значение радиуса кривизны при вершине planoидной поверхности зависит от положения «нейтральной зоны» [2] planoидной поверхности. Однако такое представление менее удобно, так как сильная зависимость коррекционного параметра P_p одновременно и от оптической силы, и от «асферичности» planoидной поверхности значительно усложняет расчеты. Рассмотренная особенность выражений для коэффициентов третьего по-

рядка систем с зеркальными planoидными элементами позволяет описывать абберации, не прибегая к введению новых коэффициентов, которые бы зависели от параметров ориентации зеркальной planoидной поверхности относительно оси системы – положения центра поверхности и угла поворота нормали к вершине поверхности.

Как известно, выражения для геометрических аббераций третьего порядка оптической системы при $\alpha_l=0$, $\alpha_k=1$, $J=-1$ (предмет находится в «бесконечности», входной зрачок совпадает с первой поверхностью), согласно [1], можно представить в следующем виде:

$$\begin{aligned}
-2\delta g' &= \frac{m_1(m_1^2 + M_1^2)}{f^2} S + \frac{3m_1^2 tg \omega + M_1^2 tg \omega + 2m_1 M_1 tg \Omega}{f} C + \\
&\quad + (3m_1 tg^2 \omega + m_1 tg^2 \Omega + 2M_1 tg \omega tg \Omega) A + \\
&\quad + (m_1 tg^2 \omega + m_1 tg^2 \Omega) P + ftg \omega (tg^2 \omega + tg^2 \Omega) D \\
-2\delta G' &= \frac{M_1(m_1^2 + M_1^2)}{f^2} S + \frac{3M_1^2 tg \Omega + m_1^2 tg \Omega + 2m_1 M_1 tg \omega}{f} C + \\
&\quad + (3M_1 tg^2 \Omega + M_1 tg^2 \omega + 2m_1 tg \omega tg \Omega) A + \\
&\quad + (M_1 tg^2 \omega + M_1 tg^2 \Omega) P + ftg \Omega (tg^2 \omega + tg^2 \Omega) D
\end{aligned} \tag{2}$$

В выражениях (2) ω и Ω – углы предметного луча соответственно в меридиональной и сагиттальной плоскостях. Воспользовавшись выражениями (4) и представляя коэффициенты аббераций третьего порядка как суммы коэффициентов элементов системы, не содержащих planoидных поверхностей, и соответствующие коэффициенты planoидных поверхностей, можно переписать выражения для зрачковых координат, учитывая параметры децентрировки. В интересующем нас случае *поворота* planoидной поверхности на угол θ в меридиональной плоскости и угол Θ в сагиттальной плоскости

$$\begin{aligned}
\bar{m}_1 &= m_1 \cos \theta + M_1 \cos \Theta \sin \theta \\
\bar{M}_1 &= M_1 \cos \Theta
\end{aligned} \tag{3}$$

В случае смещения центра planoидной поверхности в меридиональном и сагиттальном направлении соответственно на величины Δx и Δy

$$\begin{aligned}
\bar{m}_1 &= m_1 + \Delta x \\
\bar{M}_1 &= M_1 + \Delta y
\end{aligned} \tag{4}$$

В случае поворота одной planoидной поверхности в меридиональной плоскости выражения (2) примут следующий вид:

$$\begin{aligned}
-2\delta g' &= \frac{m_1(m_1^2 + M_1^2)}{f^2} S_0 + k \frac{\bar{m}_1(\bar{m}_1^2 + M_1^2)}{f^2} S_p + \\
&\quad + \frac{3m_1^2 tg \omega + M_1^2 tg \omega + 2m_1 M_1 tg \Omega}{f} C_0 + k \frac{3\bar{m}_1^2 tg \omega + M_1^2 tg \omega + 2\bar{m}_1 M_1 tg \Omega}{f} C_p + \\
&\quad + (3m_1 tg^2 \omega + m_1 tg^2 \Omega + 2M_1 tg \omega tg \Omega) A_0 + k(3\bar{m}_1 tg^2 \omega + m_1 tg^2 \Omega + 2M_1 tg \omega tg \Omega) A_p + \\
&\quad + (m_1 tg^2 \omega + m_1 tg^2 \Omega) P + ftg \omega (tg^2 \omega + tg^2 \Omega) (D_0 + kD_p) \\
-2\delta G' &= \frac{M_1(m_1^2 + M_1^2)}{f^2} S_0 + k \frac{M_1(\bar{m}_1^2 + M_1^2)}{f^2} S_p + \\
&\quad + \frac{3M_1^2 tg \omega + m_1^2 tg \omega + 2m_1 M_1 tg \omega}{f} C_0 + k \frac{3M_1^2 tg \omega + \bar{m}_1^2 tg \omega + 2\bar{m}_1 M_1 tg \omega}{f} C_p + \\
&\quad + (3M_1 tg^2 \omega + M_1 tg^2 \omega + 2m_1 tg \omega tg \Omega) A_0 + k(3M_1 tg^2 \omega + M_1 tg^2 \omega + 2\bar{m}_1 tg \omega tg \Omega) A_p + \\
&\quad + (M_1 tg^2 \omega + M_1 tg^2 \Omega) P + ftg \Omega (tg^2 \omega + tg^2 \Omega) (D_0 + kD_p)
\end{aligned} \tag{5}$$

Значения S_p , C_p , A_p и D_p вычисляются для каждой планоидной поверхности. Например, в системе «зеркальный Райт» [3] с главным зеркалом с квадратом эксцентриситета e^2 :

$$S_p = -\frac{(1+e^2)}{4}. \quad (6)$$

Коэффициент k в выражении (5) выбирается для минимизации разностей:

$$\begin{aligned} \Delta m_1 &= m_1 - \tilde{m}_1 \\ \Delta M_1 &= M_1 - \tilde{M}_1. \end{aligned} \quad (7)$$

Если планоид наклонен в одной плоскости, например в меридиональной ($\theta \neq 1$, $\Theta=1$), для минимизации Δm_1 необходимо, чтобы коэффициент k удовлетворял условию $\Delta m_1 = \Delta M_1$. Учитывая (3), можно показать, что

$$k = \frac{2}{\cos \theta (1 + \cos \theta)}. \quad (8)$$

Точность представления aberrаций будет проиллюстрирована ниже для трех оптических систем, зеркальные планоиды в которых наклонены на конечные углы. К сожалению, ограниченный объем публикации не позволяет привести здесь результаты аналогичных численных экспериментов для смещенной планоидной поверхности.

Первая система представляет собой зеркальный аналог системы Райта с главным зеркалом $e^2 < -1$ и наклонным зеркальным планоидом, установленным во входном зрачке. При расчетах приняты следующие параметры системы: $D=155$ мм, $f=720$ мм, $e^2=-1.3$, расстояние между зеркалами $d=-610$ мм, поле зрения $2\omega=1^\circ$. Точечные диаграммы системы показаны на рис. 2. Здесь, как и на последующих аналогичных рисунках в двух (на рис. 6 – трех) верхних рядах показаны aberrационные кривые, рассчитанные по формулам (5) с учетом (3) и (8), а в двух (на рис. 6 – трех) нижних рядах – точечные диаграммы, рассчитанные методами аналитической геометрии. На всех рис. 2, 4 и 6 масштабы согласованы (длина стороны квадрата равна длине вертикальной масштабной линии).

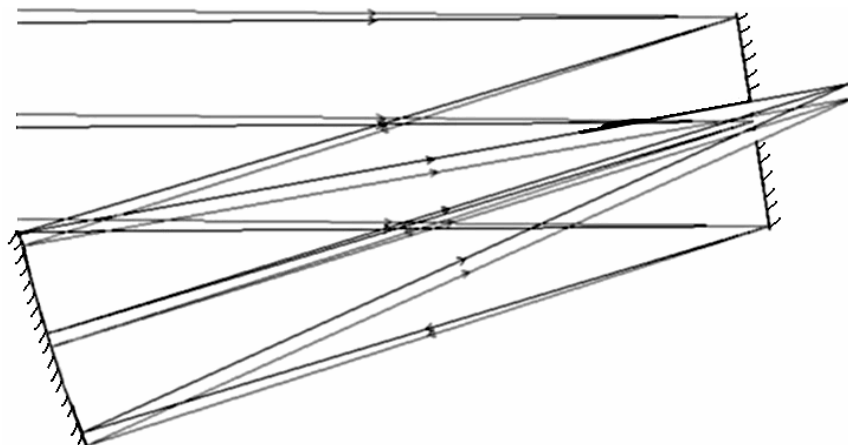


Рис. 1. Система «зеркальный Райт» с наклонным планоидом во входном зрачке

В качестве второго примера рассмотрим систему Ричи-Кретьена (см. рис. 3), в которой наклонный центральносимметричный зеркальный планоид, установленный вблизи фокальной плоскости, используется для коррекции астигматизма. Диаметр входного зрачка 155 мм, $f=1317$ мм (относительное отверстие $A=1:8.5$). Планоидное зеркало наклонено на 16° к оси двухзеркальной системы и установлено на расстоянии 55 мм ($f/23.9$) от фокальной плоскости. Точечные диаграммы системы показаны на рис. 4.

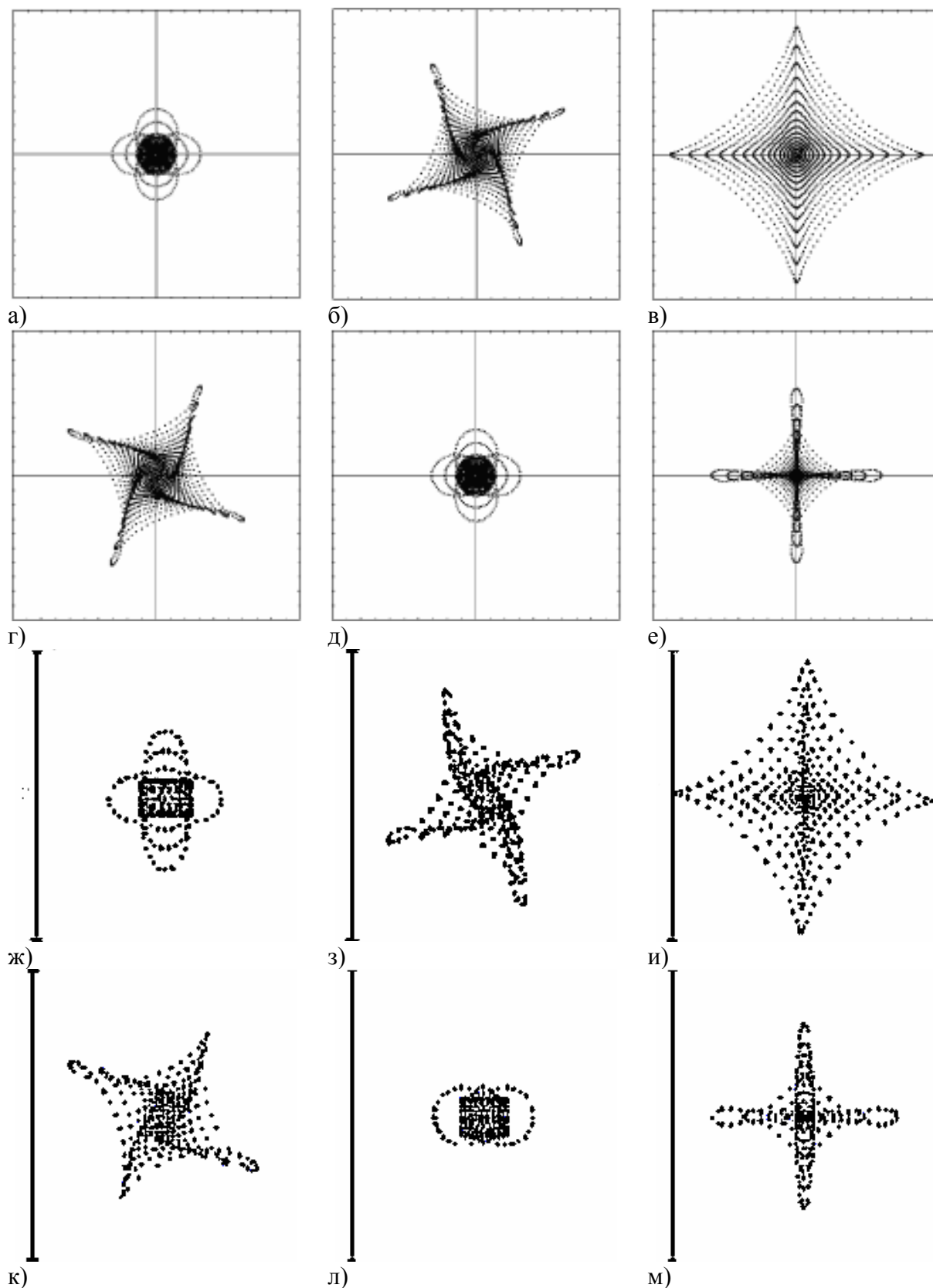


Рис. 2. Точечные диаграммы системы «зеркальный Райт», рассчитанные по формулам (5) (два верхних ряда), и методами аналитической геометрии (два нижних ряда).
 а,ж) $\omega=0.5^\circ, \Omega=0^\circ$; б,з) $\omega=0.35^\circ, \Omega=0.35^\circ$; в,и) $\omega=0^\circ, \Omega=0.5^\circ$; г,к) $\omega=-0.35^\circ, \Omega=0.35^\circ$;
 д,л) $\omega=-0.5^\circ, \Omega=0^\circ$; е,м) $\omega=0^\circ, \Omega=0^\circ$

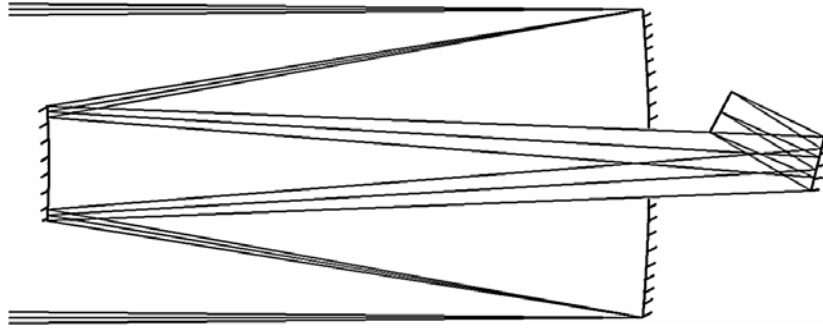


Рис. 3. Ричи-Кретьен с предфокальным корректором астигматизма

Как можно видеть из рис. 4, предлагаемый метод применим в случае произвольного расположения децентрированной planoидной поверхности в системе.

Рассмотренные свойства выражений для коэффициентов aberrаций третьего порядка planoидных поверхностей позволяют подтвердить интуитивное предположение о возможности компенсации aberrаций децентрировки на оси системы, содержащей наклонное planoидное зеркало, с помощью второго наклонного planoидного зеркала с тем же параметром P_p , плоскость наклона которого перпендикулярна. Параметр $P_{1/2}$ каждого из planoидов должен быть при этом равен половине параметра P_p эквивалентного planoида, исправляющего сферическую aberrацию последующей оптической системы:

$$P_{1/2} = \frac{P_p}{2}. \quad (9)$$

Углы наклона обоих зеркал в перпендикулярных плоскостях также одинаковы:

$$\theta = \Theta \equiv i, \quad (10)$$

где θ – угол наклона первого planoида в меридиональной плоскости, а Θ – угол наклона второго planoида в сагиттальной плоскости. Согласно (9) и (10), функции зрачковых координат при коэффициенте сферической aberrации в этом случае будут иметь вид

$$\begin{aligned} -2\delta g' &= -k \frac{m_1 \cos i (m_1^2 \cos^2 i + M_1^2)}{2f^2} S_{pl} - k \frac{m_1 (m_1^2 + M_1^2 \cos^2 i)}{2f^2} S_{pII} = \\ &= -k \frac{m_1 ((1 + \cos^3 i) m_1^2 + \cos i (1 + \cos i) M_1^2)}{2f^2} S_{pl}, \quad (11) \\ -2\delta G' &= -k \frac{M_1 (m_1^2 \cos^2 i + M_1^2)}{2f^2} S_{pl} - k \frac{M_1 \cos^2 i (m_1^2 + M_1^2 \cos^2 i)}{2f^2} S_{pII} = \\ &= -k \frac{M_1 (\cos i (1 + \cos i) m_1^2 + (1 + \cos^3 i) M_1^2)}{2f^2} S_{pl} \end{aligned}$$

где i — угол наклона каждого зеркального planoида.

Как можно заметить из выражений (11), при $S_{pl} = S_{pII}$ aberrационные кривые слабо отличаются от окружностей. Этот результат можно проиллюстрировать на примере двухзеркальной анастигматической системы Шмидта-Кассегрена с плоским полем с двумя сферическими зеркалами, внутренним расположением фокальной плоскости и двумя наклонными planoидными зеркалами, плоскости наклона которых перпендикулярны. Диаметр входного зрачка 200 мм, эквивалентный фокус 1000 мм, фокусное расстояние главного зеркала 600 мм, расстояние между главным и вторичным зеркалами 390 мм. Оба зеркала сферические, фокальная поверхность плоская, расположена между зеркалами. Угол наклона каждого из planoидов 8° , плоскости углов наклона planoидов перпендикулярны. Система и ход лучей, показаны на рис. 5, точечные диаграммы – на рис. 6.

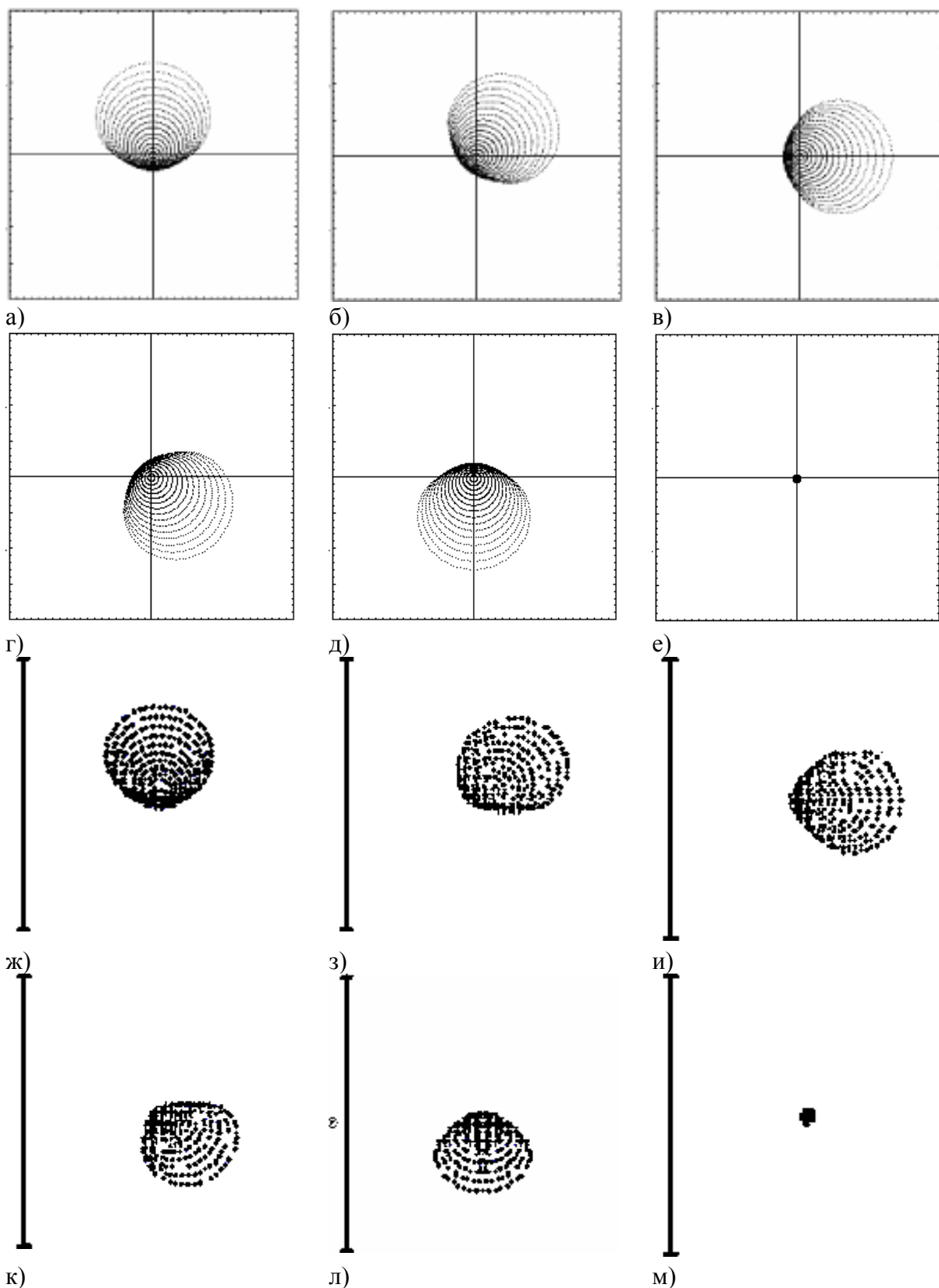


Рис. 4. Точечные диаграммы системы Ричи-Кретьена с предфокальным корректором астигматизма, рассчитанные по формулам (5) (два верхних ряда) и методами аналитической геометрии (два нижних ряда). а,ж) $\omega=0.35^\circ$, $\Omega=0^\circ$; б,з) $\omega=0.248^\circ$, $\Omega=0.248^\circ$; в,и) $\omega=0^\circ$, $\Omega=0.35^\circ$; г,к) $\omega=-0.248^\circ$, $\Omega=0.248^\circ$; д,л) $\omega=-0.35^\circ$, $\Omega=0^\circ$; е,м) $\omega=0^\circ$, $\Omega=0^\circ$

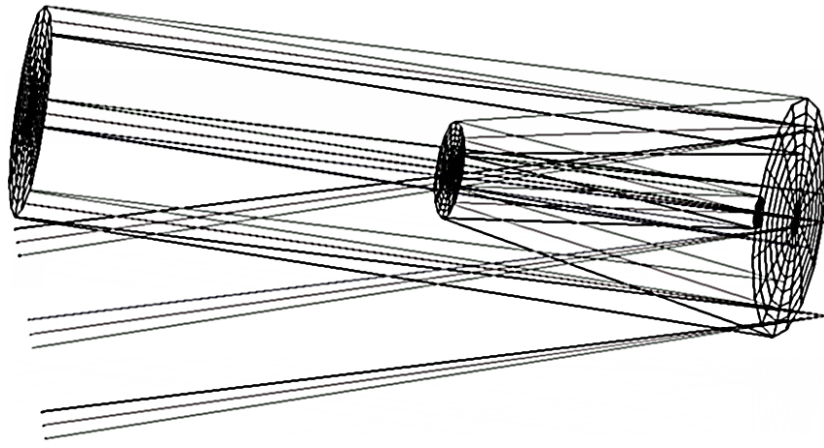
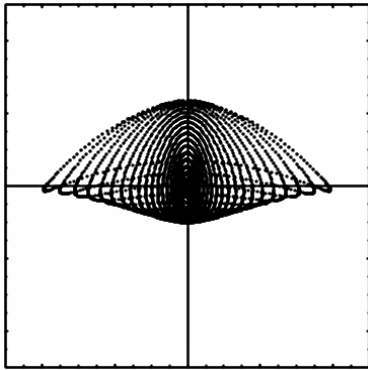
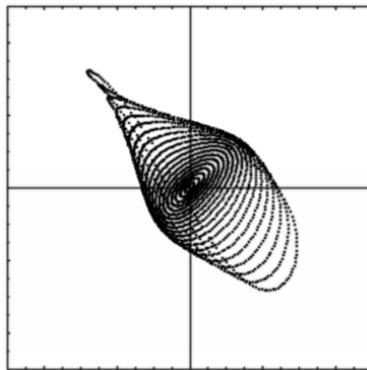


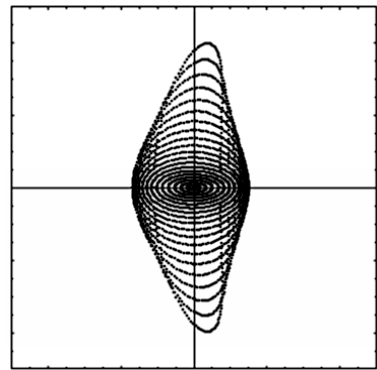
Рис. 5. Четырехзеркальная система с двумя плоскоидными зеркалами, наклоненными в перпендикулярных плоскостях



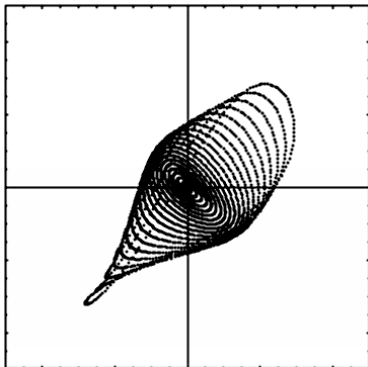
а)



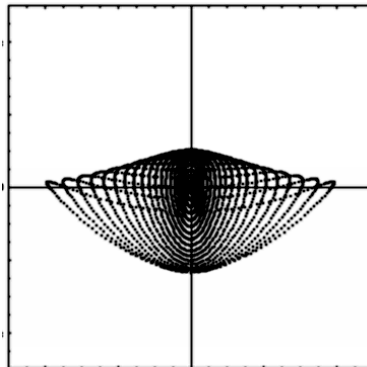
б)



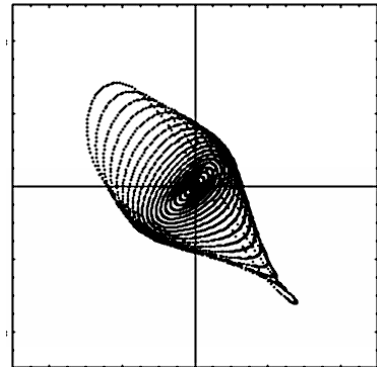
в)



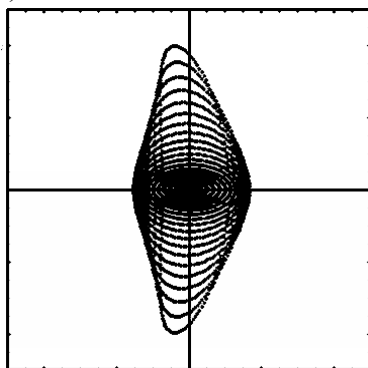
г)



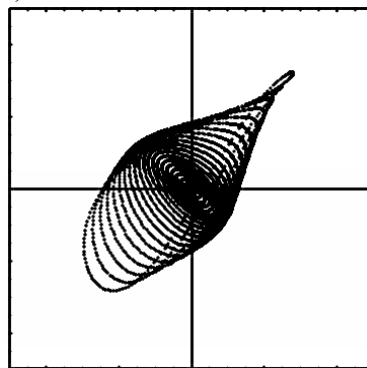
д)



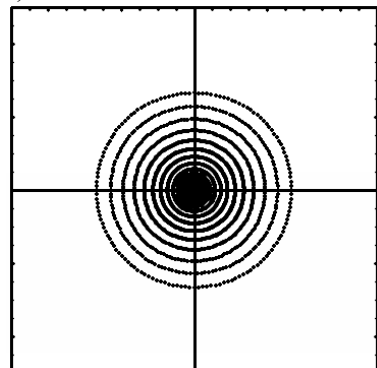
е)



ж)



з)



и)

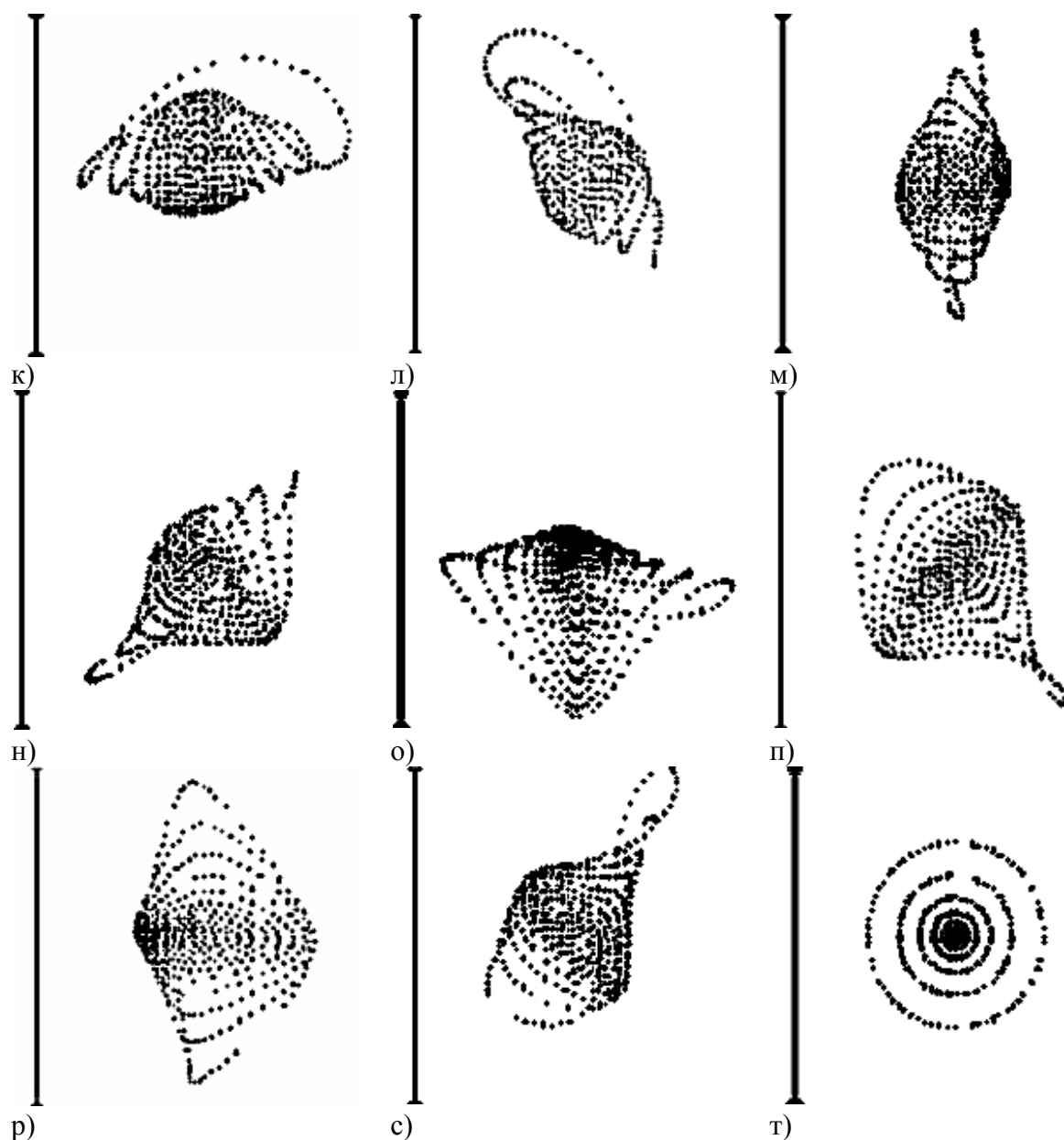


Рис. 6. Точечные диаграммы системы Шмидта-Кассегрена с двумя зеркальными planoидами, наклоненными в перпендикулярных плоскостях, показанной на рисунке 6. Три верхних ряда – расчет по формулам (5), три нижних ряда – аналитический расчет.
а,к) $\omega=1^\circ, \Omega=0$; б,л) $\omega=0.707^\circ, \Omega=0.707^\circ$; в,м) $\omega=0^\circ, \Omega=1^\circ$; г,н) $\omega=0.707^\circ, \Omega=0.707^\circ$;
д,о) $\omega=1^\circ, \Omega=0^\circ$; е,п) $\omega=0.707^\circ, \Omega=0.707^\circ$; ж,р) $\omega=0^\circ, \Omega=1^\circ$; з,с) $\omega=0.707^\circ, \Omega=0.707^\circ$;
и,т) $\omega=0^\circ, \Omega=0^\circ$

Результаты исследований показали возможность использования теории aberrаций третьего порядка для описания aberrационных свойств не только систем, в которых наклонный центральносимметричный planoид расположен во входном зрачке, но систем, в которых такой planoид расположен на значительном расстоянии от входного зрачка и даже перед фокальной плоскостью. Рассмотренная методика позволила предсказать возможность полной компенсации специфических aberrаций наклонного planoида для центра поля зрения.

Литература

1. Чуриловский В.Н. Теория хроматизма и aberrаций третьего порядка. Л.: Машиностроение, 1968.
2. Слюсарев Г.Г. Расчет оптических систем. Л.: Машиностроение, 1975.
3. Чупраков С.А., Бородин А.Н. Широкоугольный зеркальный объектив телескопа. / Патент России № 2215314, 2001.
4. Михельсон Н.Н. Оптика астрономических телескопов и методы ее расчета. М.: Наука, 1995.

ИСКАЖЕНИЯ ВОЛНОВОГО ФРОНТА ПРИ ИЗМЕНЕНИИ ПОЛОЖЕНИЯ ПЛОСКОСТИ ИНТЕРФЕРОГРАММЫ

Г.С. Полищук

Научный руководитель – д.т.н., профессор М.Н. Сокольский

В связи с развитием техники требуются все более точные объективы. Требования к точности и производительности контрольных приборов приводят к тому, что необходимо разрабатывать новые оптические схемы контроля. В связи с этим на новом витке развития оптического приборостроения необходимо проведение анализа применимости условий моделирования оптических процессов и разработки оптических схем контроля. Целью работы являлась оценка ошибки, возникающей в результате изменения положения регистрирующего устройства в схеме интерференционного контроля микрообъективов. Показано, что вид полос в интерферометрах для контроля высокоточных микрообъективов практически не зависит от положения плоскости наблюдения интерференционной картины при малых аберрациях объектива. Получена функциональная зависимость искажения вида полос от величины волновой аберрации. Получен критерий применимости данного правила. Выполненные расчеты подтверждены экспериментальными данными.

При проектировании технологического оборудования с каждым годом ужесточаются требования к точности системы и, как следствие, к точности компонентов. Для оптической части оборудования это означает ужесточение требований по точности и работа с коротковолновым излучением. Так, в микроэлектронике практически стало правилом использование излучения дальней ультрафиолетовой области спектра. На сегодняшний день большинство создаваемых установок работает на длинах волн порядка 250 нм. Это обстоятельство приводит к тому, что требования к точности оптических элементов, при прочих равных условиях, выше требований к точности в оптике для видимой области (633 нм – рабочая длина волны большинства интерферометров) более чем в 2,5 раза.

Для исследования формы поверхностей и неоднородностей в прозрачных средах широко применяются методы классической интерферометрии [1-4]. Они отличаются высокой точностью и обеспечивают получение наглядных результатов контроля. Однако высокая точность метода в реальных приборах может быть использована только при выполнении определенных требований, предъявляемых к качеству и взаимному расположению оптических элементов интерферометра. Несоблюдение этих требований приводит к приборной и методической погрешности, величина которой может значительно снизить потенциальную точность интерференционного метода.

В литературе достаточно полно освещена связь приборной или методической погрешностей с допусками на изготовление и юстировку оптических элементов интерферометров [5, 6], а также аберраций их осветительных систем при наличии разности хода, отличающейся от нуля.

В настоящей работе рассматривается случай возникновения погрешности при неточном совпадении плоскости наблюдения с плоскостью, в которой под воздействием контролируемого объекта произошла деформация оптической волновой поверхности рабочего пучка.

Существуют приборы, в которых одновременное фокусирование на объект и деформированную им волну невозможно или затруднено. К ним, в частности, относятся автоколлимационные системы для исследования неоднородностей [7, 8], в которых рабочая волна дважды проходит через исследуемый объект.

В зависимости от характера погрешности, возникающей вследствие расфокусировки, все применяемые в практике интерференционных исследований схемы можно разделить на несколько типов. Одной из них является схема с отдельными ветвями [9] и узким опорным пучком, в котором структура пучка сравнения (опорного), как правило, близка к идеальной и в процессе исследования не подвергается изменениям или возникающие изменения пренебрежимо малы. В таких системах при изучении связи

между величинами погрешности и расфокусировки достаточно выполнить анализ формы только рабочей волны.

Волновая aberrация $W(x, y)$ оптической системы для данного случая определяется как разность оптических длин хода луча $\langle L \rangle$, проходящего через координаты x, y на выходном зрачке и луча, идущего вдоль оптической оси $\langle Lo \rangle$.

$$\Delta W(x, y) = \langle L \rangle - \langle Lo \rangle.$$

Поэтому для измерения W с помощью интерферометра необходимо, чтобы интерференционная картина была локализована в выходном зрачке контролируемой системы. Пример такой системы показан на рис. 1.

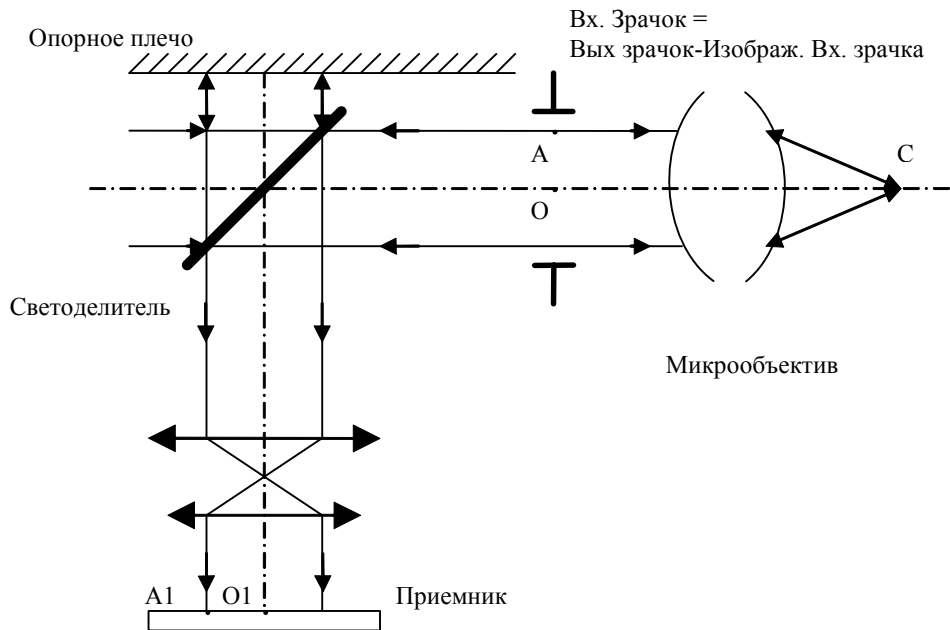


Рис. 1. Пример схемы интерференционного контроля

Волновая aberrация для этой системы равна

$$W = 2 \langle AC \rangle - 2 \langle OC \rangle.$$

Предположим, что изображение входного зрачка не совпадает с его расчетным положением, как показано на рис. 2.

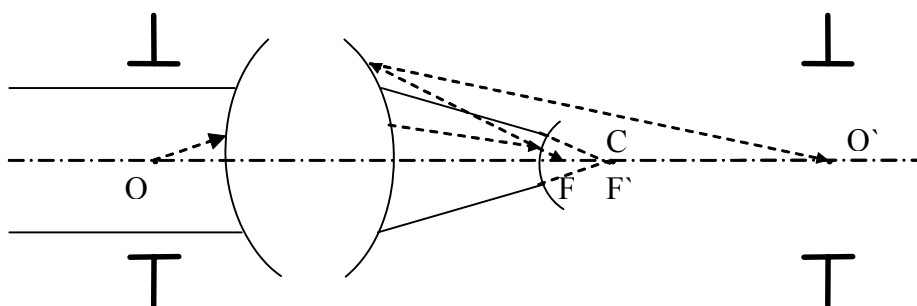


Рис. 2. Оптическая схема, в которой изображение входного зрачка не совпадает с входным зрачком

Следовательно, если для автоколлимации использовать сферический отражатель, то изображение входного зрачка не совпадает с входным зрачком. В предельном случае в системе с телецентрическим ходом луча мы получим систему, изображенную на рис. 3.

$$\text{В этом случае } xx' = f_{об}^2; \quad x = \frac{f_{об}^2}{f_{и}}.$$

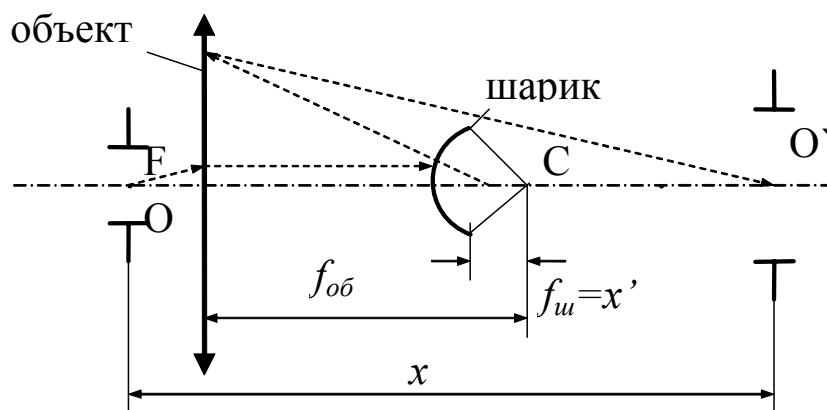


Рис. 3. Оптическая схема с автоколлимацией от шарика

Если интерферометр настроен на точку O , т.е. эта плоскость совмещена с плоскостью приёмника, то возникнет погрешность измерения (см. рис. 4.).

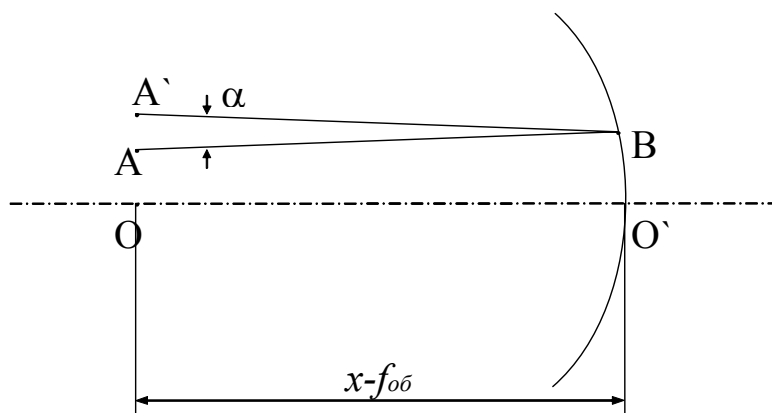


Рис. 4. К расчету погрешности измерения волнового фронта

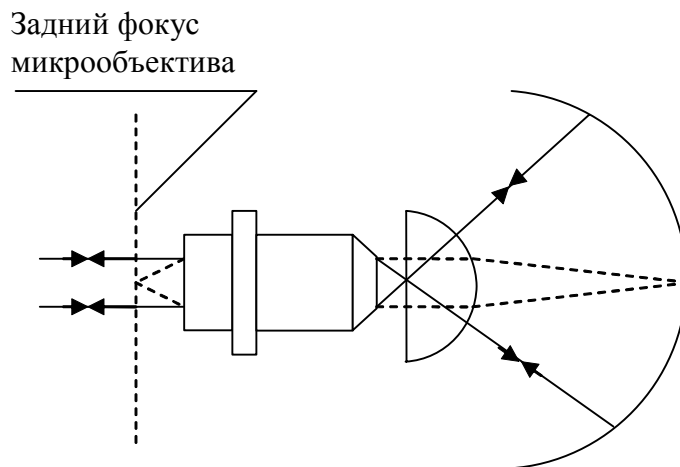


Рис. 5. Устройство Дайсона для микрообъективов

При этом угол между начальным и отраженным лучом равен $\alpha = \frac{\partial W}{\partial \rho}$, где W – волновая aberrация, ρ – относительный размер выходной апертуры.

$$\Delta W = AB - A'B = (x - f_{об}) \frac{\alpha^2}{2} = \left(\frac{f_{об}^2}{f_{у}} - f_{об} \right) \frac{\alpha^2}{2}.$$

Если $R_{шарика} = f_{об}$, то $\Delta W = 0$.

Для сравнения мы воспользуемся другим способом – устройством Дайсона [9]. Принципиальная схема данного устройства приведена на рис. 5.

Оценим погрешности измерения волновой аберрации на примере оптической системы, показанной на рис. 1, для нескольких вариантов значений сферической аберрации коэффициентов C20, C40, C60.

В табл. 1 и 2 приведены результаты моделирования схем контроля с автоколлимационным шариком и устройством Дайсона.

	Варианты исходной системы	Контроль с устройством Дайсона	Контроль с автоколлимационным шариком
1	C20 = 0,04 λ C40 = 1,8λ C60 = -0,06λ	C20=0,07 C40=3,6λ C60= - 0,1175λ	C20 = -0,58λ C40 = 2,615λ C60 = -0,802λ
2	C20 = 0,07λ C40 = 0,6λ C60 = -0,055λ	C20 = 0,13 C40 = 1,21 C60 = -0,11λ	C20 = 0,04λ C40 = 1,13λ C60 = -0,19λ
3	C20 = 0,04λ C40 = 0 C60 = -0,05λ	C20 = 0,07λ C40 = 0 C60 = -0,11λ	C20 = 0,07λ C40 = 0 C60 = -0,12λ

Таблица 1. Полиномы Цернике для оптических систем с разными значениями волновых аберраций

	Контроль насадкой Дайсона		Контроль с автоколлим. Шариком	
	ρ	W (ρ)/λ	ρ	W (ρ)/λ
1	1	-0,03	1	-1,54
	0,866	-4,28	0,866	-3,34
	0,707	-5,60	0,707	-4,71
	0,5	-4,46	0,5	-4,03
	ρ	W (ρ)/λ	ρ	W (ρ)/λ
2	1	0,10	1	0,34
	0,866	1,47	0,866	1,50
	0,707	1,90	0,707	1,93
	0,5	1,70	0,5	1,71
	ρ	W (ρ)/λ	ρ	W (ρ)/λ

Таблица 2. Искажение волнового фронта для различных оптических систем

Из табл. 1 и 2 можно сделать следующие выводы.

- Применение устройства Дайсона позволяет контролировать системы с большими аберрациями (из вар. 1 видно, что в автоколлимации аберрации удваиваются).
- Применение автоколлимационной схемы с шариком при больших аберрациях дает большую погрешность (вар. 1) и с уменьшением аберраций погрешность уменьшается.
- В вар. 2 аберрации незначительны, а в вар. 3 практически отсутствуют. В вар. 2

RMS исходной системы равна $RMS_{исх} = \sqrt{\frac{C_{40}^2}{5} + \frac{C_{60}^2}{7}} \approx \frac{C_{40}}{\sqrt{5}} = \frac{0,6}{2,236} \lambda = 0,268\lambda$, а

$RMS_{системы} = 0,5 \sqrt{\frac{1,13^2}{5} + \frac{0,19^2}{7}} = 0,255\lambda$, т.е. с точки зрения RMS погрешность составляет $\approx 0,01\lambda$.

- Можно считать, что если RMS контролируемой системы менее 0,15–0,2λ, то погрешность в автоколлимационной схеме с шариком достаточно мала.

Литература

1. Духовел И.И., Федина Л.Г. // ОМП. 1973. № 8. С.50.
2. Духовел И.И., Константиновская Н.В., Федина Л.Г. // ОМП. 1975. № 7. С.64.
3. Белозеров А.Ф., Кузилин Ю.Е. // ОМП. 1971. № 12. С.39.
4. Овечкин А.И. и др. // ОМП. 1975. № 12. С.34.
5. Коломийцова Т.С., Федина Л.Г. // ОМП. 1976. №2. С.32.
6. Овечкин А.П. и др. // ОМП. 1976. №4. С.68.
7. Орлов А.А. и др. // ОМП. 1974. №2. С.24.
8. Овечкин А.П. и др. // ОМП. 1976. №4. С.68.
9. Малакара Д. Оптический производственный контроль. 1985, 46 с.

ВАРИАНТЫ КОМПОЗИЦИИ ЗЕРКАЛЬНО-ЛИНЗОВОГО ОБЪЕКТИВА ПЛАННАСТИГМАТА

Е.В. Ермолаева

Научный руководитель – д.т.н., профессор В.А. Зверев

При наблюдении через телескопическую систему важно обеспечить высокое качество изображения по всему полю зрения. Поэтому актуальной проблемой является повышение качества изображения внеосевых точек предмета. Качество изображения внеосевых точек предмета определяется прежде всего величинами комы и астигматизма.

В данной работе рассмотрены основные направления повышения качества изображения внеосевых точек.

Рассмотрим компенсацию комы и астигматизма в классических двухзеркальных системах.

Классической двухзеркальной системой телескопа называется система, в которой главное зеркало имеет форму параболоида вращения. Классические двухзеркальные системы, предложенные Ньютоном, Грегори и Кассегреном, свободны от сферической aberrации, но обладают значительной комой.

Для компенсации комы в изображении точки, образованном классической двухзеркальной системой, профессор В.Н. Чуриловский предложил применить мениск, ограниченный двумя сферическими поверхностями, имеющими равные радиусы, при этом осевая точка предмета должна быть расположена в переднем фокусе первой поверхности, а ее изображение, образованное мениском, оказывается расположенным в заднем фокусе второй поверхности мениска. В некоторых случаях толщина мениска может оказаться слишком большой. В этих случаях целесообразно заменить мениск двумя линзами, как бы вводя в него ограниченный плоскими поверхностями воздушный промежуток. Вариант схемы применения компенсатора представлен на рис. 1. Положив расстояние между плоскими поверхностями (воздушный промежуток) равным нулю, естественным образом можем от двух линз перейти к мениску.

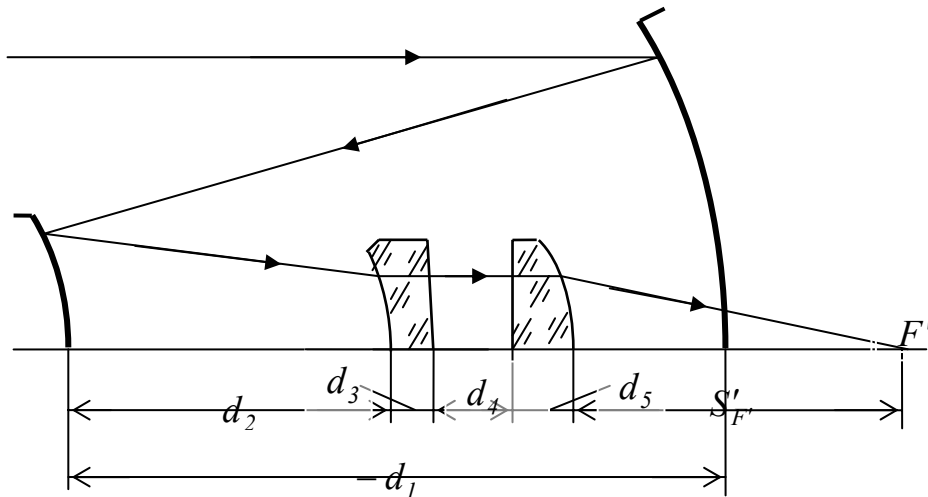


Рис. 1. Компенсатор комы в классической двухзеркальной системе

Коэффициент комы классической двухзеркальной системы равен $K_0 = -\frac{1}{2}$, т.е. не зависит от вида системы. Поэтому компенсатор Чуриловского является универсальным в применении в классических двухзеркальных системах.

При условии компенсации комы, т.е. при $K_0 = 0$, толщина мениска определяется выражением:

$$d_m = k \frac{(n-1)^2}{2n},$$

где n – показатель преломления материала мениска; k – величина, равная заднему фокальному отрезку, выраженному в масштабе фокусного расстояния, т.е. определяющая, как далеко расположен мениск от задней фокальной плоскости.

Данный мениск обладает апохроматической коррекцией хроматизма положения. Однако он не свободен от хроматизма увеличения.

Хроматическая aberrация увеличения в данном случае определяется формулой:

$$\frac{\delta L'_k}{l'_k} = - \frac{d_3 + nd_4 + d_5}{kn\mu}.$$

Отсюда следует, что при апохроматической коррекции хроматизма положения aberrация увеличения, вносимая в изображение компенсатором комы, тем меньше, чем меньше толщина компенсатора, чем дальше он расположен от задней фокальной плоскости и чем больше показатель преломления и коэффициент дисперсии материала мениска или линз компенсатора.

Итак, мениск с равными радиусами кривизны поверхностей решает задачу коррекции комы в изображении, образованном системой анаберрационных поверхностей. Однако применение рассмотренного мениска позволяет компенсировать не только кому изображения, но кому и астигматизм, если сферические поверхности мениска заменить одинаковыми несферическими поверхностями вращения второго порядка.

До настоящего времени соотношения, определяющие параметры мениска, для данного случая не были получены. Однако они очень важны, так как применение данного мениска позволяет получить лучшее качество изображения в классических двухзеркальных системах. Нами были получены соотношения, определяющие параметры компенсатора комы и астигматизма изображения точки, образованного классической двухзеркальной системой.

При условии компенсации комы и астигматизма толщина компенсатора и коэффициент деформации поверхностей определяются следующими выражениями:

$$d_k = 2kn \frac{n-1}{3n+1}$$

$$\sigma = - \left(\frac{n+1}{2} \right)^2.$$

Так как показатель преломления материала мениска $n > 1$, величина коэффициента деформации $\sigma < -1$, т.е. несферические поверхности линз компенсатора имеют форму гиперболоида вращения. При этом при изготовлении одна из линз может служить пробным стеклом для изготовления другой.

Как видно из приведенных формул, параметры корректора комы и астигматизма не зависят от вида классической двухзеркальной системы. Одним и тем же мениском можно устранить кому и астигматизм любой классической двухзеркальной системы.

Кривизна поверхности изображения, образованного всей системой в целом, полностью определяется параметрами двухзеркальной системы. Параметры коррекционного мениска не влияют на кривизну поверхности изображения.

Если отказаться от строгого стигматизма в изображении осевой точки, то путем изменения коэффициентов деформации отражающих поверхностей можно получить апланатическую коррекцию aberrаций третьего порядка в изображении, образованном двухзеркальной системой, которую при подобном исправлении aberrаций принято называть зеркальной системой Ричи-Кретьена.

В системе Ричи-Кретьена при хорошей коррекции сферической aberrации и комы весьма велики остаточные астигматизм и кривизна поверхности изображения. Для

компенсации кривизна поверхности изображения в системе Ричи-Кретьена применяется конфокальная линза.

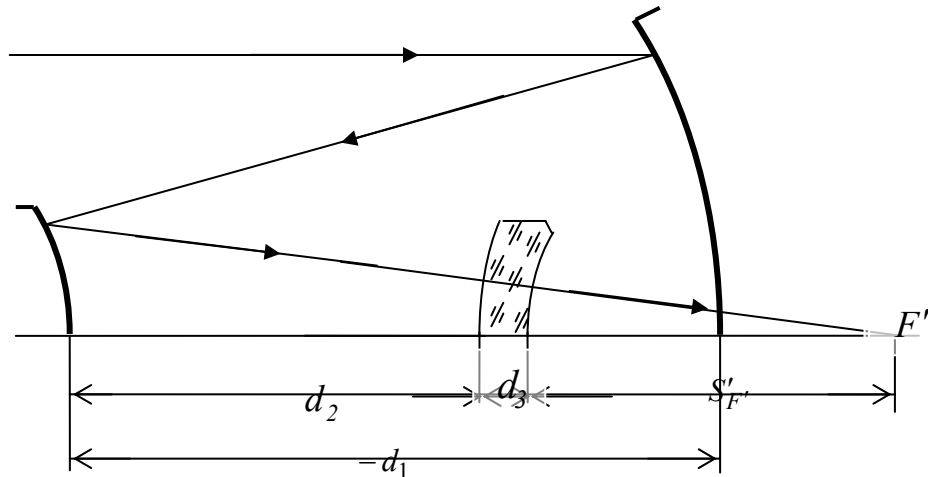


Рис. 2. Компенсатор кривизны поверхности изображения в системе Ричи-Кретьена

При исправленной педвалеовой кривизне поверхности изображения величина астigmatизма по абсолютной величине примерно вдвое превышает астigmatизм изображения, образованного системой Ричи-Кретьена без корректора кривизны поверхности изображения. При исправленной кривизне величину остаточного астigmatизма определяют только параметры самой зеркальной системы. Таким образом, в рассматриваемой системе нет коррекционных параметров для исправления астigmatизма.

При условии компенсации кривизны поверхности изображения оптическая сила мениска равна

$$\varphi_M = \frac{\eta - (\eta - 1)^2 k_s}{\eta^2},$$

где $k_s = -\frac{S'_{F'}}{d_1}$. При этом величина радиуса одной из поверхностей выбирается произ-

вольно в пределах разумного диапазона, а величина радиуса другой поверхности следует из соотношения

$$\varphi_M = \frac{n-1}{n} \left(\frac{1}{r_3} - \frac{1}{r_4} \right).$$

Для компенсации остаточного астigmatизма в качестве коррекционного параметра можно использовать форму одной из поверхностей конфокального мениска. В этом случае коэффициент деформации асферизуемой поверхности определяется выражением

$$\sigma_3 = -\frac{1}{2} \frac{h_3}{n-1} \frac{2(1-\eta)^2 k_s^2 - \eta^2}{[h_3 \eta + k_s (\eta - h_3)] (\eta - h_3)}.$$

Например, при $k_s = 1$, $\eta = 0.3$, $h_3 = 0.075$, $\sigma_3 = -1.1586$.

В этом случае требуется пересчет коэффициентов деформации поверхностей зеркал. В случае асферизации одной из поверхностей компенсатора они будут определяться выражениями:

$$\sigma_1 = \frac{2\eta^2}{k_s (\eta - 1)^3} \frac{\eta(\eta - h_3) + 2h_3 k_s (1 - \eta)^2}{h_3 \eta + k_s (\eta - h_3)} - 1$$

$$\sigma_2 = -\left[\frac{\eta + (1 - \eta)k_s}{\eta - (1 - \eta)k_s} \right]^2 + \frac{2\eta}{[\eta + (\eta - 1)k_s]^3} \left[k_s \eta + h_3 \frac{2(1 - \eta)^2 k_s^2 - \eta^2}{\eta - h_3} \right].$$

Компенсатор для системы Ричи-Кретьена был рассчитан Г.И. Цукановой.

Далее приведены aberrации системы Кассегрена без компенсатора и aberrации системы Кассегрена с компенсатором комы и астигматизма, aberrации системы Ричи-Кретьена без компенсатора и aberrации системы Ричи-Кретьена с компенсатором кривизны поверхности изображения и астигматизма при одних и тех же параметрах системы.

Аберрации системы Кассегрена без компенсатора. Параметры системы: $f' = 1000$ мм, $2\omega = 4^\circ$, $\eta = 0.3$.

Отн. зр. коор.	Продольные aberrации $\Delta S'$, мм	Поперечные aberrации $\Delta y'$, мм	Волновые aberrации W , дл. волн	Неизопланатизм η , (%)
1	0.000963	0.000121	0.00688	0.391
0.866	0.000724	0.0000786	0.00388	0.293
0.707	0.000484	0.0000429	0.00173	0.195
0.5	0.000243	0.0000152	0.435-3	0.0977

Таблица 1. Аберрации осевой точки системы Кассегрена без компенсатора

Размер предм., $\sin \sigma$	Кома	Астигматизм	Кривизна поверхности изображения
0.0248	0.295	-1.60	-0.001
0.0350	0.425	-3.20	-0.01

Таблица 2. Аберрации внеосевой точки системы Кассегрена без компенсатора

Аберрации системы Кассегрена с компенсатором комы и астигматизма. Параметры системы: $f' = 1000$ мм, $2\omega = 4^\circ$, $\eta = 0.3$.

Отн. зр. коор.	Продольные aberrации $\Delta S'$, мм	Поперечные aberrации $\Delta y'$, мм	Волновые aberrации W , дл. волн	Неизопланатизм η , (%)
1	-0.0462	-0.0058	-0.215	0.0103
0.866	-0.0254	-0.0028	-0.0875	0.00563
0.707	-0.0108	-0.00096	-0.0242	0.00241
0.5	-0.0023	-0.00015	-0.0024	0.000548

Таблица 3. Аберрации осевой точки системы Кассегрена с компенсатором комы и астигматизма

Размер предм., $\sin \sigma$	Кома	Астигматизм	Кривизна поверхности изображения
0.0248	-0.0045	-0.183	0.056
0.0350	-0.024	-0.707	0.216

Таблица 4. Аберрации внеосевой точки системы Кассегрена с компенсатором комы и астигматизма

Аберрации осевой точки системы Ричи-Кретьена без компенсатора. Параметры системы: $f' = 1000$ мм, $2\omega = 4^\circ$, $\eta = 0.3$.

Отн. зр. коор.	Продольные aberrации $\Delta S'$, мм	Поперечные aberrации $\Delta y'$, мм	Волновые aberrации W , дл. волн	Неизопланатизм η , (%)
1	0.0243	0.00306	0.118	0.00640
0.866	0.0138	0.00151	0.0505	0.00363
0.707	0.00626	0.000555	0.0153	0.00162
0.5	0.00163	0.000102	0.00201	0.000410

Таблица 5. Абберации осевой точки системы Ричи-Кретьена без компенсатора

Размер предм., $\sin \sigma$	Кома	Астигматизм	Кривизна поверхности изображения
0.0248	0.012	-1.73	- 0.655
0.0350	0.023	-3.46	- 1.31

Таблица 6. Абберации внеосевой точки системы Ричи-Кретьена без компенсатора

Абберации системы Ричи-Кретьена с компенсатором кривизны поверхности изображения и астигматизма. Параметры системы: $f' = 1000$ мм, $2\omega = 4^\circ$, $\eta = 0.3$.

Отн. зр. коор.	Продольные aberrации $\Delta S'$, мм	Поперечные aberrации $\Delta y'$, мм	Волновые aberrации W , дл. волн	Неизопланатизм η , (%)
1	0.0439	0.00553	0.149	0.0133
0.866	0.0200	0.00217	0.0369	0.00809
0.707	0.00458	0.407-3	- 0.0046	0.00414
0.5	- 0.0021	- 0.13-3	- 0.0064	0.00144

Таблица 7. Абберации осевой точки системы Ричи-Кретьена с компенсатором кривизны поверхности изображения и астигматизма

Размер предм., $\sin \sigma$	Кома	Астигматизм	Кривизна поверхности изображения
0.0248	-0.0644	0.111	-0.0853
0.0350	-0.2242	-0.632	0.178

Таблица 8. Абберации внеосевой точки системы Ричи-Кретьена с компенсатором кривизны поверхности изображения и астигматизма

СХЕМА КОНТРОЛЯ ВЫПУКЛЫХ НЕСФЕРИЧЕСКИХ ПОВЕРХНОСТЕЙ ВРАЩЕНИЯ ВТОРОГО ПОРЯДКА МЕТОДОМ АНАБЕРРАЦИОННЫХ ТОЧЕК

Е.В. Ермолаева

Научный руководитель – д.т.н., профессор В.А. Зверев

Зеркальные системы состоят, как правило, из ограниченного числа отражающих поверхностей. Основным коррекционным параметром при этом является коэффициент асферизации поверхностей. Основной проблемой при изготовлении несферических поверхностей является контроль их формы.

Идеальной схемой контроля является схема с использованием анаберрационных точек. Если поверхность имеет идеальную форму, а точечный источник света расположен точно в одном из геометрических фокусов, то лучи, отраженные от поверхности, образуют безабберационное изображение источника в другом геометрическом фокусе. По величине и знаку деформации волнового фронта можно определить погрешности асферической поверхности, вызвавшие эту деформацию.

Однако обычно один из фокусов труднодоступен, и для реализации контроля требуются дополнительные оптические элементы. Особую сложность представляет контроль выпуклых поверхностей. Лучи, отраженные от выпуклой поверхности, образуют расходящийся пучок, поэтому дополнительный элемент должен быть, по крайней мере, того же размера, что и испытываемая деталь.

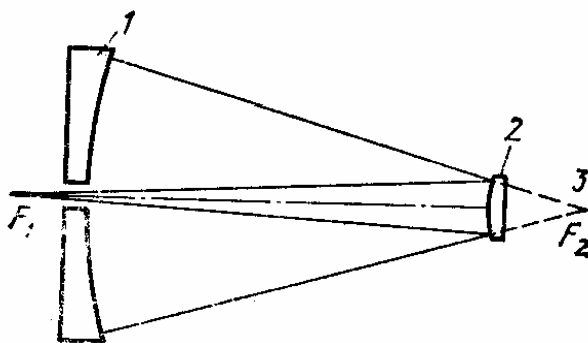


Рис. 1. Автоколлимационная схема контроля выпуклой поверхности гиперboloида вращения

Схемой с использованием анаберрационных точек является схема Хиндла для контроля выпуклой поверхности гиперboloида вращения. Хиндл показал, как посредством отражения от сферы с центром вблизи недоступного фокуса гиперболического зеркала можно реализовать автоколлимационную схему контроля (рис. 1).

Аналогичные схемы существуют для контроля выпуклой поверхности параболоида вращения, выпуклой поверхности эллипсоида вращения, а также для контроля вогнутой поверхности гиперboloида вращения. При контроле эллипсоидов необходимо использовать дополнительный объектив, формирующий пучок, сходящийся в одном из фокусов, а при контроле параболоидов – коллиматор (рис. 2–4).

Недостатком представленных схем является наличие экранирования. Центральная часть испытываемого зеркала остается неисследованной. Чтобы она не была большой, расстояние между зеркалами должно быть значительным, а, значит, и контрольное зеркало должно иметь большой диаметр.

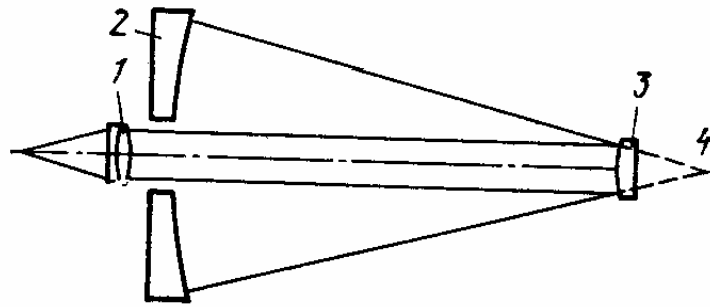


Рис. 2. Автоколлимационная схема контроля выпуклой поверхности параболоида вращения

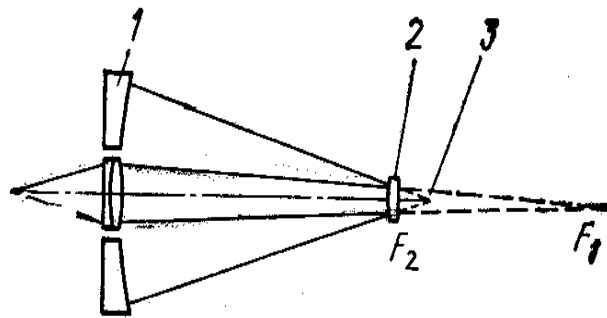


Рис. 3. Автоколлимационная схема контроля выпуклой поверхности эллипсоида вращения

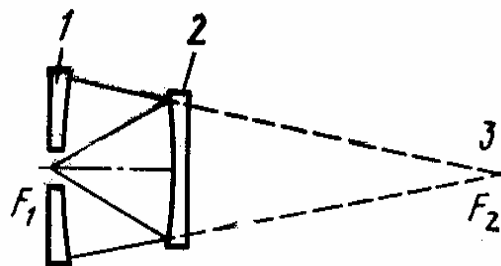


Рис. 4. Автоколлимационная схема контроля вогнутой поверхности гиперboloида вращения

Сферическую поверхность контрольного зеркала необходимо изготовлять с высокой точностью, так как она принимает участие в формировании волнового фронта. Как видно из представленных схем, лучи дважды отражаются от контролируемой поверхности и один раз от контрольного сферического зеркала. Таким образом, точность поверхности контрольного сферического зеркала может быть в два раза ниже заданной точности контролируемой поверхности.

Причиной деформации волнового фронта может быть не только погрешность изготовления контролируемой несферической поверхности, но и погрешность совмещения точечного источника света с одним из геометрических фокусов – расфокусировка. Отделить погрешность изготовления несферической поверхности от погрешности установки практически невозможно. Поэтому предъявляются высокие требования к взаимному положению элементов в контрольной схеме. Таким образом, осуществление этих схем на практике является весьма проблематичным.

Поэтому представляет интерес схема с менисковым компенсатором для контроля выпуклой поверхности гиперboloида вращения. Данная схема представлена на рис. 5.

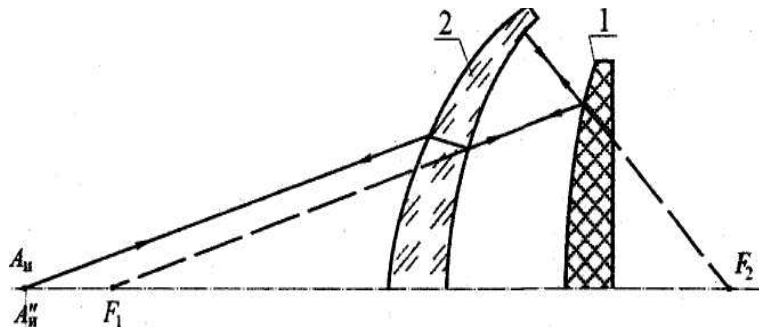


Рис. 5. Автоколлимационная схема контроля выпуклой поверхности гиперboloида вращения с мениском (компенсатором)

Схема представляет собой модификацию схемы Хиндла. На этой схеме в качестве компенсатора применен мениск 2, вогнутая поверхность которого, расположенная непосредственно вблизи поверхности гиперboloида, играет роль вспомогательной сферы, концентричной геометрическому фокусу F_2 контролируемого гиперboloида 1, а радиус кривизны выпуклой поверхности мениска служит коррекционным параметром для исправления сферической aberrации в изображении геометрического фокуса F_1 , образованного вогнутой поверхностью мениска. Основное преимущество ее в том, что диаметр вспомогательной линзы практически равен диаметру контролируемого зеркала. В схеме полностью отсутствует центральное экранирование.

Схема была предложена профессором Пуряевым для случая мениска с равными радиусами. Подобная схема была также разработана Симпсоном, Оландом и Меккелем.

При произвольной толщине мениска и при равных радиусах кривизны его поверхностей рассматриваемая оптическая система позволяет контролировать форму поверхности гиперboloида вращения при $3 \leq e < \infty$.

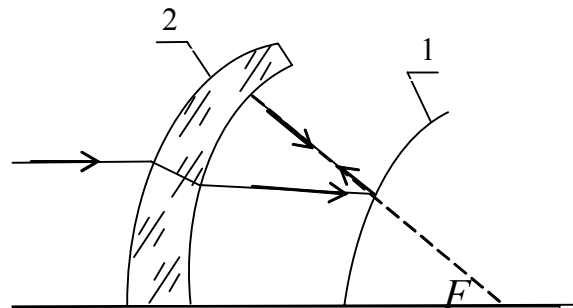


Рис. 6. Автоколлимационная схема контроля выпуклой поверхности параболоида вращения с мениском (компенсатором): 1 – мениск; 2 – выпуклое параболическое зеркало

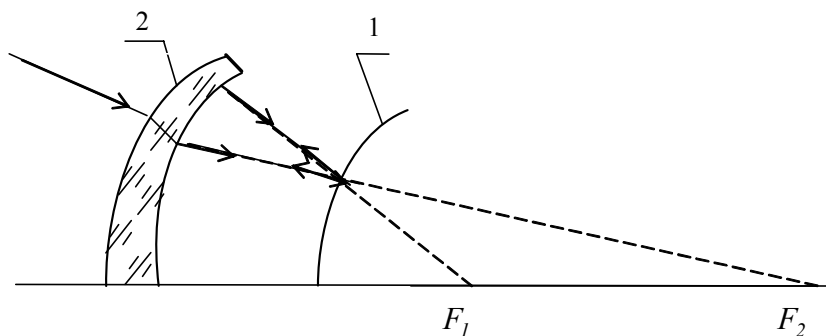


Рис. 7. Автоколлимационная схема контроля выпуклой поверхности эллипсоида вращения с мениском (компенсатором): 1 – мениск; 2 – выпуклое эллиптическое зеркало

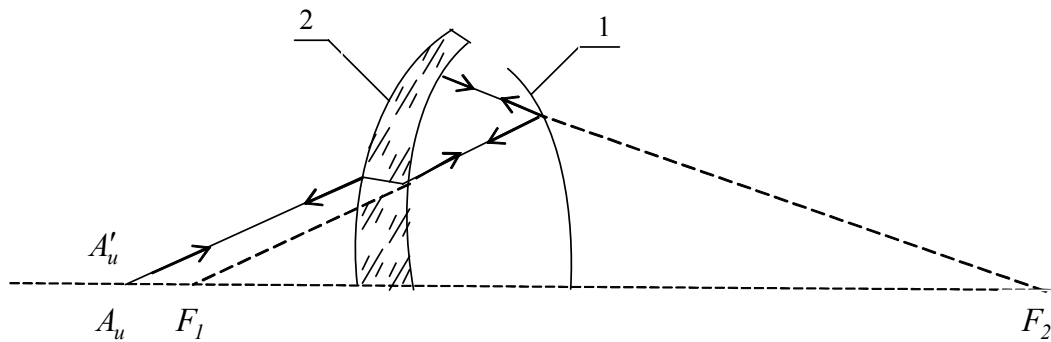


Рис. 8. Автоколлимационная схема контроля вогнутой поверхности гиперboloида вращения с мениском (компенсатором): 1 – мениск; 2 – вогнутое гиперболическое зеркало

В случае неравных радиусов кривизны поверхностей мениска можно контролировать форму поверхности гиперboloида вращения при $1 < e < 3$. Показано, что при этом конструктивные параметры оптической системы зависят от толщины мениска; с ростом толщины мениска плоскость изображения удаляется. Поэтому в области малых эксцентриситетов следует применять мениски с наименьшей толщиной.

Мы расширили область применения схемы контроля с менисковым компенсатором, доказав возможность ее использования для контроля выпуклой поверхности параболоида вращения, выпуклой поверхности эллипсоида вращения, а также для контроля вогнутой поверхности гиперboloида вращения. На рис. 6–8 представлены схемы для контроля формы этих поверхностей.

Заменив мениском контрольное зеркало в схемах, рассмотренных ранее для этих поверхностей, получаем оптические системы, свободные от экранирования, при этом диаметр дополнительного оптического элемента в них не превышает диаметра контролируемой поверхности.

Для данных поверхностей нами получены схемы контроля, диаметр дополнительного оптического элемента в которых практически не превышает диаметра контролируемой несферической поверхности. При этом полностью отсутствует экранирование.

Аберрации, вносимые мениском достаточно малы, чтобы контролировать выпуклые поверхности гиперboloида вращения с высокой точностью. Если требуется более высокая точность, в данную схему можно включить киноформный элемент.

ЗЕРКАЛЬНО-ЛИНЗОВЫЙ КОНЦЕНТРИЧЕСКИЙ ОБЪЕКТИВ

Е.В. Ермолаева, К.Ю. Мацерук, А.А. Репин
 Научный руководитель – д.т.н., профессор В.А. Зверев

Исследовано влияние применения пластинки с конической преломляющей поверхностью в зеркально-линзовых объективах на поперечную сферическую aberrацию.

В простейшем случае зеркальная оптическая система состоит из одной отражающей поверхности. Схема такой системы, состоящей из сферической и плоской отражающих поверхностей, представлена на рис. 1.

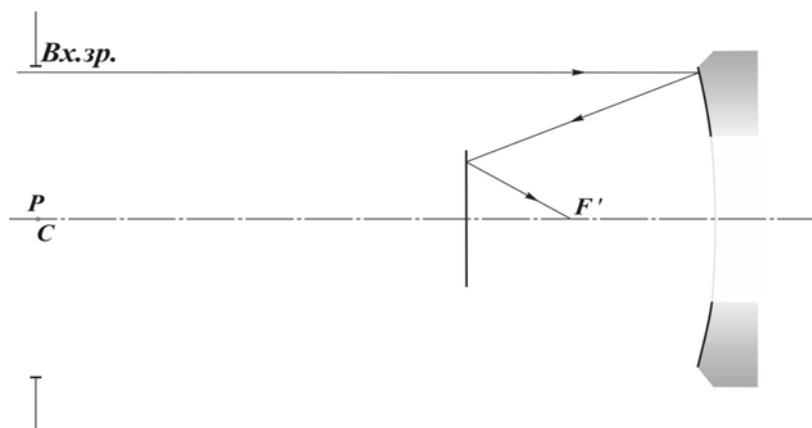


Рис. 1. Пример зеркально-линзовой оптической системы

Будем считать, что центр входного зрачка системы совмещен с центром кривизны C отражающей поверхности. Параметры системы приведен в табл. 1, а остаточные aberrации в изображении, образованном рассматриваемой системой при $f_0'=50$ мм, $D: f_0'=1:1$ и $2\omega = 2^\circ 30'$, приведены в табл. 2.

Фокусное расстояние в мм	50.0000
Дифрагменное число	1.00000
Угловое поле (2 омега)	2°30'00"
Положение предмета относительно первой поверхности	Бесконечность
Положение изображения относительно последней поверхности	15.0000
Диаметр входного зрачка	50.0000
Положение входного зрачка относительно первой поверхности	-100.000
Положение выходного зрачка относительно последней поверхности	65.0000
Основная длина волны в нм	2600.00
Диапазон ахроматизации	1800.00 – 3200.00

№ пов.	Радиусы	D	Конструктивные параметры			Стрелки
			Марки стекол	Показатели преломления	Световые диаметры	
				1.000000		
1	-100.000	-35.00		-1.000000	54.22	-3.75
2	0.000			1.000000	17.83	0.00

Таблица 1. Паспорт оптической системы

S0	Z0	Z'0(0)	S'0(0)	F'(0)	V3P(0)
	-100.0	65.00	15.000	50.000	-1.000
MU=0	H	DS'(0)	TGC'	Y'(0)	W(0)
	25.0	-1.64	0.553	-0.907	-38.8
	21.7	-1.21	0.466	-0.567	-21.6
	17.7	-0.800	0.371	-0.297	-9.54
	12.5	-0.395	0.256	-0.101	-2.37

Таблица 2. Остаточные aberrации

По данным табл. 2 построена кривая сферической aberrации, представленная на рис. 2. Величину остаточной aberrации можно существенно уменьшить, если сместить плоскость наблюдения изображения относительно плоскости Гаусса на некоторое расстояние Δ' . При этом изменение величины поперечной сферической aberrации определяется соотношением вида:

$$\Delta g \Delta' = \Delta' \operatorname{tg} \sigma' . \quad (1)$$

Таким образом, в результате смещения плоскости наблюдения остаточная сферическая aberrация в изображении точки будет равна:

$$\Delta g'_{\text{ост}} = \delta g' - \Delta' \operatorname{tg} \sigma' . \quad (2)$$

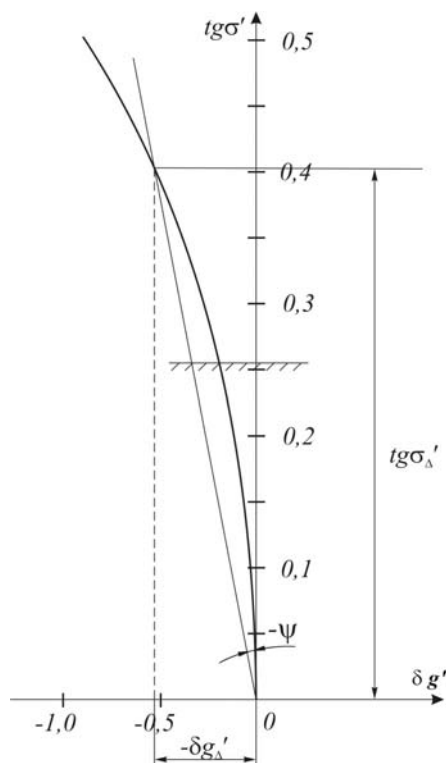


Рис. 2. Кривая сферической aberrации по данным табл. 2

Пусть при некотором значении угла $\sigma' = \sigma \Delta'$ остаточная aberrация $\delta g'_{\text{ост}} = \delta g \Delta' - \Delta' \operatorname{tg} \sigma \Delta' = 0$, (3)

тогда

$$\Delta' = \delta g \Delta' / \operatorname{tg} \sigma \Delta' . \quad (4)$$

На рис. 2 величине Δ' соответствует тангенс угла ψ между линией $\delta g' = \Delta' \operatorname{tg} \sigma'$ и осью ординат, т.е. $\Delta' = \operatorname{tg} \psi$. В результате получаем, что после смещения плоскости наблюдения изображения на расстояние Δ' кривая зависимости $\delta g'_{\text{ост}} = \delta g'_{\text{ост}}(\operatorname{tg} \sigma')$ примет вид, показанный на рис. 3. Вполне очевидно, что эта кривая выходит из начала координат под углом, равным ψ , к оси ординат. Кривая дает наглядное представление об эффективности смещения плоскости наблюдения изображения для уменьшения остаточных aberrаций. Пусть в рассматриваемой оптической системе коэффициент экраниро-

вания зрачка по диаметру $\eta = 0,5$. Введем смещение плоскости наблюдения, при котором поперечная сферическая aberrация на краю зрачка равна сферической aberrации с обратным знаком на краю экранируемой зоны зрачка. При этом при смещении $\Delta' = -1,246$ мм изображение осевой точки будет обладать сферической aberrацией, величины которой приведены в табл. 3.

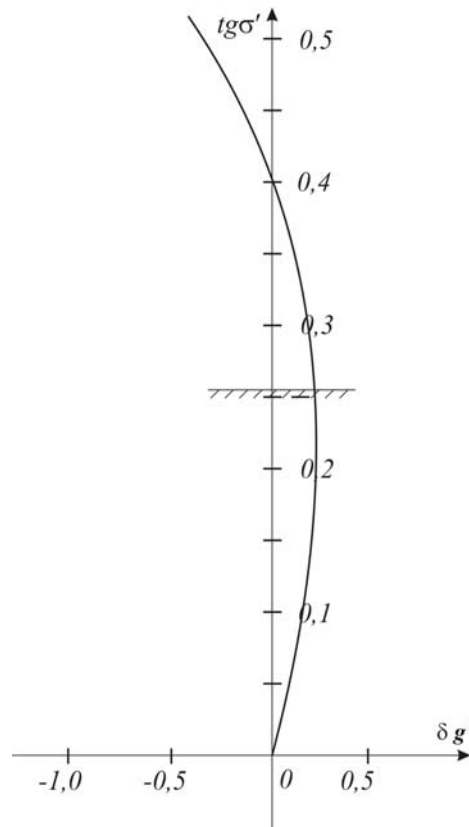


Рис. 3. Кривая зависимости $\delta g'_{\text{ост}} = \delta g'_{\text{ост}}(\text{tg}\sigma')$

S0	Z0	Z'0(0)	S'0(0)	F'(0)	VЗР(0)
	-100.0	65.00	15.000	50.000	-1.000
MU=0	H	ДС'(0)	TGC'	Y'(0)	W(0)
	25.0	-0.394	0.553	-0.218	21.1
	21.7	0.0313	0.466	0.0146	23.3
	17.7	0.446	0.371	0.166	20.4
	12.5	0.851	0.256	0.218	12.6

Таблица 3. Остаточные aberrации при $\Delta' = -1.246$ мм

Для оптических систем с центральным экранированием зрачка в основу дальнейшего улучшения изображения могут быть положены следующие соображения. Пусть кривая поперечной сферической aberrации имеет вид, показанный на рис. 4.

Через крайнюю точку кривой А и точку В, соответствующую крайней точке экрана, проведем прямую, пересекающую ось абсцисс в точке с координатами $\delta g'_0, 0$. Из рис. 4 следует

$$\frac{\delta g'_K - \delta g'_Э}{\text{tg}\sigma'_K - \text{tg}\sigma'_Э} = \frac{\delta g'_K - \delta g'_Э}{\text{tg}\sigma_K} \quad (5)$$

Отсюда находим, что

$$\delta g'_0 = \delta g'_K - \frac{\delta g'_K - \delta g'_Э}{\text{tg}\sigma'_K - \text{tg}\sigma'_Э} \cdot \text{tg}\sigma_K \quad (6)$$

или

$$\delta g_o' = \frac{\delta g_{\sigma_3}' \operatorname{tg} \sigma_{\kappa}' - \delta g_{\kappa}' \operatorname{tg} \sigma_{\sigma_3}}{\operatorname{tg} \sigma_{\kappa}' - \operatorname{tg} \sigma_{\sigma_3}}. \quad (7)$$

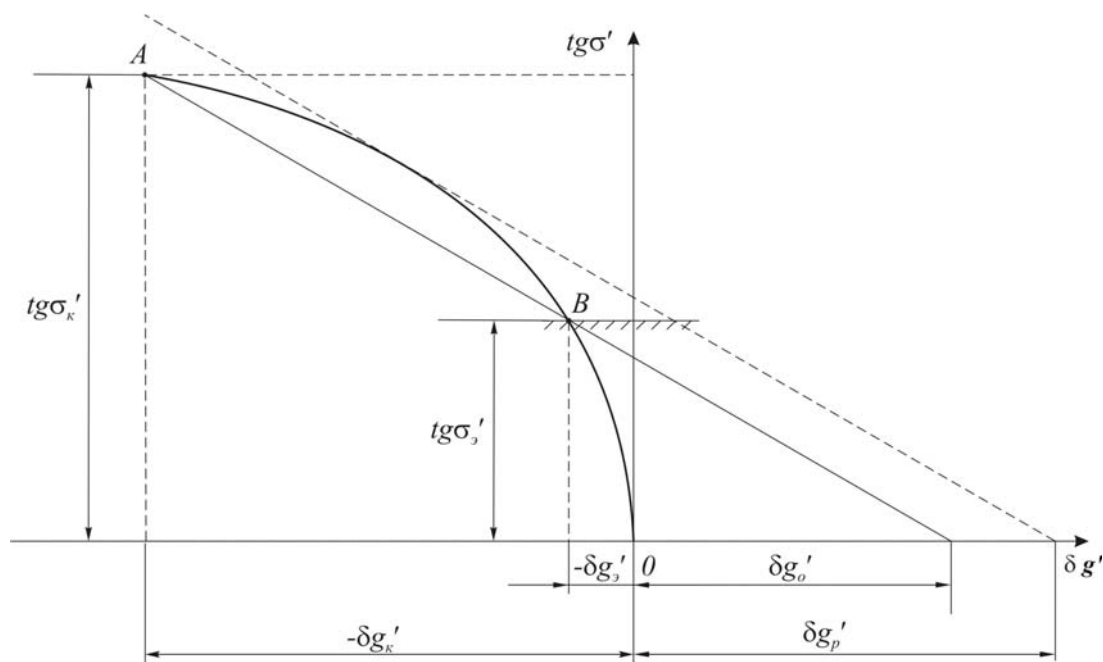


Рис. 4. Кривая поперечной сферической аберрации

В этом случае геометрическое изображение точки будет иметь вид кольца шириной, равной разности $\delta g'_{\rho} - \delta g'_{\sigma}$, т.е. расстоянию между пунктирной и сплошной наклонными линиями, как показано на рисунке, при этом радиус внутренней зоны кольца равен $\delta g'_{\sigma}$. Для определения положения плоскости, в которой изображение точки имеет вид узкого кольца, обратимся к рис. 5, из которого находим:

$$-\Delta' = -\Delta'_{SK} + \frac{\delta g'_{\sigma}}{\operatorname{tg} \sigma_{\kappa}'} = -\Delta'_{S\sigma_3} + \frac{\delta g'_{\sigma}}{\operatorname{tg} \sigma_{\sigma_3}'}.$$

Эти соотношения можно представить в виде

$$\Delta' = \frac{\delta g_{\kappa}' - \delta g_{\sigma}'}{\operatorname{tg} \sigma_{\kappa}} = \frac{\delta g_{\sigma_3}' - \delta g_{\sigma}'}{\operatorname{tg} \sigma_{\sigma_3}'}. \quad (8)$$

При этом в общем случае имеем

$$\delta g' = \delta g'_{\sigma} + \Delta' \operatorname{tg} \sigma'. \quad (9)$$

Подставив соответствующие величины из табл. 2 (при $\eta = 0,5$) в формулу (6) или (7), получаем $\delta g'_{\sigma} = 0,594$ мм. При этом, используя формулу (8), находим, что $\Delta' = -2,714$ мм. Величины аберраций в изображении осевой точки при $s' = s'_0 + \Delta' = 15,0 - 2,714 = 12,286$ мм, приведены в табл. 4.

Из анализа данных таблицы следует, что поперечная сферическая аберрация в изображении точки измеряется в пределах кольца от $\delta g' = 0,594$ мм до $\delta g' = 0,71$ мм, что можно записать в виде: $\delta g'_{\text{ср}} \pm 0,058$ мм, где $\delta g'_{\text{ср}} = 0,652$ мм. Замена величины $\delta g'_{\sigma}$ величиной $\delta g'_{\text{ср}}$ эквивалентна параллельному смещению наклонной линии на рис. 4; при этом наибольшее значение положительной сферической аберрации будет равно наименьшему значению отрицательной, а сохранение наклона линии определит сохранение положения плоскости наблюдения изображения. Для получения изображения точки в виде пятна, радиус которого равен $\delta g' = 0,058$ мм, необходимо окружность радиуса $\delta g'_{\text{ср}} = 0,652$ мм «стянуть» в точку. Для этого применим коническую преломляющую

поверхность, вершину конуса которой совместим с центром кривизны С отражающей поверхности.

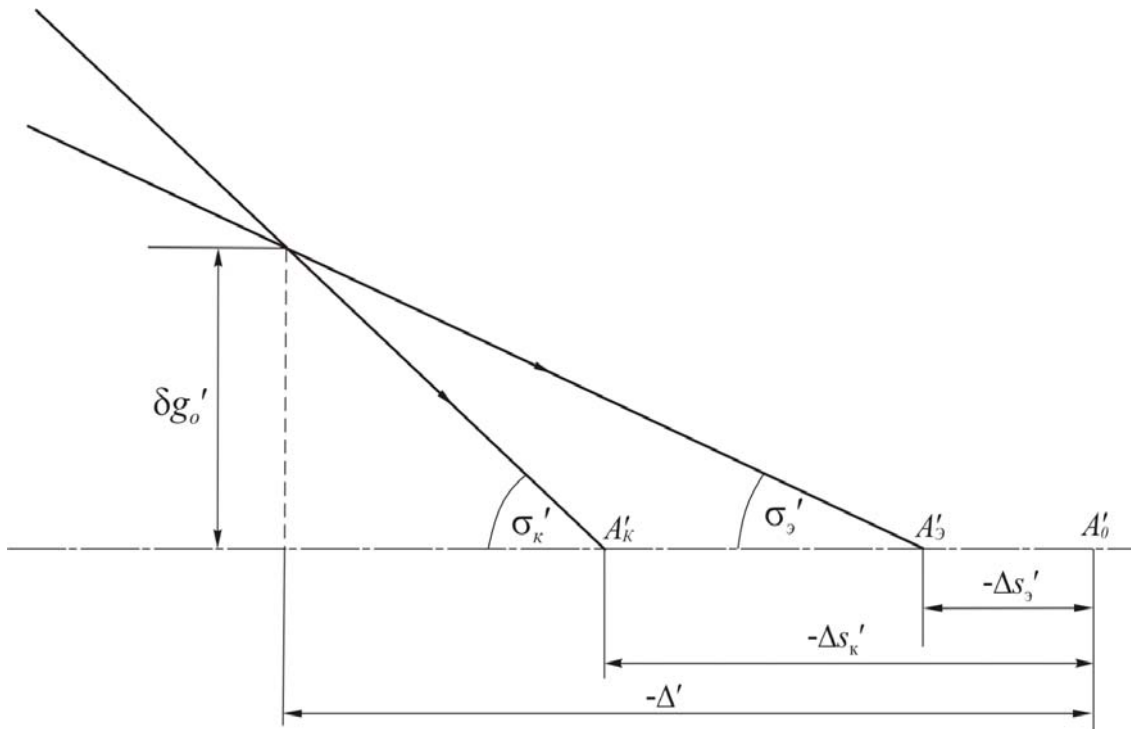


Рис. 5. Геометрическое изображение точки в виде кольца

S0	Z0	Z'0(0)	S'0(0)	F'(0)	V3P(0)	S'A
	-100.0	65.00	15.000	50.000	-1.000	12.286
MU=0	H	DS'(0)	TGC'	Y'(0)	W(0)	ЭТА %
	25.0	1.07	0.553	0.594	91.7	0
	21.7	1.50	0.466	0.699	76.2	0
	17.7	1.91	0.371	0.710	55.7	0
	12.5	2.32	0.256	0.594	30.3	0

Таблица 4. Остаточные aberrации

Коническая преломляющая поверхность не имеет параксиальной области, и формальная возможность ее применения не очевидна. Однако известно, что коническая поверхность представляет собой асимптотическую поверхность гиперboloида. Сечение поверхности второго порядка меридиональной плоскостью можно определить уравнением вида:

$$y^2 = 2rz - (1+b)z^2, \quad (10)$$

где $b = -e^2$, e – эксцентриситет кривой.

Тангенс угла наклона (угла с осью OZ) касательной к кривой равен

$$\operatorname{tg} \gamma = \frac{dy}{dz} = \frac{r - (1+b)z}{yz} = \frac{y^2 - rz}{yz};$$

$$\operatorname{tg} \gamma = -\frac{r}{y} + \frac{y}{z}. \quad (11)$$

При $\left| \frac{r}{y} \right| \ll 1$, $\left| \frac{r}{y} \right| \gg 1$ можно принять, что

$$\operatorname{tg} \gamma \approx \frac{y}{z}. \quad (12)$$

Прямая линия, определяемая уравнением $y = z \cdot \operatorname{tg} \gamma$, является образующей конуса, уравнение которого можно записать в виде

$$x^2 + y^2 = z^2 \cdot \operatorname{tg}^2 \gamma. \quad (13)$$

Запишем уравнение (10) в виде

$$\frac{y^2}{z^2} = 2 \frac{r}{z} - (1+b).$$

Но $\left| \frac{y}{z} \right| \gg 2 \left| \frac{r}{z} \right|$. Тогда

$$1+b \approx -\frac{y^2}{z^2} = -\operatorname{tg}^2 \gamma. \quad (14)$$

Тогда уравнение (10) можно представить в виде

$$y^2 = 2rz + z^2 \operatorname{tg}^2 \gamma. \quad (15)$$

В сечении меридиональной плоскости пластинки, образованной сочетанием плоской и конической преломляющих поверхностей, будем иметь два клина равной величины, но разного знака. Пусть α – угол одного из клиньев. Тогда будем иметь $\alpha + \gamma = 90^\circ$. При этом $\operatorname{tg} \gamma = \operatorname{ctg} \alpha$. Угол отклонения луча клином от начального направления равен

$$\omega = (n-1)\alpha, \quad (16)$$

где n – показатель преломления материала пластины. Требуемый угол отклонения параллельного пучка лучей равен

$$\omega = -\frac{\delta g'_o}{f'_o}. \quad (17)$$

Заметим, что знак угла ω определяет знак радиуса кривизны в вершине гиперболы, описываемой уравнением (15).

Используя данные табл. 4, имеем:

$$\delta g'_o = \delta g'_{cp} = 0,652 \text{ мм, а } f'_o = 50 \text{ мм.}$$

При этом

$$\omega = -\frac{0,652}{50} = -0,01304.$$

При $n=1,42053$: $\alpha = -0,031$; $\operatorname{ctg} \alpha = \operatorname{tg} \gamma = -32,25$.

№ пов.	Радиусы кривизны	Осевые расст.	Световые высоты	Марки стекол	Показатели преломления		
					Воздух		
1	0.000	5.000		CAF2	1.420526	1.416424	1.424881
2	-0.020*	100.000		Воздух			
3	-100.000	-35.000		Воздух			
4	0.000			Воздух			

* Несферическая поверхность вращения $e^2 = 1218.00000$

ПРЕДМЕТ: Удал. размер $Y = 1.15000$ гр.мнсек $S_0 = 0.000$ дптр от вх.зрачка

ИЗОБРАЖ: Близ. размер $Y' = 0.00000$ плоск. $SI = 12.320$ мм от посл.пов.

ДИАФР.: $ND = 2$ совп. с 2-й пов. $SD = 0.0000$ мм передн. апертура = 25.0000 мм

Таблица 5. Конструктивные параметры оптической системы

Подставив эти значения $\operatorname{tg} \gamma$ в уравнение (15) и получив $r = -0,01$ мм, имеем:

$$x^2 + y^2 = -0,02z - (1 - e^2), \quad (18)$$

где $e^2 = 1041$. Конструктивные параметры рассматриваемой системы, дополненной пластинкой с конической поверхностью, приведены в табл. 5, а остаточные аберрации в

изображении, образованном неэкранируемой частью оптической системы, полученные при $s'=12,32$ мм, представлены в табл. 6.

Из сопоставления величин aberrаций в табл. 2 и 6 следует, что изменение конической преломляющей поверхности позволило уменьшить величину поперечных aberrаций не меньше, чем в 15 раз.

S0	Z0	Z'0(0)	S'0(0)	F'(0)	V3P(0)	S'A
	-100.0	65.00	15.000	50.000	-1.000	12.320
MU=0	H	DS'(0)	TGC'	Y'(0)	W(0)	ЭТА %
	25.0	-0.113	0.537	-0.0608	-26.9	0
	21.7	0.103	0.452	0.0465	-26.9	0
	17.7	0.170	0.357	0.0609	-28.7	0
	12.5	-0.210	0.243	-0.0512	-29.3	0

Таблица 6. Остаточные aberrации

Рассмотрим систему концентрических поверхностей, две из которых образуют мениск, а две – двухзеркальную систему, при этом будем считать, что вторая поверхность мениска и вторая поверхность зеркальной системы физически представляют собой одну и ту же поверхность, как показано на рис. 6.

Такую систему поверхностей можно записать в виде:

$$\begin{aligned}
 n_1 &= 1 \\
 r_1 &= \\
 d_1 &= r_1 - r_2 \quad n_2 = n \\
 r_2 &= \\
 d_2 &= r_2 - r_3 \quad n_3 = 1 \\
 r_3 &= \\
 d_3 &= r_3 - r_4 = -d_2 \quad n_4 = -1 \\
 r_4 &= r_2 \\
 n_5 &= 1
 \end{aligned}$$

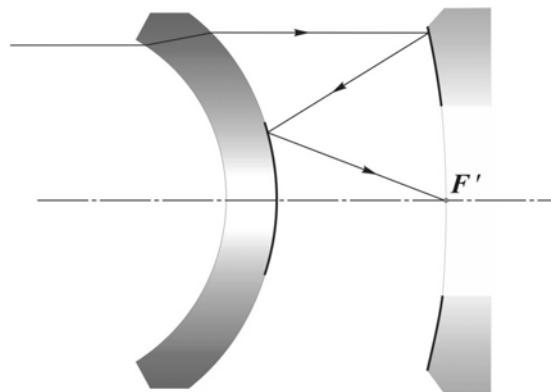


Рис. 6. Пример концентрической зеркально-линзовой оптической системы

Оптическая сила концентрического мениска равна

$$\Phi_m = (n-1) \left(\frac{1}{r_1} - \frac{1}{r_2} \right) + \frac{(n-1)d_1^2}{nr_1r_2} = -\frac{n-1}{n} \frac{d_m}{r_1r_2}. \quad (19)$$

Для концентрической системы из двух отражающих поверхностей (при $n=-1$) имеем

$$\Phi_{zc} = \frac{2}{r_3r_4} d_{zc} - \frac{4}{r_3r_4} d_{zc} = -2 \frac{d_{zc}}{r_3r_4}. \quad (20)$$

Поскольку центры кривизны всех поверхностей лежат в одной и той же точке, то и главные плоскости лежат в одной и той же плоскости, проходящей через общий центр кривизны поверхностей. При этом оптическая сила зеркально-линзовой системы равна

$$\varphi_{злс} = \varphi_{.м} + \varphi_{зс} = -\frac{n-1}{n} \frac{d_{.м}}{r_1 r_2} - 2 \frac{d_{зс}}{r_3 r_4}.$$

Но $d_{.м} = r_1 - r_2$, а $d_{зс} = r_3 - r_4$. Получив $r_4 = r_2$, имеем:

$$\varphi_{злс} = \frac{n-1}{nr_1} - \frac{3n-1}{nr_2} + \frac{2}{r_3}. \quad (21)$$

При заданной величине оптической силы $\varphi_{злс}$ свободными параметрами остаются показатель преломления материала мениска и два из трех радиусов кривизны поверхностей.

При выборе радиусов кривизны поверхностей можно руководствоваться следующими габаритными ограничениями. Выразим задний фокальный отрезок в расстояниях $d_{зс}$ между отражающими поверхностями через коэффициент k_s в виде:

$$s'_{f'} = -k_s d_{зс} = k_s (r_2 - r_3). \quad (22)$$

При этом расстояние от главной плоскости до заднего фокуса равно:

$$f'_{злс} = -r_2 + s'_{f'} = (k_s - 1)r_2 - k_s r_3. \quad (23)$$

Отсюда находим что:

$$r_3 = -r_2 + s'_{f'} = (k_s - 1)r_2 - k_s r_3, \quad (24)$$

где $\tilde{r}_2 = r_2 \varphi_{злс}$. Заметив, что при $k_s = 1$ $r_3 = -f'_{злс}$.

Умножив выражение (21) на r_2 и подставив соотношение (24), получаем:

$$\tilde{r}_2 = \frac{n-1}{n} \frac{\tilde{r}_2}{r_2} - \frac{3n-1}{n} - \frac{2k_s \tilde{r}_2}{1 - (k_s - 1)\tilde{r}_2}. \quad (25)$$

Преобразуем это выражение к виду

$$\frac{\tilde{r}_2}{\tilde{r}_1} = \frac{3n-1}{n-1} + \frac{n}{n-1} \tilde{r}_2 \left[1 + \frac{2k_s}{1 - (k_s - 1)\tilde{r}_2} \right]. \quad (26)$$

В рассматриваемом случае $\frac{\tilde{r}_2}{\tilde{r}_1} \geq 1$. При этом из условия

$$\frac{3n-1}{n-1} + \frac{n}{n-1} \tilde{r}_2 \left[1 + \frac{2k_s}{1 - (k_s - 1)\tilde{r}_2} \right] \geq 1$$

находим что

$$\tilde{r}_2 \left[1 + \frac{2k_s}{1 - (k_s - 1)\tilde{r}_2} \right] \geq -2. \quad (27)$$

Из выражения (24) имеем

$$1 - (k_s - 1)\tilde{r}_2 = -k_s \tilde{r}_3.$$

Но $\tilde{r}_3 < 0$. Следовательно, $1 - (k_s - 1)\tilde{r}_2 > 0$. Учитывая это, выражение (27) можно представить в виде:

$$(k_s - 1)\tilde{r}_2^2 - 3\tilde{r}_2 - 2 \leq 0. \quad (28)$$

При $k_s \neq 1$ в частном случае выражение (28) принимает вид уравнения:

$$\tilde{r}_2^2 - \frac{3}{k_s - 1} \tilde{r}_2 - \frac{2}{k_s - 1} = 0,$$

решение которого можно записать в виде

$$\tilde{r}_2 = \frac{3}{2(k_s - 1)} \pm \sqrt{\frac{9}{4(k_s - 1)^2} + \frac{2}{k_s - 1}}.$$

Условию $r_2 < 0$ удовлетворяет решение:

$$\tilde{r}_2 = \frac{1}{2} \frac{3}{k_s - 1} \left(1 - \sqrt{1 + 8 \frac{k_s - 1}{9}} \right). \quad (29)$$

Из выражения (28) следует, что при $k_s = 1$ должно соблюдаться условие $\tilde{r}_2 \geq -\frac{2}{3}$.

Следовательно, в общем случае формулу (29) можно представить условием

$$\tilde{r}_2 \geq \frac{1}{2} \frac{3}{k_s - 1} \left(1 - \sqrt{1 + 8 \frac{k_s - 1}{9}} \right). \quad (30)$$

Выбрав величину \tilde{r}_2 , удовлетворяющую условию (30), и подставив в выражение (24), получаем величину

$$\tilde{r}_3 = -\frac{1 - (k_s - 1)\tilde{r}_2}{k_s}. \quad (31)$$

И, наконец, из выражения (26) находим, что

$$\tilde{r}_1 = \frac{1}{N} \tilde{r}_2, \quad (32)$$

где
$$N = \frac{3n-1}{n-1} + \frac{n}{n-1} \tilde{r}_2 \left[1 + \frac{2k_s}{1 - (k_s - 1)\tilde{r}_2} \right].$$

Полученные соотношения существенно упрощаются, если принять $k_s = 1$. При этом выражение (31) примет вид: $\tilde{r}_3 = -1$; из выражения (28) следует, что $\tilde{r}_2 \geq -\frac{2}{3}$, а в выражении (32) величина N становится равной

$$N = \frac{3n(1 + \tilde{r}_2) - 1}{n - 1}. \quad (33)$$

Заметим, что при $\tilde{r}_2 \geq -\frac{2}{3}$ и $k_s = 1$ величина $N = 1$. При этом $\tilde{r}_3 = \tilde{r}_2$, т.е. в этом случае мениск вырождается, и остается лишь одна система из двух отражающих концентрических поверхностей.

Пусть $k_s = 1$. Тогда при $f' = 50$ мм $r_3 = -50$ мм; принимаем $\tilde{r}_2 = -0,65 > -\frac{2}{3}$, т.е. принимаем $r_2 = -32,5$ мм. При $n = 1,42053$ (флюорит, $\lambda = 2,6$ мкм) величина $N = 1,169$. Следовательно, $r_1 = -27,8$. В результате вычислений оказалось, что кривая зависимости $\delta g' = \delta g'(\text{tg}\sigma')$ более линейна в неэкранируемой зоне зрачка при $r_1 = -25,6$ мм, $r_2 = r_4 = -32,1$ мм, $r_3 = -50$. Конструктивные параметры системы приведены в табл. 7, а остальные aberrации – в табл. 8.

Подставив соответствующие величины из табл. 8 (при $\eta = 0,5$) в формулу (6) или (7), получаем $\delta g'_o = 0,402$ мм. При этом, применив формулу (8), находим, что $\Delta' = -2,633$ мм. Величины aberrаций в изображении осевой точки при $s' = s_o' + \Delta' = 17,991 - 2,633 = 15,358$ мм приведены в табл. 9. Из анализа величин, приведенных в таблице, следует, что поперечная сферическая aberrация в изображении точки изменяется в пределах кольца от $\delta g' = 0,359$ мм до $\delta g' = 0,408$ мм, что можно записать в виде: $\delta g' = \delta g'_{cp} \pm 0,025$ мм, где $\delta g'_{cp} = 0,384$ мм. Используя данные табл. 9, имеем $\delta g'_o = \delta g'_{cp} = 0,384$ мм. При этом при $f_o' = 50,091$ мм угол

$$\omega = -\frac{0,384}{50,091} = -0,007666.$$

Фокусное расстояние в мм	50.0000
Дифрагменное число	1.25000
Угловое поле (2 омега)	2°30'00"
Положение предмета относительно первой поверхности	бесконечность
Положение изображения относительно последней поверхности	18.0000
Диаметр входного зрачка	40.0000
Положение входного зрачка относительно первой поверхности	4.2600
Положение выходного зрачка относительно последней поверхности	-13.4000
Основная длина волны в нм	2600.00
Диапазон ахроматизации	1800.00 – 3200.00

№ пов.	Радиусы	D	Конструктивные параметры			Стрелки
			Марки стекло	Показатели преломления	Световые диа- метры	
				1.000000		
1	-25.600	6.50	Флюорит	1.420526	40.64	-10.03
2	-32.100	17.90		1.000000	45.31	-9.36
3	-50.000	-17.90		1.000000	48.85	-6.37
4	-32.100			1.000000	16.33	-1.06

Таблица 7. Паспорт оптической системы

S0	Z0	Z'0(0)	S'0(0)	F'(0)	V3P(0)	S'A
	4.256	-13.39	17.99	50.09	0.6266	17.99
MU=0	H	ДС'(0)	TGC'	Y'(0)	W(0)	ЭТА %
	20.0	-1.75	0.454	-0.793	-40.8	2.08
	17.3	-1.70	0.383	-0.650	-25.0	2.02
	14.1	-1.29	0.303	-0.389	-11.4	1.53
	10.0	-0.693	0.207	-0.143	-2.84	0.824

Таблица 8. Остаточные aberrации

S0	Z0	Z'0(0)	S'0(0)	F'(0)	V3P(0)	S'A
	-25.60	-32.10	17.99	50.091	1.000	15.358
MU=0	H	ДС'(0)	TGC'	Y'(0)	W(0)	ЭТА %
	20.0	0.887	0.454	0.403	49.9	0
	17.3	0.936	0.383	0.359	42.1	0
	14.1	1.350	0.303	0.408	32.0	0
	10.0	1.940	0.207	0.401	18.1	0

Таблица 9. Остаточные aberrации в изображении осевой точки

При $n = 1,42053$ в соответствии с формулой (16) имеем
 $\alpha = -0,01823$; $\text{ctg } \alpha = \text{tg } \gamma = -54,856$.

Подставив это значение $\text{tg } \gamma$ в уравнение (15) и положив $r = -0,01$ мм, получаем уравнение поверхности

$$x^2 + y^2 = -0,02z + 3009z^2. \quad (34)$$

Конструктивные параметры рассматриваемой системы, дополненной пластинкой с конической поверхностью, приведены в табл. 10, а остаточные aberrации в изображении осевой точки, образованном неэкранируемой частью оптической системы, полученные при $s' = 15,358$ мм, представлены в табл. 11. Схема оптической системы показана на рис. 7.

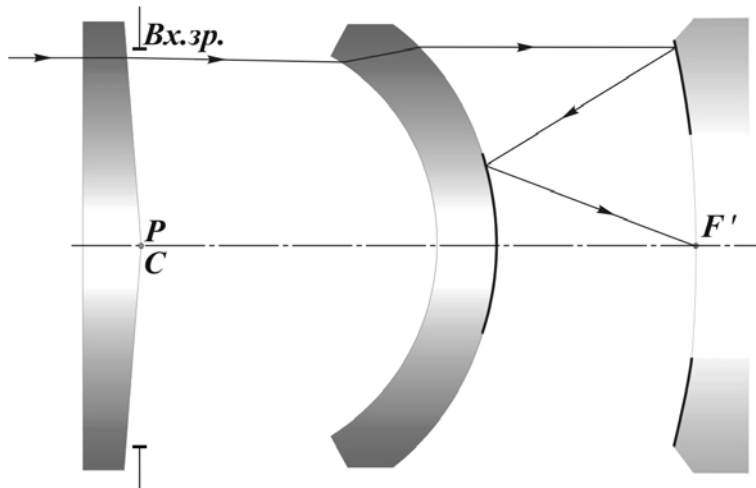


Рис. 7. Схема оптической системы по данным табл. 10

№ пов.	Радиусы кривизны	Осевые расст.	Световые высоты	Марки стекол	Показатели преломления		
1	0.000	4.000		Воздух			
2	-0.010*	25.600		CAF2	1.420526	1.416424	1.424881
3	-25.600	6.500		Воздух			
4	-32.100	17.900		CAF2	1.420526	1.416424	1.424881
5	-50.000	-17.900		Воздух			
6	-32.100			- Воздух			

* Несферическая поверхность вращения $e^2=3010.00000$

ПРЕДМЕТ: Удал. размер $Y = 1.15000$ гр.мнсек $S_0 = 0.000$ дптр от вх.зрчка

ИЗОБРАЖ: Близ. размер $Y' = 0.00000$ плоск. $SI = 15.358$ мм от посл.пов.

ДИАФР.: $ND = 2$ совп. с 2-й пов. $SD = 0.0000$ мм передн. апертура = 20.0000 мм

Таблица 10. Конструктивные параметры оптической системы

S_0	Z_0	$Z'_0(0)$	$S'_0(0)$	$F'(0)$	$V_{ЗР}(0)$	$S'A$
	-25.60	-32.10	17.99	50.091	1.000	15.358
$MU=0$	H	$DS'(0)$	TGC'	$Y'(0)$	$W(0)$	ЭТА %
	20.0	0.0862	0.464	0.0400	-8.48	0
	17.3	-0.011	0.392	-0.0043	-8.49	0
	14.1	0.144	0.311	0.0447	-9.02	0
	10.0	0.175	0.215	0.0375	-10.7	0

Таблица 11. Остаточные аберрации

Анализируя приведенные в таблице величины поперечной сферической аберрации, видим, что следовало бы принятую величину $\delta g_0'$ увеличить на 0,0202 мм. Принимая $\delta g_0' = 0,404$, получаем $\text{tg } \omega \approx \omega = -0,008069$; $\text{tg } \alpha \approx \alpha = -0,01919$; $\text{tg } \gamma = -52,11$. При этом уравнение конической поверхности (гиперboloида) принимает вид

$$x^2 + y^2 = -0,02 z + 2700 z^2.$$

Полученные при этом остаточные аберрации изображения приведены в табл. 12.

S0	Z0	Z'0(0)	S'0(0)	F'(0)	VЗР(0)	S'A
	-25.60	-32.10	17.99	50.091	1.000	15.358
MU=0	H	ДС'(0)	TGC'	Y'(0)	W(0)	ЭТА %
	20.0	0.0424	0.464	0.0197	-11.8	0
	17.3	-0.0627	0.393	-0.0246	-11.3	0
	14.1	0.0785	0.312	0.0245	-11.3	0
	10.0	0.0803	0.215	0.0173	-12.4	0

Таблица 12. Остаточные aberrации

Из сопоставления величин aberrаций, приведенных в табл. 8 и 12, следует, что применение пластинки с конической преломляющей поверхностью позволило уменьшить величину поперечной сферической aberrации не менее, чем в 30 раз. Рассматриваемая оптическая система формирует изображение вполне приемлемого качества на сферической поверхности радиуса $R = -46,5$ мм в пределах углового поля $2w = 25^\circ$. Конструктивные параметры системы приведены в табл. 13, а остаточные aberrации – в табл. 14.

№ пов.	Радиусы кривизны	Осевые расст.	Световые высоты	Марки стекол	Показатели преломления		
				Воздух			
1	0.000	4.000		CAF2	1.420526	1.416424	1.424881
2	-0.010*	25.600		Воздух			
3	-25.600	6.500		CAF2	1.420526	1.416424	1.424881
4	-32.100	17.900		Воздух			
5	-50.000	-17.900		- Воздух			
6	-32.100			Воздух			

* Несферическая поверхность вращения $e^2=2700.00000$

ПРЕДМЕТ: Удал. размер $Y = 12.30000$ гр.мнсек $S0 = 0.000$ дптр от вх.зрачка
 ИЗОБРАЖ: Близ. размер $Y' = 0.00000$ мм $RIM = -46.5000$ $SI = 15.358$ мм от посл.пов.
 ДИАФР.: $ND = 2$ совп. с 2-й пов. $SD = 0.0000$ мм передн. апертура = 20.0000 мм

Таблица 13. Конструктивные параметры оптической системы

S0	Z0	Z'0(0)	S'0(0)	F'(0)	VЗР(0)	S'A
	-25.60	-32.10	17.99	50.091	1.000	15.358
MU=0	H	ДС'(0)	TGC'	Y'(0)	W(0)	ЭТА %
	20.0	0.0424	0.464	0.0197	-11.8	0
	17.3	-0.0627	0.393	-0.0246	-11.3	0
	14.1	0.0785	0.312	0.0245	-11.3	0
	10.0	0.0803	0.215	0.0173	-12.4	0

Таблица 14. Остаточные aberrации

Высокое относительно отверстие первой отражающей поверхности, обусловленное конструкцией рассмотренной оптической системы, определяет ее сравнительно малую длину, чуть превышающую фокусное расстояние. Но именно высокая светосила первого зеркала определяет большую кривизну первого мениска, что не позволяет получить более высокое относительно отверстие, чем 1:1,25, всей системы в целом.

ПАРАМЕТРИЧЕСКИЙ СИНТЕЗ ТРЕХЗЕРКАЛЬНОЙ ОПТИЧЕСКОЙ СИСТЕМЫ НА ОСНОВЕ БАЗОВОЙ ДВУХЗЕРКАЛЬНОЙ

А.Н. Шепелевич

Научный руководитель – д.т.н., профессор В.А. Зверев

Разработанная профессором Г.Г. Слюсаревым теория расчета оптических систем на основе введенного им понятия тонкого компонента [1] определила широкое применение теории первичных аберраций не только для расчета, но и для абберационного анализа оптических систем и составляющих их элементов. Оптическая система из двух отражающих поверхностей сферической или несферической формы при нулевом расстоянии между ними практического смысла не имеет. Однако, дополнив отражающую поверхность плоской, при $d = 0$ формально получаем идеально тонкий компонент [2], обладающий абберационными свойствами первой по ходу луча отражающей поверхности. Оптическая сила i -го тонкого зеркального компонента

$$\varphi_i = (-1)^i \frac{2}{r_i}, \quad (1)$$

где r_i – радиус кривизны первой по ходу луча отражающей поверхности компонента. При таком представлении оптической системы из отражающих поверхностей для определения положения изображения относительно тонких компонентов применимы известные формулы параксиальной оптики и, в частности, например, известная формула отрезков. Оптическая сила системы из двух тонких зеркальных компонентов равна

$$\varphi = \varphi_1 + \varphi_2 - \varphi_1 \varphi_2 d_{3i}, \quad (2)$$

где $d_i = (-1)^i d_i$; d_i – расстояние между отражающими поверхностями в исходной системе. При этом $\varphi_1 = -\frac{2}{r_1}$, $\varphi_2 = \frac{2}{r_2}$; $d_{3i} = -d_1$.

Расстояние от оптической оси системы до точки пересечения осевого виртуального луча со вторым компонентом равно

$$h_2 = h_1 (1 - \varphi_1 d_{3i}).$$

Отношение $k_{ос} = \frac{h_2}{h_1}$ определяет параксиальный коэффициент экранирования осе-

вого пучка лучей (зрачка) по диаметру. При этом

$$\varphi = \varphi_1 + k_{ос} \varphi_2, \quad (3)$$

где $k_{ос} = 1 - \varphi_1 d_{3i}$. (4)

Из формулы (3) и (4) следует, что

$$\varphi_1 = \frac{1 - k_{ос}}{d_{3i}}, \quad (5)$$

$$\varphi_1 = \frac{\varphi - \varphi_2}{k_{ос}}. \quad (6)$$

Задний фокальный отрезок при $\alpha' = 1$ определяется выражением

$$s'_{F'} = h_2 = \frac{h_2}{h_1} h_1 = k_{ос} f' = \frac{k_{ос}}{\varphi}.$$

Этот отрезок удобно выразить через расстояние между компонентами, положив $s'_{F'} = k_s d_{3i}$, где k_s – коэффициент, значение которого выбирается из конструктивных соображений. При этом

$$d_{\text{э1}} = \frac{k_{\text{оэ}}}{k_s}, \quad (8)$$

$$\varphi_1 = k_s \frac{1 - k_{\text{оэ}}}{k_{\text{оэ}}}, \quad (9)$$

$$\varphi_2 = \frac{1}{k_{\text{оэ}}} \left(1 - k_s \frac{1 - k_{\text{оэ}}}{k_{\text{оэ}}} \right). \quad (10)$$

Кривизна поверхности изображения при равном нулю астигматизме (нулевая кривизна) определяется коэффициентом

$$S_{IV} = \sum_{i=1}^{i=k} \frac{\varphi_i}{n_i}.$$

В рассматриваемом случае $S_{IV} = -\varphi_1 - \varphi_2$. Подставив в это выражение соотношения (9) и (10) и преобразовав его, получаем

$$k_{\text{оэ}}^2 - \frac{2k_s + 1}{k_s - S_{IV}} k_{\text{оэ}} + \frac{k_s}{k_s - S_{IV}} = 0. \quad (11)$$

При $0 < k_s < 1$ изображение, образованное двухзеркальной системой, расположено в сходящихся пучках лучей в промежутке между зеркалами. Используя параметры оптической системы из двух тонких зеркальных компонентов, находим, что разность $d_p = d_{\text{э1}} - s'_{F'} = d_{\text{э1}} - k_s d_{\text{э1}} = (1 - k_s) d_{\text{э1}}$. При этом для осевого виртуального луча при $h_1 = 1$ имеем

$$h_p = h_1 - \alpha_2 d_p = h_1 - h_1 \varphi_1 d_p = 1 - \frac{1 - k_s}{k_s} k_{\text{оэ}} \varphi_1.$$

Используя формулу (9), получаем

$$h_p = k_{\text{оэ}} + k_s (1 - k_{\text{оэ}}). \quad (12)$$

Пусть D_p – диаметр сечения сходящегося осевого пучка лучей в плоскости расположения изображения. При этом приближенно справедливо следующее геометрическое соотношение:

$$\frac{D_p}{D} = \frac{h_p}{h} = h_p,$$

где D – диаметр входного зрачка. Отсюда следует, что $D_p = Dh_p = k_{\text{оэ}} D + (1 - k_{\text{оэ}}) k_s D$.

Диаметр изображения, образованного рассматриваемой оптической системой, равен $2y' = -2 \text{tg} W$, где W – угол поля в пространстве предметов. При этом экранирование осевого пучка лучей поверхностью изображения определим коэффициентом $k_{\text{wэ}}$, равным

$$k_{\text{wэ}} = \frac{2y'}{D_p} = -\frac{2 \text{tg} W}{k_{\text{оэ}} + k_s (1 - k_{\text{оэ}})} \frac{f'}{D} = -\frac{2F \cdot \text{tg} W}{k_{\text{оэ}} + k_s (1 - k_{\text{оэ}})}, \quad (13)$$

где F – диафрагменное число. Вполне очевидно, что должно соблюдаться условие $k_{\text{wэ}} \leq k_{\text{оэ}}$. При этом из выражения (13) следует, что

$$2 \text{tg} W \geq -\frac{k_{\text{оэ}}^2}{F} \left(1 + k_s \frac{1 - k_{\text{оэ}}}{k_{\text{оэ}}} \right). \quad (14)$$

При $k_s = 1$ $2 \text{tg} W = -\frac{k_{\text{оэ}}}{F}$. Таким образом, чем меньше диафрагменное число F (чем светосильнее система), тем угловое поле больше.

Кривую сечения несферической поверхности вращения меридиональной плоскостью в декартовой системе координат определим уравнением

$$y^2 = 2rz - (1 + b)z^2, \quad (15)$$

где r – радиус кривизны поверхности в осевой точке, b – эксцентриситет кривой второго порядка, образующей поверхность вращения.

Можно показать, что при апланатической коррекции aberrаций третьего порядка, т.е. при отсутствии в изображении сферической aberrации ($S_I = 0$) и комы третьего порядка ($S_{II} = 0$), коэффициенты деформации отражающих поверхностей соответственно равны:

$$b_1 = -1 - 2 \frac{k_s}{\varphi_1^3} = -1 - \frac{2}{k_s^2} \left(\frac{k_{o3}}{1 - k_{o3}} \right)^3, \quad (16)$$

$$b_2 = \frac{1}{k_{o3}^3 \varphi_2^3} \left[\frac{2k_s}{k_{o3}} - k_{o3} \varphi_2 (1 + \varphi_1)^2 \right] =$$

$$= \frac{1}{\left(1 - k_s \frac{1 - k_{o3}}{k_{o3}} \right)^3} \left[2 \frac{k_s}{k_{o3}} - \left(1 - k_s \frac{1 - k_{o3}}{k_{o3}} \right) \left(1 + k_s \frac{1 - k_{o3}}{k_{o3}} \right) \right] \quad (17)$$

При $S_I = 0$ и $S_{II} = 0$ коэффициент

$$S_{III} = - \frac{1 - 2(\varphi_1^2 - k_s S_{IV})}{2k_s}, \quad (18)$$

где $S_{IV} = -\varphi_1 - \varphi_2$.

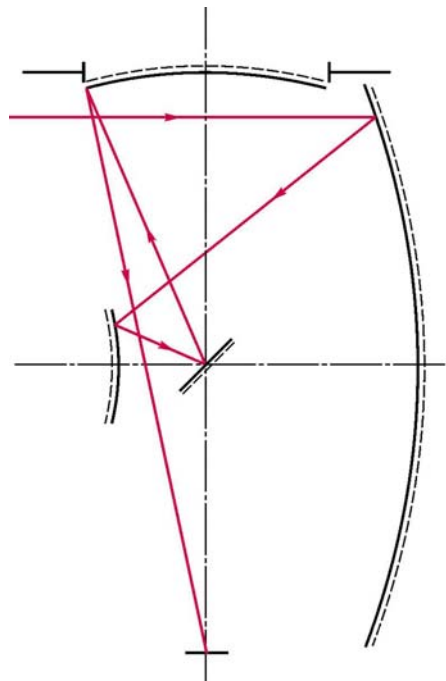


Рис. 1. Схема трехзеркальной оптической системы

В общем случае $S_{III} \neq 0$ и $S_{IV} \neq 0$. Для достижения планастигматической коррекции aberrаций изображения, при которой $S_I = S_{II} = S_{III} = S_{IV} = 0$, двухкомпонентную зеркальную систему можно дополнить третьим зеркальным компонентом $\varphi_3 = -\frac{2}{r_3}$. Рассмотрим оптическую систему, схема которой показана на рис. 1, для которой $f' : D = -F$.

При диаметре третьего компонента, равном D_3 , в соответствии с уравнением (15) имеем:

$$D_3^2 = 8r_3\Delta_3 - 4(1 + b_3)\Delta_3^2,$$

где Δ_3 – «прогиб» вогнутой поверхности третьего компонента ($\Delta_3 < 0$).

В соответствии с рисунком, на расстояние от третьего компонента до изображения, образованного двухкомпонентной системой, и на расстояние от третьего компонента до изображения, образованного третьим компонентом (и всей системой в целом), должны быть наложены следующие очевидные геометрические ограничения:

$$a_3 < \Delta_3 - \frac{1}{2}D,$$

$$a'_3 > D - \Delta_3,$$

где D – диаметр входного зрачка системы. В общем случае отрезки a_3 и a'_3 можно выразить через диаметр D соотношениями вида:

$$a_3 = -\frac{1}{2}k_1D, \quad a'_3 = \frac{1}{2}k_1D - \frac{1}{2}D + k_2D = \left(\frac{k_1-1}{2} + k_2\right)D,$$

где $k_1 > 1$, $k_2 > 1$. При этом, используя формулу отрезков, находим, что оптическая сила третьего компонента, приведенная к оптической силе всей системы ($\varphi < 0$), равна

$$\tilde{\varphi}_3 = f' \frac{a_3 - a'_3}{a_3 a'_3} = f' \frac{2}{k_1 D} \frac{2(k_1 + k_2) - 1}{k_1 + 2k_2 - 1} = -2 \frac{F}{k_1} \frac{2(k_1 + k_2) - 1}{k_1 + 2k_2 - 1}. \quad (19)$$

Отсюда следует, что при $k_1 = 1$, $k_2 \rightarrow \infty$: $\tilde{\varphi}_3 \rightarrow -2F$; при $k_1 \rightarrow \infty$, $k_2 = 1$: $\tilde{\varphi}_3 \rightarrow 0$. Положив $k_1 = k_2 = 1$, получаем $\tilde{\varphi}_3 \rightarrow -3F$.

Поперечное увеличение изображения, образованного третьим компонентом, равно

$$V_3 = \frac{a'_3}{a_3} = -1 - \frac{2k_2 - 1}{k_1}. \quad (20)$$

При $k_1 = 1$, $k_2 \rightarrow \infty$: $V_3 \rightarrow -\infty$; при $k_1 \rightarrow \infty$, $k_2 = 1$: $V_3 \rightarrow -1^\times$. При $k_1 = 2k_2 - 1$: $V_3 \rightarrow -2^\times$, при этом при $k_2 = 1$ коэффициент $k_1 = 1$.

Таким образом, в пределах разумного конструктивного решения задачи построения трехзеркальной системы можно принять, что

$$\tilde{\varphi}_3 > 3F, \quad a V_3 < -2^\times.$$

При выбранной величине V_3 фокусное расстояние двухкомпонентной системы равно

$$f'_{12} = \frac{1}{V_3} f', \quad (21)$$

а оптические силы первых двух компонентов, приведенные к оптической силе всей системы, определяются соотношениями:

$$\tilde{\varphi}_1 = \frac{f'}{f'_1} = \frac{f'_{12}}{f'_1} \frac{f'}{f'_{12}} = \varphi_1 \frac{f'}{f'_{12}} = V_3 \varphi_1;$$

аналогично находим, что $\tilde{\varphi}_2 = V_3 \varphi_2$.

Как следует из формулы (20), диафрагменное число двухкомпонентной системы в составе трехкомпонентной равно

$$F_{12} = \frac{f'_{12}}{D} = \frac{1}{V_3} \frac{f'}{D} = -\frac{1}{V_3} F.$$

Величина F_{12} принципиально не может быть меньше 0,5 при соблюдении условия синусов. Следовательно, диафрагменное число $F = -V_3 F_{12}$ не может быть меньше значения, определяемого приемлемыми величинами F_{12} и V_3 .

Кривизна поверхности изображения, образованного рассматриваемой оптической системой в целом, определяется коэффициентом

$$S_{IV\Sigma} = -\tilde{\varphi}_1 - \tilde{\varphi}_2 - \tilde{\varphi}_3 = -V_3(\varphi_1 + \varphi_2) - \tilde{\varphi}_3 = V_3 S_{IV} - \tilde{\varphi}_3,$$

где S_{IV} – коэффициент, определяющий кривизну поверхности изображения, образованного двухкомпонентной оптической системой. При $S_{IV\Sigma} = 0$ коэффициент

$$S_{IV} = \frac{1}{V_3} \tilde{\varphi}_3 = 2F \frac{2(k_1 + k_2) - 1}{(k_1 + 2k_2 - 1)^2}. \quad (22)$$

Подставив соотношение (22) в уравнение (11), при выбранном значении коэффициента k_s получаем уравнение относительно величины k_{o3} , решив которое, находим все параметры рассчитываемой системы.

Пусть, например, требуется определить конструктивные параметры трехзеркального объектива при $f' = -1000$ мм, $F = 2,5$, $2W = 2^\circ$. Диаметр входного зрачка системы $D = -\frac{f'}{F} = 400$ мм. Из конструктивных соображений принимаем $k_1 = 1,2$, $k_2 = 1,1$. При

этом из выражений (19), (20), (21) и (22) находим, что $\tilde{\varphi}_3 = -6,25$, $V_3 = -2^x$, $f'_{12} = 500$ мм, $S_{IV} = 3,125$. Приняв $k_s = 0,3$ и подставив найденное значение коэффициента S_{IV} в уравнение (11), получаем

$$k_{o3}^2 + 0,5664k_{o3} - 0,1062 = 0.$$

Параметрам рассматриваемой системы соответствует корень этого уравнения $k_{o3} = 0,1485$. Положив в соотношениях (8), (9) и (10) коэффициент $k_{o3} = 0,2$, находим, что $k_{s1} = 0,495$, $\varphi_1 = 1,7202$, $\varphi_2 = -4,8498$. При этом $d_1 = d_{s1} \cdot f'_{12} = -247,5$, $r_1 = \frac{2}{\varphi_1} f'_{12} = -581,33$ мм, $r_2 = \frac{2}{\varphi_2} f'_{12} = -206,2$ мм. Кроме того, $r_3 = -\frac{2}{\tilde{\varphi}_3} f' = -320$ мм.

Расстояние между второй и третьей отражающими поверхностями рассчитываемой оптической системы равно

$$d_2 = -k_s d_1 + \frac{1}{2} k_1 D = 314,25 \text{ мм.}$$

В результате выполненных вычислений получаем оптическую систему, которую можно записать в виде:

$$\begin{array}{lll} r_1 = -581,33 & d_1 = -247,5 & n_1 = 1 \\ r_2 = -206,2 & & n_2 = -1 \\ r_3 = -320 & d_2 = 314,25 & n_3 = 1 \\ & & n_4 = -1 \end{array}$$

В рассматриваемом случае коэффициент $k_{o3} = 0,1485 \approx 0,15$. Принято считать допустимой максимальную величину $k_{o3} = 0,3$. Поэтому выходной зрачок рассматриваемой оптической системы (апертурную диафрагму) можно считать расположенным в плоскости, проходящей через вершину третьей отражающей поверхности.

Последующая оптимизация значений коэффициентов деформации b_1 , b_2 и b_3 отражающих поверхностей по критерию минимизации значений сферической аберрации, комы и астигматизма приводит к получению системы, параметры которой приведены в табл. 1, конструктивные параметры – в табл. 2, а остаточные аберрации – в табл. 3.

Фокусное расстояние в мм	1000.0
Диафрагменное число	2.500
Угловое поле (2 омега)	2°00'20"
Положение предмета относительно первой поверхности	бесконечность
Положение изображения относительно последней поверхности	-479.98
Диаметр входного зрачка	400.000
Положение входного зрачка относительно первой поверхности	-2741.98
Положение выходного зрачка относительно последней поверхности	0
Основная длина волны в нм	546.070

Таблица 1. Параметры системы

#	РАДИУСЫ	D	МАРКИ СТЕКЛОЛ	ПОКАЗАТ. ПРЕЛОМЛ.	СВЕТОВЫЕ ДИАМЕТРЫ	СТРЕЛКИ
1	-581.330A	-247.50		1.000000	493.68	-52.64
2	-206.200A	314.25		1.000000	89.78	-4.46
3	-320.000A			-1.000000	192.31	-14.70

Таблица 2. Конструктивные параметры

S0	Z0	Z'0(0)	S'0(0)	F'(0)	VЗР(0)	S'A		
	-2742.	0	-480.0	1000.0	-.4800	-479.9		
MU=0	H	ДС'(0)	TGC'	Y'(0)	W(0)	ЭТА %		
	200.	-.0241	0.204	-.0049	.716-5	-.0424		
	173.	.00281	0.176	.494-3	-.0873	-.0436		
	141.	0.0130	0.143	.00186	.654-6	-.0375		
	100.	.00375	0.101	.377-3	0.0937	-.0232		
MU	Z	Z'	TGC'	Y'(0)	ДИС	Z'M	Z'S	Z'M-Z'S
0.0124	-2732.	0	-.0258	-12.4	.00639	-.0259	-.0279	0.00203
0.0175	-2722.	0	-.0364	-17.5	0.0177	-.0322	-.0322	-.706-6
MU	H	TGC'	ДТGC'	ДУ'	W(0)			
0.0175	200.	0.166	0.203	-.0331	0.562			
	173.	0.138	0.175	-.0108	-0.460			
	141.	0.106	0.142	.00195	-0.658			
	100.	0.0640	0.100	.00501	-0.326			
	-100.	-0.138	-0.102	0.0135	0.987			
	-141.	-0.181	-0.145	0.0215	2.32			
	-173.	-0.215	-0.179	0.0274	3.76			
	-200.	-0.244	-0.208	0.0327	5.23			
MU	M	TGC'	TGD'	ДХ'	W(0)	ДУ'		
0.0175	200.	-.0373	0.204	-.0204	1.54	.00297		
	173.	-.0371	0.176	-.0100	0.820	.00419		
	141.	-.0368	0.143	-.0042	0.435	.00419		
	100.	-.0366	0.101	-.0020	0.223	.00284		
MU	H	TGC'	ДТGC'	ДУ'	W(0)			

Таблица 3. Остаточные аберрации

Конструктивно малое плоское зеркало расположено в сходящемся пучке лучей, отраженном поверхностью третьего зеркала, а, следовательно, экранирует центральную зону пучка лучей. Заметим, что линейные размеры плоского зеркала определяются линейным полем промежуточного изображения. Следовательно, экранируемая плоским зеркалом центральная зона пучка лучей определяет угловое поле рассматриваемой оптической системы в целом.

Диаметр светового пучка лучей в плоскости, перпендикулярной оптической оси третьего компонента и содержащей оптическую ось двухкомпонентной системы, равен

$$D_p = (a'_3 + a_3) \cdot 2|tg\sigma'| = \frac{2k_2 - 1}{\sqrt{1 - \sin^2 \sigma'}} D |\sin \sigma'|. \text{ Но } |\sin \sigma'| = \frac{1}{2} \frac{D}{|f'|} = \frac{1}{2F}. \text{ При этом}$$

$$D_p = \frac{2k_0 - 1}{\sqrt{4F^2 - 1}} D. \quad (23)$$

Полагая $2y' \approx k_{o3} D_p$, где $2y'$ – диаметр изображения, образованного двухкомпонентной системой, получаем

$$2|\operatorname{tg}W| = \frac{2y'}{f'_{12}} = \left| \frac{v_3}{f'} \right| k_{o3} D_p = \left| \frac{v_3}{f'} \right| \frac{2k_2 - 1}{\sqrt{4F^2 - 1}} k_{o3} D$$

или $2|\operatorname{tg}W| \leq \left| \frac{v_3}{F\sqrt{4F^2 - 1}} \right| k_{o3}$. (24)

Отсюда следует, что чем больше величина коэффициента k_2 и чем меньше величина F , тем больше угловое поле системы. Однако следует иметь в виду, что чем больше k_2 , тем больше абсолютная величина v_3 , а, следовательно, тем больше светосила двухкомпонентной системы.

Заметим, что при $F = 2,5$, $V_3 = -2^\times$ и $k_2 = 1,1$ величина $2|\operatorname{tg}W| = 0,196k_{o3}$. При $k_{o3} = 0,3$ угловое поле $2W = 3^\circ 22'$.

Литература

1. Слюсарев Г.Г. Методы расчета оптических систем. Л.: Машиностроение, 1969, 672 с.
2. Зверев В.А., Шепелевич А.Н. Понятие тонкого компонента в системе отражающих поверхностей. // Оптический журнал. 2006. Т.73. №12. С. 21–26.

РАЗРАБОТКА ПРОГРАММНОГО КОМПЛЕКСА ДЛЯ НОВОГО ПОКОЛЕНИЯ ОФТАЛЬМОЛОГИЧЕСКИХ ПРИБОРОВ. ПРИБОР ДЛЯ ИССЛЕДОВАНИЯ РЕАКЦИИ ГЛАЗА НА ВНЕШНИЕ ИНТЕНСИВНЫЕ ЗАСВЕТКИ

К.В. Каряев

Научный руководитель – д.т.н., профессор Э.С. Путилин

Рассматриваются требования, предъявляемые к программным комплексам учета в клинике. Предложена оригинальная платформа построения программного комплекса, позволяющая интегрировать разнородные диагностические данные и учет историй болезни пациентов в рамках единой системы. Рассмотрено применение модели для реализации модуля для прибора для исследования реакции глаза на внешние интенсивные засветки.

Введение

Современное диагностическое оборудование с точки зрения организации пользовательского интерфейса, как правило, можно отнести к одной из двух категорий. Первую составляют приборы для экспресс-диагностики, обладающие собственными дисплеями, упрощенной встроенной информационной системой и системой обработки и анализа данных. Вторую составляют высокоточные комплексы для клиник и диагностических центров, а их информационные системы и системы обработки данных, как правило, используют персональные компьютеры. Обмен данными производится по последовательному порту RS-232, а в последнее время также с использованием шины USB или FireWire.

С другой стороны, в настоящий момент внедрение автоматизированных систем учета историй болезни пациентов в клиниках становится все более распространенным и в России. Однако, как правило, эти системы ориентируются на автоматизацию платежного документооборота и подчас не имеют даже электронных баз историй болезни пациентов. Таким образом, количество работы персонала не сокращается, а возрастает, а дополнительные возможности, предоставляемых вычислительными системами и системами баз данных, не используются.

Также стоит отметить превалирование индивидуальных заказов и индивидуальных программных решений в данной области. С учетом необходимости массового переоснащения государственных клиник и перехода на повсеместное использование информационных технологий целесообразна разработка единого стандарта и единой программной платформы. Возможна и разработка единого масштабируемого и расширяемого программного комплекса на основе этой платформы.

В данной работе обсуждается один из возможных подходов к построению общей архитектуры информационной системы клиники на примере информационной системы (ИС) офтальмологического центра. Предлагаемая платформа ориентирована на регистрацию и обработку данных клинических исследований. Особое внимание уделяется разработке платформы для взаимодействия с диагностическими приборами.

Информационная платформа

Предлагаемая платформа (рис. 1) состоит из нескольких основных структурных элементов, обладающих определенными особенностями:

- a) база данных (БД), хранящая как текстовые, так и графические данные;
- b) пользовательский интерфейс различных автоматизированных рабочих мест (АРМ), типичных для клиники;
- c) модульная структура АРМ врача-специалиста:

- a. модули взаимодействия с приборами;
- b. различные диалоговые окна для различных задач (интерфейс к диагностическим приборам, выборка данных из истории болезни, внесение диагноза и других текстовых данных в историю болезни);
- c. расширяемая архитектура комплекса, позволяющая устанавливать модули для взаимодействия с новыми приборами и реализации новых методов анализа клинических данных.

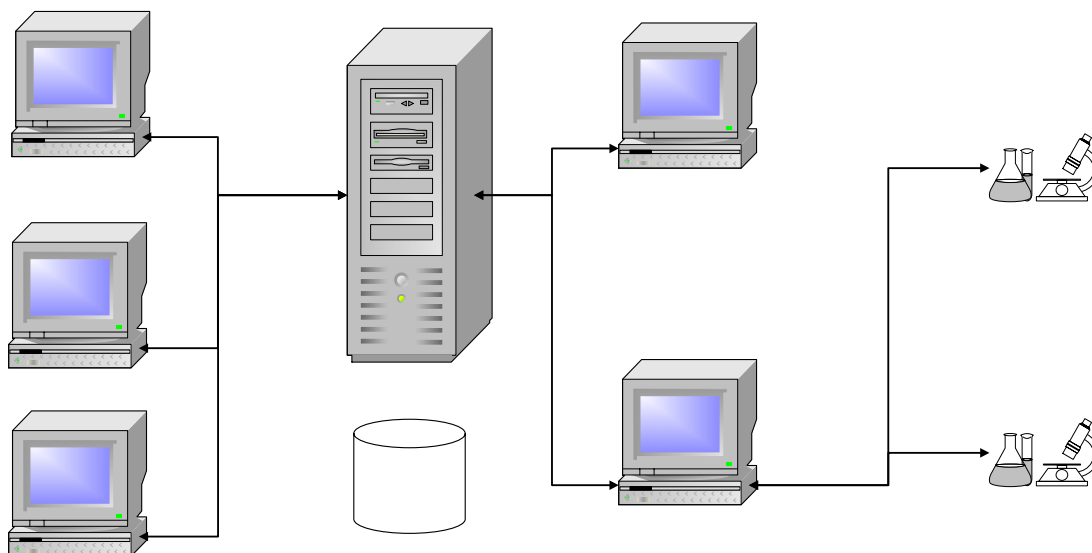


Рис. 1. Общая схема информационной сети клиники

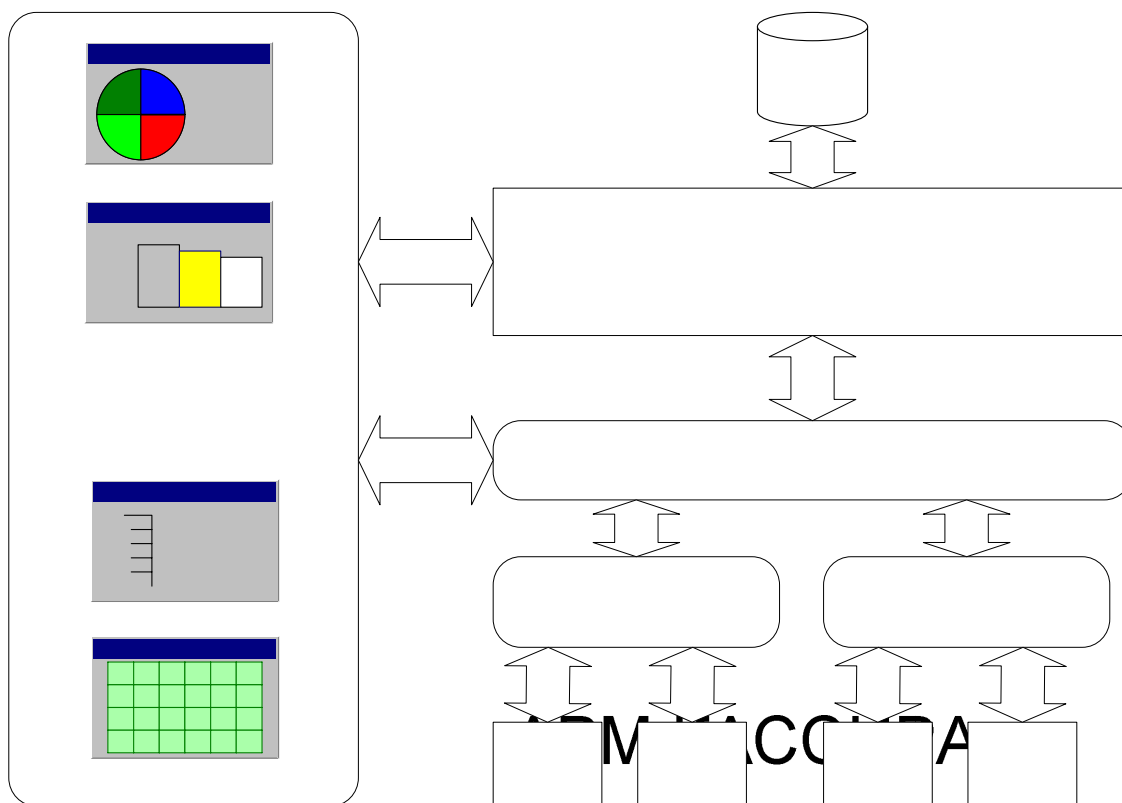


Рис. 2. Схема программно-аппаратного комплекса АРМ врача-специалиста

Современные системы автоматизации учета начального уровня в клиниках акцентируют внимание на АРМ регистратуры и АРМ кассира. Задачи автоматизации учета

новых пациентов и оформление оплаты услуг являются хорошо разработанными и на данном этапе не представляют серьезной сложности. Тем не менее, необходимо отметить вопрос ограничения доступа на АРМ к чувствительным данным о пациенте. Таким образом, каждое АРМ должно обеспечивать доступ только к той информации, которая необходима, обеспечивая наиболее удобный интерфейс, позволяющий с наименьшими затратами времени производить обслуживание пациентов.

Мы предлагаем обратить особое внимание на АРМ врача-специалиста (рис. 2). Для этого АРМ характерны специфические задачи, а именно:

- доступ к истории болезни пациента (в том числе поиск отдельных данных по категориям и ключевым словам), включая анамнез, данные предыдущих лабораторных исследований, диагнозы и т.п.;
- возможность вносить данные исследований;
- возможность вносить данные диагноза.

На практике, однако, возникает проблема, связанная с неприспособленностью большинства существующих приборов к автоматизированному учету результатов клинических исследований. Это приводит к необходимости, наряду с электронными историями болезни, и бумажных карточек, включающих распечатки результатов клинических исследований. Это затрудняет и восстановление истории течения болезни.

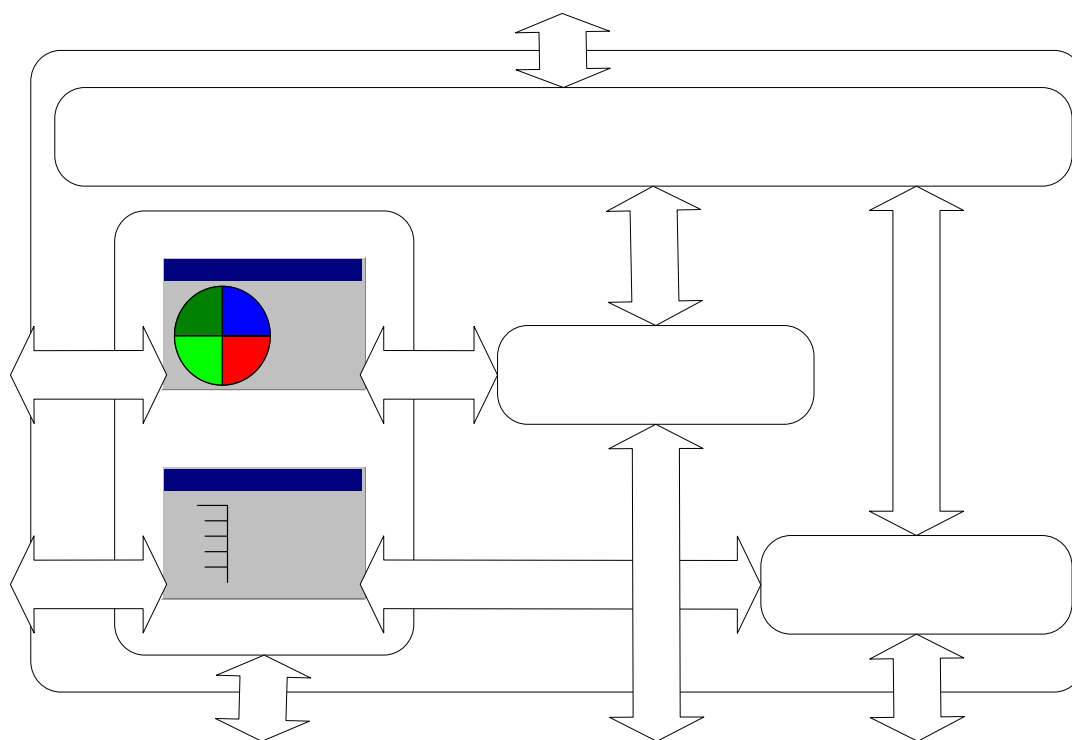


Рис. 3. Схема модуля ИС для взаимодействия с отдельным прибором

Таким образом, внедрение единого стандарта передачи данных для интерфейса приборов с вычислительной системой является целесообразным решением, способствующим дальнейшему развитию информационных систем клиник. С другой стороны, многие современные приборы укомплектованы разъемом для подключения к компьютеру. Это позволяет рассчитывать на возможность разработки отдельных модулей для информационной системы, которые будут объединять функциональность, уникальную для конкретного прибора, или группы приборов (например, последовательных модификаций одного прибора). Единственным условием является открытость протокола обмена данными с прибором и алгоритмов ее обработки.

Структура модуля информационной платформы

Целесообразно выделить в модуле (рис. 3) логические блоки, отвечающие за низкоуровневое взаимодействие с прибором, за специфические компоненты пользовательского интерфейса и за специальные процедуры анализа и обработки входных данных.

Важнейшим фактором является взаимодействие различных компонент модуля между собой и с ядром АРМ специалиста. Выбранная нами схема представлена на рис. 3. Она позволяет разделить низкоуровневые алгоритмы обмена данными с прибором и пользовательский интерфейс, а также производить анализ, обработку и сохранение полученных данных в БД в фоновом для пользователя режиме.

Реализация модуля прибора для исследования реакции глаза на внешние интенсивные засветки

Как известно, реакцией глаза на посторонние засветки, в том числе импульсные, является изменение диаметра зрачка и биохимических процессов, проходящих в сетчатке. Скорость реакции глаза на меняющиеся условия освещенности, особенно при перепаде освещенности на несколько порядков, практически не исследованы.

Исследование динамики изменения диаметров зрачков глаз в реальном времени с высоким пространственным и временным разрешением позволит определить профессиональную пригодность лиц, допущенных к управлению транспортными средствами. Выделим основные требования к установке, позволяющей исследовать реакцию глаза на внешние интенсивные засветки:

- установка должна отвечать требованиям эргономики и безопасности для здоровья испытуемого;
- установка должна обеспечивать имитацию широкого диапазона условий освещения, встречающихся в повседневной жизни человека, включая воздействие импульсных источников в условиях дневного, ночного и сумеречного зрения, в том числе в нестандартных (экстремальных) ситуациях;
- установка должна обеспечивать регистрацию и измерение диаметра зрачка глаза исследуемого пациента в реальном времени, а также сохранение и интерпретацию данных измерений.

Целесообразно оставить в базовой конструкции прибора только регистрирующие функции, а функции анализа и расчета передать программному комплексу. В таком случае входными данными программного модуля будет последовательность кадров (десятки или сотни, со скоростью 20–40 кадров в секунду), считываемых с ПЗС-матрицы. Целями обработки будут определение диаметра зрачка глаза и построение функции его изменения со временем, соотношенной с функцией изменения во времени интенсивности освещения.

Целесообразным будет отображение как графической, так и табличной информации. Опционально возможно и отображение видеопотока, полученного с ПЗС-матрицы. В БД необходимо сохранить табличные данные, дополнительно возможно сохранение видеопотока.

Пользователь (врач-специалист) должен иметь возможность выбрать или задать функцию изменения освещения, а также запустить процедуру исследования в момент готовности пациента. Необходимо предусмотреть и функцию отображения предыдущих исследований, запрашиваемых из базы данных.

На рис. 4 приведена итоговая схема данного модуля.

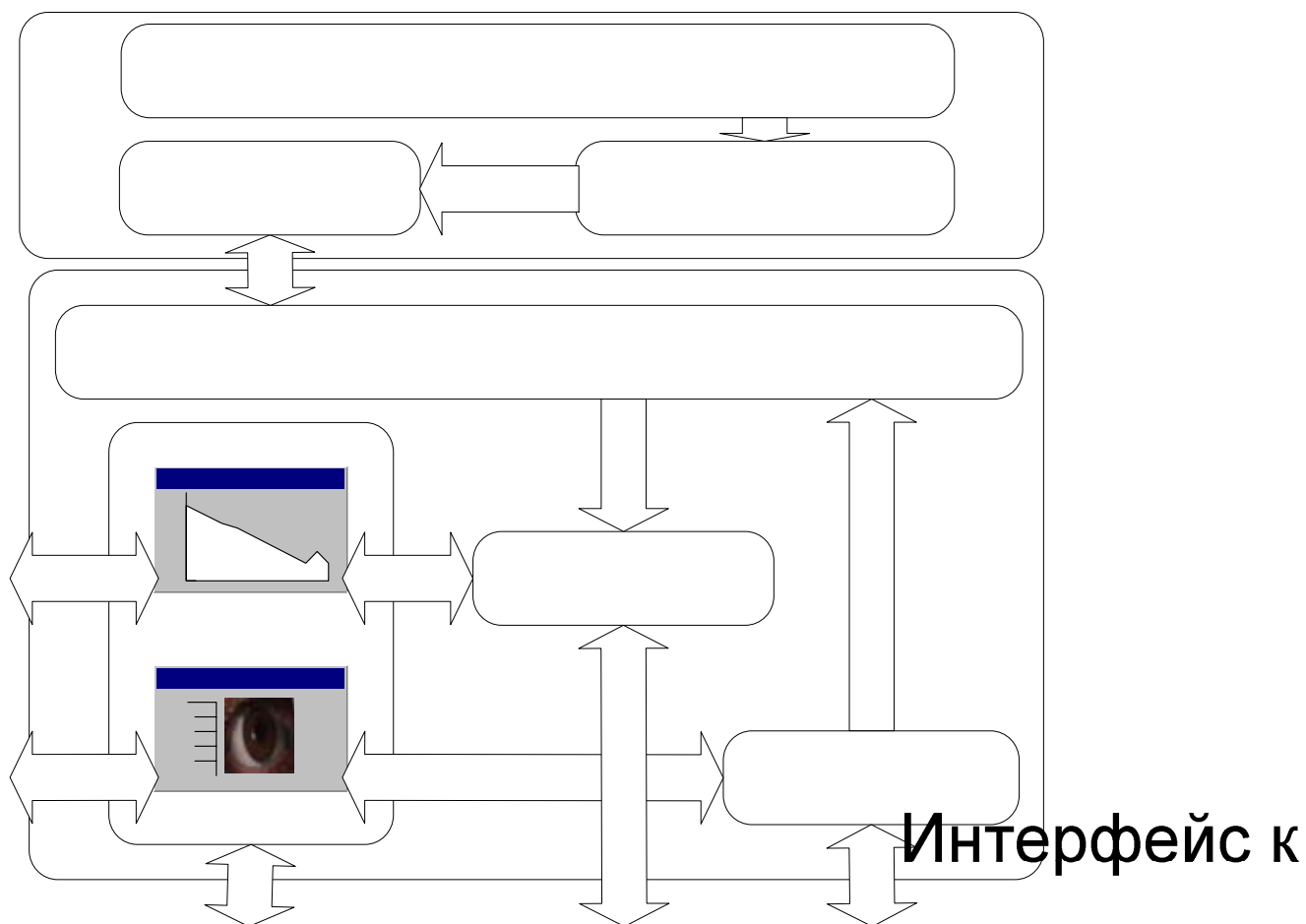


Рис. 4. Схема модуля ИС для прибора для исследования реакции глаза на внешние интенсивные засветки

Заключение

Разработка современных программных комплексов для интегрирования всего документооборота клиники является важной областью, критической для дальнейшего совершенствования оказания медицинской помощи и повышения качества и эффективности работы специалистов-врачей. В результате анализа выявлены технические требования и разработана платформа программного комплекса для интегрирования системы учета с диагностическими приборами. На примере прибора для исследования реакции глаза на внешние интенсивные засветки показаны основные детали данного комплекса и подтверждена возможность реализации предложенной модели на практике. В ходе дальнейшего развития работы возможна разработка универсального комплекса, позволяющего обеспечить интеграцию различных диагностических данных в рамках единой базы данных, и тем самым повысить эффективность работы специалистов клиники.

Литература

1. Урмахер Л.С., Айзенштат Л.И. Офтальмологические приборы. М.: Медицина, 1988.
2. Москалев В.А., Нагибина И.М., Полушкина Н.А., Рудин В.Л. Прикладная физическая оптика. СПб: Политехника, 1995.
3. Закер К. Компьютерные сети: Пер. с англ. СПб: БХВ-Петербург, 2002.
4. Кренке Д. Теория и практика построения баз данных. 8 изд. СПб: Питер, 2003.

УСТАНОВКА ДЛЯ ИССЛЕДОВАНИЯ РЕАКЦИИ ГЛАЗА НА ВНЕШНИЕ ИНТЕНСИВНЫЕ ЗАСВЕТКИ

Г.В. Макаричев

Научный руководитель – д.т.н., профессор Э.С.Путилин

Рассматриваются технические требования, предъявляемые к установке для исследования реакции глаза человека на меняющиеся условия освещения, включая внешние интенсивные засветки. Предложена и реализована оригинальная схема установки для исследования скорости изменения диаметра зрачка глаза человека при изменении условий внешнего освещения с различной скоростью.

Введение

Оценка диаметра зрачка и поведения глаза в условиях различных уровней освещенности является важной офтальмо-диагностической процедурой. Кроме определения профессиональной пригодности пациента, оценка диаметра зрачка при различных внешних засветках является важной частью предоперационного обследования при рефракционной хирургии. Некорректное определение диаметра зрачка при рефракционной хирургии может привести к потере трудоспособности пациента.

В настоящее время ведущие зарубежные компании-производители офтальмо-диагностического оборудования выпускают приборы близкого назначения (Procuon, NeurOptics, Oasis Medical, Torcon и др.), однако все эти приборы обладают недостатками, ухудшающими достоверность получаемых результатов, и не позволяют проводить комплексное обследование пациентов для определения их профессиональной пригодности. Кроме определения профессиональной пригодности пациента, оценка диаметра зрачка является важной частью предоперационного обследования при рефракционной хирургии. Некорректное определение диаметра зрачка при рефракционной хирургии может привести к потере трудоспособности пациента. В литературе описан случай возмещения страховыми компаниями иска на 5 млн. долларов, поданного пациентом (коммерческим пилотом), утратившим трудоспособность в результате рефракционной операции по некорректным данным измерения диаметра зрачка.

Отечественная промышленность не выпускает устройства для определения реакции глаза на внешние интенсивные засветки, предназначенные для функциональной диагностики состояния органа зрения лиц, управляющих транспортными средствами различного назначения, и работающие в условиях сильно меняющегося освещения.

Таким образом, возникает задача создания установки, позволяющей исследовать реакцию глаза на внешние интенсивные засветки.

Теоретические предпосылки разработки установки для исследования реакции глаза на внешние интенсивные засветки.

Как известно, реакцией глаза на посторонние засветки, в том числе импульсные, является измерение диаметра зрачка и биохимических процессов, проходящих в сетчатке. Скорость реакции глаза на меняющиеся условия освещенности, особенно при перепаде освещенности на несколько порядков, практически не исследованы.

Исследование динамики изменения диаметров зрачков глаз в реальном времени с высоким пространственным и временным разрешением позволит определить профессиональную пригодность лиц, допущенных к управлению транспортными средствами. В зависимости от уровней яркости, при которых работает глаз, функционируют или колбочки, или палочки сетчатки. В соответствии с этим различают дневное, ночное и сумеречное зрение. Дневным зрением называют зрение нормального глаза при адаптации его к уровням яркости большим, чем $2-3 \text{ кд/м}^2$. Считается, что в этих условиях начинают действовать в основном колбочки сетчатки, и предметы представляются окра-

шенными. Ночное зрение – это зрение нормального глаза при адаптации его к уровням яркости, меньшим, чем $0,02-0,03$ кд/м². Считается, что в этих условиях начинают действовать, главным образом, палочки сетчатки, и предметы представляются неокрашенными. Максимальная относительная спектральная световая эффективность в этом случае соответствует более короткой длине волны, чем при дневном зрении.

Зрение, которое является промежуточным между дневным и ночным зрением, называется сумеречным. В приведенных определениях введен термин адаптация. Адаптация – это процесс изменения свойств глаза под воздействием яркостных и цветовых стимулов. Кроме того, под адаптацией понимают конечное состояние этого процесса. В частности, говорят о световой или темновой адаптации в зависимости от того, превышает ли яркость несколько кандел на квадратный метр или не достигает нескольких сотых канделы на квадратный метр.

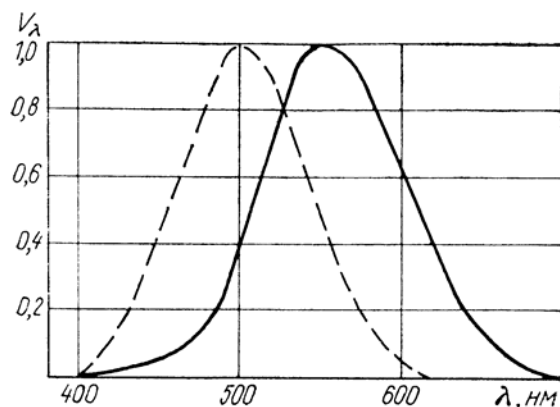


Рис. 1. График функции относительной спектральной световой эффективности глаза человека для дневного (-) и ночного (---) зрения

Для адаптации к низкому уровню яркости требуется больше времени, чем для адаптации к высокому уровню яркости. По мере пребывания в темноте чувствительность все время возрастает. Для достижения полной чувствительности глаза необходим примерно часовой–полтора часовой отдых от света. При переходе от дневного зрения к ночному зрению меняется спектральная чувствительность глаза, которая определяется нормализованной функцией относительной спектральной световой эффективности излучения, показанной на рис. 1.

Глаз чувствителен к излучению в диапазоне длин волн с $\lambda=380-780$ нм. Максимум функции относительной спектральной световой эффективности $V(\lambda)$ при дневном зрении находится вблизи $\lambda_{\max}=555$ нм, а при ночном – смещается к $\lambda_{\max} = 507$ нм. Это явление носит название эффекта Пуркинье; оно состоит в уменьшении светлоты красного света по сравнению со светлотой синего света, когда яркости уменьшены в одинаковой пропорции без изменения спектрального состава.

Пороговая чувствительность глаза соответствует работе палочкового аппарата зрения. В состоянии полной темновой адаптации глаз крайне чувствителен и способен реагировать на потоки энергии.

Одной из важнейших характеристик при исследованиях является спектральный состав зондирующего источника освещения. Существующие приборы с близким назначением осуществляют исследования только в широком спектральном диапазоне (белый свет), что не позволяет полностью адекватно определить профессиональную пригодность лиц в реальных условиях и ситуациях. В ряде случаев от точности исследования, позволяющего определить профессиональную пригодность лиц, допущенных к управлением транспорта, зависит здоровье и жизнь большого количества людей. Дальнейший технический прогресс в области комфортности и безопасности управления транспортными средствами невозможен без подобных устройств.

Основными потребителями и предполагаемыми применениями разрабатываемых установок могут быть:

- офтальмологические клиники – для исследований при подготовке к операциям по рефракционной хирургии;
- врачебные комиссии (авто-, аэро-, судо-водительские, специальные) – для определение профессиональной пригодности кандидатов;
- судебно-медицинская и врачебная экспертиза – для анализа причин и обследования участников чрезвычайных ситуаций, аварий и катастроф;
- поликлиники и семейные врачи – комплексные экспресс-исследования органа зрения пациентов с целью назначения поддерживающей терапии;
- фирмы-изготовители транспортных средств различного назначения – оценка корректности функционирования осветительных систем разрабатываемых/производимых транспортных средств.

Рассмотрим требования к установке, позволяющей исследовать реакцию глаза на внешние интенсивные засветки:

- отвечать требованиям эргономики и безопасности для здоровья испытуемого;
- обеспечивать имитацию широкого диапазона условий освещения, встречающихся в повседневной жизни человека, включая воздействие импульсных источников в условиях дневного, ночного и сумеречного зрения, в том числе в нештатных (экстремальных) ситуациях;
- обеспечивать регистрацию и измерение диаметра зрачка глаза исследуемого пациента в реальном времени, а также сохранение и интерпретацию данных измерений.

Конструкция установки для исследования реакции глаза на внешние интенсивные засветки

Исходя из сформированных требований, была предложена и реализована оригинальная схема установки для исследования реакции глаза на внешние интенсивные засветки. Подход к разработке установки направлен на увеличение достоверности получаемой информации о состоянии органа зрения, его реакции на сильно меняющиеся условия освещения, особенно при переходе от ночного к сумеречному и дневному зрению и обратно. В сочетании с комплексным подходом к импульсным и стабильным во времени источникам излучения с меняющейся во времени по заданному либо случайному закону интенсивностью излучения, имитирующего засветки от встречного транспорта или случайных источников света, и видеосистемами контроля диаметра зрачков глаз, а также программным обеспечением для управления параметрами засветки и анализа результатов измерения положения и диаметров зрачков глаз, этот подход позволяет рассчитывать на высокую адекватность получаемых результатов.

Установка состоит из следующих систем:

- система освещения, представляющая собой импульсные и стабильные во времени источники излучения с меняющейся во времени по заданному, либо случайному закону интенсивностью излучения, которое имитирует засветки от встречного транспорта или случайных источников света;
- система регистрации, включающая ИК-светодиоды подсветки для измерения диаметра зрачка, объектив, формирующий изображение зрачка пациента на черно-белой ПЗС-матрице. Также в системе регистрации присутствует автоматизированный модуль смены светофильтров;
- система управления и обработки данных, состоящая из персонального компьютера со специализированным программным обеспечением для управления источниками освещения и блоком светофильтра, а также для записи видеопоследовательности с ПЗС-матрицы и ее анализа в реальном времени.

Блок-схема взаимосвязи основных деталей и узлов установки приведена на рис. 2.

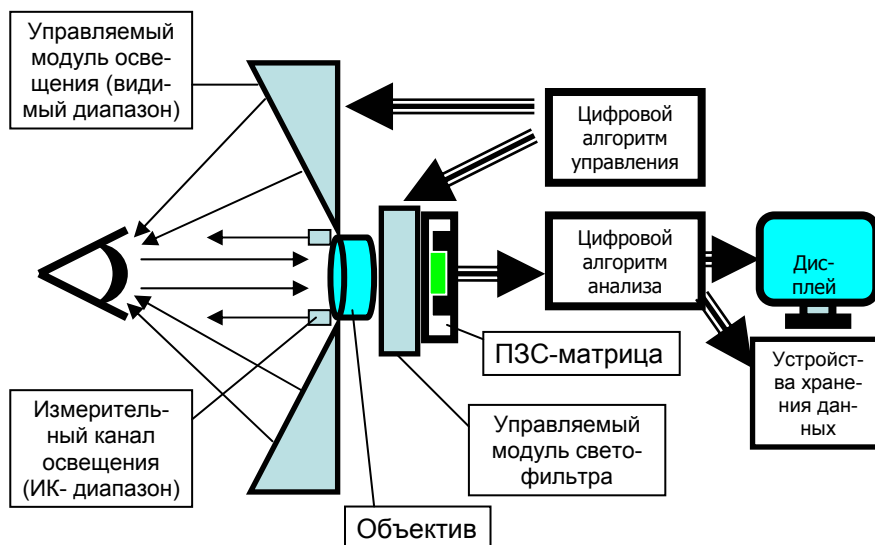


Рис. 2. Блок-схема установки

В настоящее время авторами разрабатывается конструкторская документация для изготовления данной установки, проводятся работы по созданию соответствующего программного обеспечения. Ряд дополнительных идей в настоящее время проходит стадию проработки. Технические решения, положенные в основу предлагаемого проекта, подготовлены в виде заявок на полезные модели и изобретения, подаваемые в Российское патентное ведомство.

Заключение

Оценка диаметра зрачка и поведения глаза в условиях различных уровней освещенности является важной офтальмо-диагностической процедурой, критической для определения профессиональной пригодности пациента, а также для рефракционной хирургии. В результате работы выявлены технические требования и разработана оригинальная схема установки для определения реакции глаза на внешние интенсивные зачатки. В случае практической реализации, возможно появление уникального прибора, позволяющего проводить комплексное обследование пациентов для определения их профессиональной пригодности.

Литература

1. Сомов Е.Е. Клиническая анатомия органа зрения человека. Учебное пособие. СПб: ПМИ, 1992.
2. Глазные болезни. Под ред. Т.И. Ерошевского, А.А. Бочкаревой. 2-е изд., перераб. и доп. М.: Медицина, 1983.
3. Грегори Р.Л. Глаз и мозг. Психология зрительного восприятия. М.: Прогресс, 1970.
4. Урмахер Л.С., Айзенштат Л.И. Офтальмологические приборы. М.: Медицина, 1988.
5. Москалев В.А., Нагибина И.М., Полушкина Н.А., Рудин В.Л. Прикладная физическая оптика. СПб: Политехника, 1995.
6. Джадд Д., Вышецки Г. Цвет в науке и технике. М., Мир, 1978.
7. Физиология зрения в нормальных и экстремальных условиях. Проблемы физиологической оптики. Т. 15. Л.: Наука, 1969.

ИССЛЕДОВАНИЕ ВЛИЯНИЯ ЗАЗОРА В БАШМАЧНОЙ ПРИЗМЕ НА КАЧЕСТВО ИЗОБРАЖЕНИЯ

К.Т. Чан

Научный руководитель – д.т.н., профессор А.П. Грамматин

Зазор между двумя частями башмачной призмы представляет собой наклоненную к оптической оси прибора плоскопараллельную воздушную пластину, расположенную в стеклянной среде, вызывающую появление аберраций децентрировки. Выполнено численное исследование комы, астигматизма в зависимости от величины зазора и числовой апертуры в пространстве, где расположена призма.

Введение

Зазор между двумя частями башмачной призмы имеет разные значения в различных источниках [1, 2]. При разных значениях этого зазора аберрации, вносимые башмачной призмой, имеют различные значения, в том числе появляются аберрации, превышающими допустимые значения по критериям Марешаля. Поэтому необходимо исследование влияния зазора на качество изображения в башмачной призме. На основе этого исследования конструктор сможет выбрать рациональные величины зазора для своей задачи с целью достижения хорошего качества изображения и облегчения в процессе изготовления. В этой работе путем непосредственного расчета хода лучей выполнено численное исследование комы, астигматизма в зависимости от величины зазора и числовой апертуры в пространстве, где расположена призма.

1. Численный расчет аберраций, вносимых башмачной призмой с различными величинами зазора, которые представлены в справочниках

В качестве примера численного исследования влияния зазора на качества оптического изображения расчет аберраций, вносимых башмачной призмой с различными величинами зазора, выполнен в следующих условиях.

- Башмачная призма располагается в пространстве изображения объектива с числовой апертурой в этом пространстве $A' = 0.1$.
- Призма выполнена из стекла К8, Ф2 или ТФ1.

В справочнике [1] представлена башмачная призма с зазором $d = 0.1$ мм. При этом зазоре были рассчитаны аберрации децентрировки, вносимой наклоненной к оптической оси прибора плоскопараллельной воздушной пластиной. Аберрации показаны в табл. 1–3.

Призма из стекла К8 (табл. 1)

Кома в волновой мере: $W_k = 1.17$

Астигматизм в волновой мере: $W_{act} = 2.87$

		Н ПУЧКА= 0									
		NLU	MU	LA	DTGS'	DTGD'	DY'	DX'	W(0)	W(0)	
0	Y	0	0	0	0	0	0.101	0	0.101	0	
		1	0.100	0	-0.101	0	0.0171	0	0.0171	-1.17	
0	X	3	0.0707	0	-0.0709	0	.00604	0	.00604	-0.280	
		6	-0.0707	0	0.0709	0	.00621	0	.00621	0.489	
0		8	-0.100	0	0.101	0	0.0106	0	0.0106	1.17	
		9	0	0.100	0	-0.101	.00161	-0.0206	0.0206	2.87	
		11	0	0.0707	0	-0.0709	.791-3	-0.0146	0.0146	1.44	
		41	0.0707	0.0707	-0.0711	-0.0711	.00770	-0.0123	0.0145	0.935	
		43	0.0500	0.0500	-0.0501	-0.0501	.00277	-0.0093	0.00975	0.595	
		49	-0.0707	0.0707	0.0711	-0.0711	.00671	-0.0159	0.0172	2.05	
		51	-0.0500	0.0500	0.0501	-0.0501	.00394	-0.0110	0.0117	0.973	

Таблица 1. Призма из стекла К8

Призма из стекла Ф2 (табл. 2)

Кома в волновой мере: $W_k = 3.25$
 Астигматизм в волновой мере: $W_{act} = 6.35$

		Н ПУЧКА= 0									
		NLU	MU	LA	DTGC'	DTGD'	DY'	DX'	J***	W(0)	
0	Y	0	0	0	0	0	0.120	0	0.120		
		1	0.100	0	-0.101	0	0.0556	0	0.0556	-3.25	
		3	0.0707	0	-0.0709	0	0.0145	0	0.0145	-0.534	
	X	6	-0.0707	0	0.0709	0	0.0161	0	0.0161	1.43	
		8	-0.100	0	0.101	0	0.0261	0	0.0261	3.25	
		9	0	0.100	0	-0.101	0.0297	-0.0427	0.0428	6.35	
		11	0	0.0707	0	-0.0709	0.0145	-0.0303	0.0303	3.18	
		41	0.0707	0.0707	-0.0711	-0.0711	0.0190	-0.0251	0.0315	2.12	
		43	0.0500	0.0500	-0.0501	-0.0501	0.0518	-0.0194	0.0201	1.43	
		49	-0.0707	0.0707	0.0711	-0.0711	0.0169	-0.0325	0.0366	4.83	
		51	-0.0500	0.0500	0.0501	-0.0501	0.0104	-0.0226	0.0249	2.31	

Таблица 2. Призма из стекла Ф2

Призма из стекла ТФ1 (табл. 3)

Кома в волновой мере: $W_k = 4.82$
 Астигматизм в волновой мере: $W_{act} = 8.58$

		Н ПУЧКА= 0									
		NLU	MU	LA	DTGC'	DTGD'	DY'	DX'	J***	W(0)	
0	Y	0	0	0	0	0	0.127	0	0.127		
		1	0.100	0	-0.101	0	0.0941	0	0.0941	-4.82	
		3	0.0707	0	-0.0709	0	0.0196	0	0.0196	-0.569	
	X	6	-0.0707	0	0.0709	0	0.0233	0	0.0233	2.16	
		8	-0.100	0	0.101	0	0.0371	0	0.0371	4.82	
		9	0	0.100	0	-0.101	0.0364	-0.0566	0.0567	8.58	
		11	0	0.0707	0	-0.0709	0.0177	-0.0401	0.0402	4.30	
		41	0.0707	0.0707	-0.0711	-0.0711	0.0263	-0.0331	0.0423	3.01	
		43	0.0500	0.0500	-0.0501	-0.0501	0.0577	-0.0258	0.0264	2.04	
		49	-0.0707	0.0707	0.0711	-0.0711	0.0242	-0.0427	0.0491	6.73	
		51	-0.0500	0.0500	0.0501	-0.0501	0.0152	-0.0298	0.0335	3.23	

Таблица 3. Призма из стекла ТФ1

Из табл. 1–3 видно, что эти aberrации превышают допустимые значения по критериям Марешаля для малых aberrаций. По Марешалю [3], допустимая кома в волновой мере равна 0.61, а допустимый астигматизм равен 0.17. Поэтому при этих условиях этот зазор должен быть уменьшен.

В справочнике [2] представлена башмачная призма с зазором $d = 0.002$ мм. Тогда aberrация, вызванная этим зазором, имеет следующие значения.

- Призма из стекла К8
 Кома в волновой мере: $W_k = 0.0233$
 Астигматизм в волновой мере: $W_{act} = 0.0574$
- Призма из стекла Ф2
 Кома в волновой мере: $W_k = 0.0649$
 Астигматизм в волновой мере: $W_{act} = 0.127$
- Призма из стекла ТФ1
 Кома в волновой мере: $W_k = 0.0964$
 Астигматизм в волновой мере: $W_{act} = 0.17$

Очевидно, что при этом зазоре (0.002 мм) башмачная призма создает идеальное изображение.

2. Исследование отношения между aberrациями и величиной зазора, числовой апертурой в пространстве, где расположена призма

Зависимость комы и астигматизма от величины зазора. Выполнив расчет aberrаций с различными величинами зазора, можно получить зависимость комы и астигматизма от зазора. Эта зависимость – линейная:

$$W_k = M_1 \cdot d,$$

$$W_{act} = M_2 \cdot d.$$

Здесь M_1 и M_2 – неизвестные величины, которые зависят от числовой апертуры A' в пространстве, где расположена призма, и от показателей преломления материала призмы.

Зависимости комы и астигматизма от числовой апертуры – следующие:

$$W_k = M_1' \cdot d \cdot (A')^{3.45},$$

$$W_{act} = M_2' \cdot d \cdot (A')^2,$$

где M_1' и M_2' – коэффициенты, которые зависят от показателей преломления материалов.

Числовое исследование зависимости комы и астигматизма от показателей преломления материалов, из которых выполнена призма. Это отношение исследовано путем расчета aberrаций, вносимых призмой, которая выполнена из некоторых различных стекол. Была рассчитана aberrация для стекол: К8, Ф2 и ТФ1 для случая, где числовая апертура $A' = 0.1$ и зазор с величиной $d = 0.1$ мм. Показатели преломления стекол и соответствующие aberrации представлены в табл. 4.

Материалы	Показатели преломления при $\lambda_0 = 0.5461$ нм	Кома в волновой мере	Астигматизм в волновой мере
К8	1.518	1.17	2.87
Ф2	1.621	3.25	6.35
ТФ1	1.652	4.82	8.58

Таблица 4. Показатели преломления стекол и соответствующие aberrации

Из числовых значений в табл. 4 можно сделать вывод о зависимости aberrаций от показателей преломления материалов, что кома и астигматизм прямо пропорциональны показателям преломления.

Заключение

Из исследования следует, что зазор между двумя частями башмачной призмы не должен обязательно быть постоянным, чтобы призма создала идеальное изображение. Этот зазор зависит от числовой апертуры A' в пространстве, где расположена призма, а также от показателей преломления материалов, из которых выполнена призма.

Литература

1. Кругер М.Я. Справочник конструктора оптико-механических приборов. М.-Л., 1963.
2. Кругер М.Я. Справочник конструктора оптико-механических приборов. М., 1980.
3. Марешаль А., Франсон М. Структура оптического изображения. М., 1964.
4. Грамматин А.П. Автоматизация проектирования оптических систем. Л., 1989.

РАЗРАБОТКА И РАСЧЕТ ОПТИЧЕСКОЙ СИСТЕМЫ ЛАЗЕРНОЙ ТЕХНОЛОГИЧЕСКОЙ УСТАНОВКИ

И.Ю. Суворова

Научный руководитель – д.т.н., профессор В.А. Зверев

Оптическая система служит для передачи излучения с заданным распределением интенсивности от выходного зрачка лазерного усилителя с последующей передачей (фокусировкой) на металлический образец.

Актуальность темы напрямую связана с необходимостью разработки лазерных установок для прodelывания отверстий в материалах диаметром порядка одной десятой мм. Это позволит существенно повысить технологичность в машиностроении.

В данной работе предлагается способ изменения параметров лазерного пучка с помощью оптической системы.

Таким образом, целью работы является разработка и расчет оптической системы, позволяющей преобразовать лазерное излучение в световую трубку диаметром 0.1 мм и длиной 3 мм (см. рис. 1).

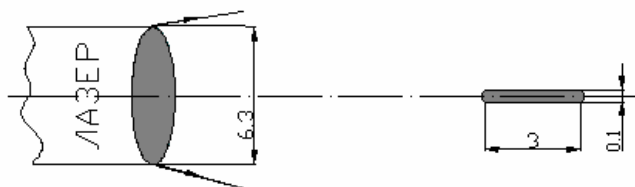


Рис. 1. Схема преобразования

Исходными данными разработки служат параметры лазерного пучка:

- диаметр пучка при дифракционной угловой расходимости $D_{\text{лаз}} = 6.3$ мм;
- длина волны излучения $\lambda = 1.064$ мкм;
- энергия в импульсе $W_{\text{имп}} = 100$ мДж;
- длительность импульса $\tau = 350$ нс.

При разработке оптической системы необходимо учесть ее особенности.

- 1) Отношение требуемой величины угла расходимости пучка лучей к его дифракционной расходимости должно быть больше единицы (см. рис. 2).

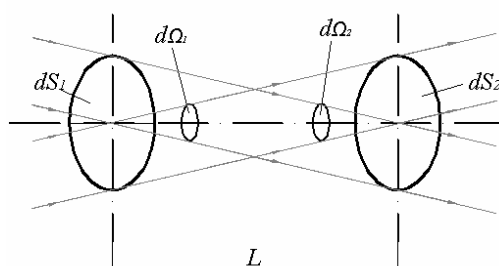


Рис. 2. Соотношение расходимостей в системе

$$d\Omega_T = \pi/4 D_T^2/L_T^2 = \pi\omega_T^2.$$

Угловая расходимость пучка световой трубки: $\omega_T = 1/2 D_T/L_T$. Дифракционная угловая расходимость: $\omega_{dT} = 1.22\lambda/D_T$. Тогда $\omega_T/\omega_{dT} = D_T^2/2.44\lambda L_T > 1$.

В соответствии с исходными данными имеем: $\omega_T/\omega_{dT} = 1.284$.

- 2) Необходимость согласования апертурного и полевого углов.

Из рис. 3 видно, что $2y' = -2(f\theta' - p')\text{tg}\omega$ и $D = (f\theta' - p')\text{tg}\sigma'$. Положив диаметр выходного зрачка равным размеру изображения $D = 2y'$, получим световую трубку. При этом апертурный и полевой углы равны между собой, т.е. $\omega = -\sigma'$.

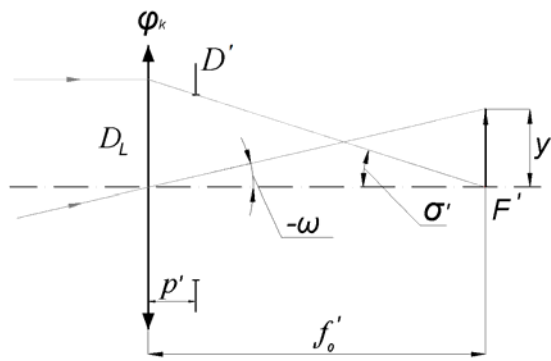


Рис. 3. Согласование апертурного и полевого углов

- 3) В условиях эксплуатации оптическая система должна обладать достаточным запасом лучевой прочности, поэтому плотность потока излучения в любом сечении лазерного пучка лучей на поверхностях и внутри оптической системы не должна превышать предельного значения $E_{\text{пред}}$.

$$E = \Phi_0 / S_p \leq E_{\text{пред}},$$

где E – плотность потока лазерного излучения; $\Phi_0 = 4 e^2 / (e^2 - 1) W_{\text{имп}} / \tau$ – поток излучения; S_p – допустимая площадь поперечного сечения лазерного пучка лучей. При этом предельно допустимая площадь сечения лазерного пучка лучей:

$$S_{p\text{пред}} = \Phi_0 / E_{\text{пред}} = 4 e^2 / (e^2 - 1) W_{\text{имп}} / \tau E_{\text{пред}}.$$

Полагая любое сечение пучка лучей круглым, имеем:

$$S_{p\text{пред}} = \frac{1}{4} \pi D_{\text{пред}}^2, \text{ тогда } D_{\text{пред}} = 4 (e^2 / (e^2 - 1) W_{\text{имп}} / \tau E_{\text{пред}})^{1/2}.$$

Отсюда получаем, что предельный диаметр $D_{\text{пред}}$ равен 0.41 мм.

Для решения поставленной задачи применяется афокальная система. При разработке такой системы рассматривается два варианта. В первом варианте в качестве первой подсистемы используется телескопическая система. Оптическая схема, конструкторский выпуск и конструктивные параметры системы представлены на рис. 4, 5 и в табл. 1 и 2, соответственно.

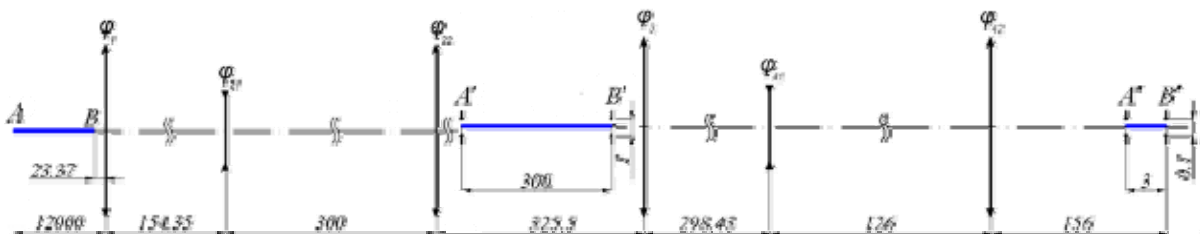


Рис. 4. Оптическая схема

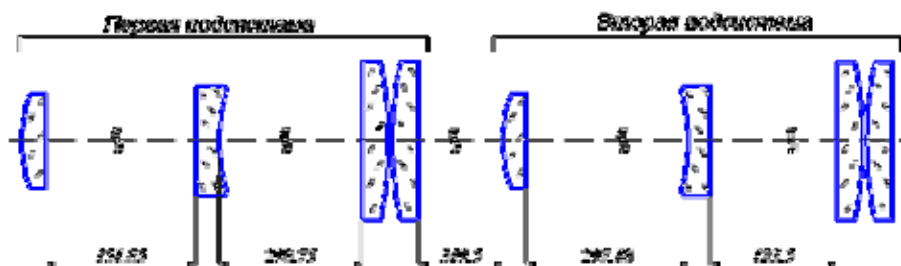


Рис. 5. Конструкторский выпуск

Линейное увеличение	0.016
Числовая апертура (A)	0.00026354
Линейное поле предмета (2Y)	6.3
Положение предмета относительно первой поверхности	-23.37
Положение изображения относительно последней поверхности	154.781
Диаметр входного зрачка	6.24533
Положение входного зрачка относительно первой поверхности	-11872.3
Положение выходного зрачка относительно последней поверхности	151.786
Основная длина волны в нм	1064.0

Таблица 1. Конструкторский выпуск

	R	D	N(0) / КОДЫ
1	106.7	2	1
2	0	151.85	ТК4
3	0	2	1
4	17.36	298.75	Ф2
5	0	2	1
6	-194.5	0.01	ТК4
7	194.5	2	1
8	0	349.25	ТК4
9	190.31	2	1
10	0	297.18	ТК4
11	-19.75	2	1
12	0	123.5	Ф2
13	0	2	1
14	-93.36	0.01	ТК4
15	93.36	2	1
16	0		ТК4
17			1

Таблица 2. Конструктивные параметры

Общая длина всей системы составляет 1.4 м.

Поперечное увеличение изображения, образованного афокальной системой, будет равняться отношению диаметра световой трубки к диаметру лазерного излучения (см. рис. 6), с другой стороны, оно не зависит от расстояния от оптической системы до плоскости предмета и равно отношению оптических сил компонентов с обратным знаком.

Поперечное увеличение изображения: $V = -D_T/D_L = -(1/63)^*$, при этом с другой стороны $V = -f_2/f_1$. Поэтому применение во-втором случае телеобъектива (см. рис. 7) позволяет уменьшить фокусное расстояние первой подсистемы. При этом расстояние от световой трубки до первого компонента во много раз превышает его фокусного расстояния.

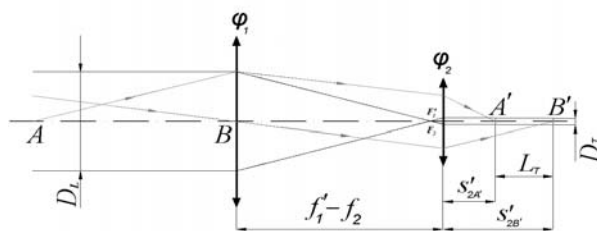


Рис. 6. Отношение диаметров световой трубки и лазерного излучения

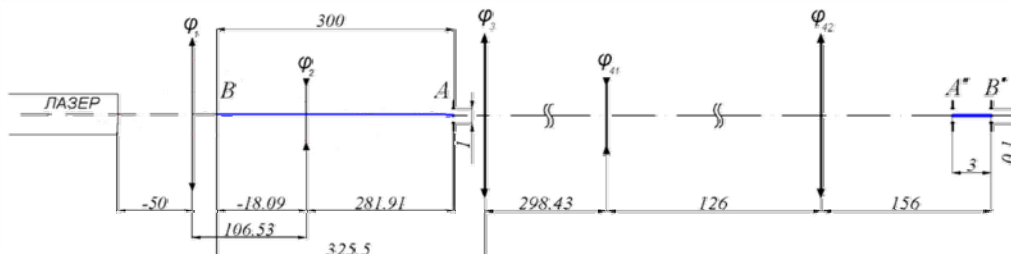


Рис. 7. Оптическая схема телеобъектива

Оптическая схема, конструкторский выпуск и конструктивные параметры системы представлены на рис. 8, 9 и в табл. 3 и 4, соответственно.

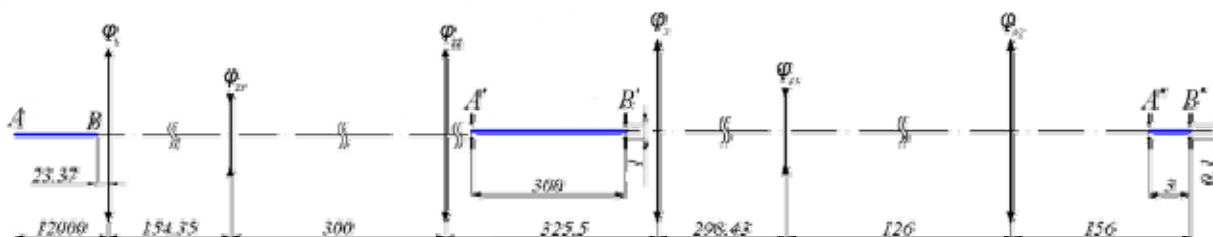


Рис. 8. Оптическая схема

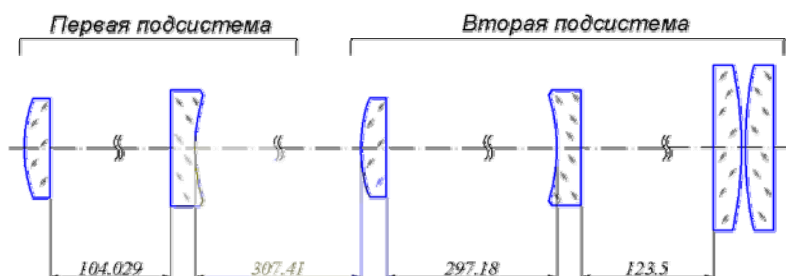


Рис. 9. Конструкторский выпуск

Фокусное расстояние в мм	-188.882
Диафрагменное число	29.9813
Угловое поле (2 омега)	1'49.2"
Положение предмета относительно первой поверхности	∞
Положение изображения относительно последней поверхности	154.784
Диаметр входного зрачка	6.30000
Положение входного зрачка относительно первой поверхности	-50.0000
Положение выходного зрачка относительно последней поверхности	151.788
Основная длина волны в нм	1064.00

Таблица 3. Конструкторский выпуск

	R	D	N(0) / КОДЫ
1	74.93	2	1
2	0	104.029	ТК4
3	0	2	1
4	11.97	307.41	Ф2
5	190.31	2	1
6	0	297.18	ТК4
7	-19.75	2	1
8	0	123.5	Ф2
9	0	2	1
10	-93.36	0.01	ТК4
11	93.36	2	1
12	0		ТК4
13			1

Таблица 4. Конструктивные параметры

В этом варианте общая длина оптической системы лазерной установки составляет чуть больше метра.

Таким образом, в случае применения в качестве первой подсистемы телеобъектива расстояние от выходного окна лазера до сформированной световой трубки существенно меньше, чем при применении второй афокальной системы. Однако при первом варианте решения материальные диафрагмы, реально определяющие световую трубку, расположены в одном и том же пространстве и имеют одинаковый диаметр.

Таким образом, в работе:

- определены условия формирования световой трубки с заданными параметрами.
- определены базовые схемы формирования и преобразования цилиндрической световой трубки.
- определены требования к элементам оптической системы из условия лучевой прочности.
- разработана последовательность формирования исходной световой трубки и ее преобразование до требуемых параметров.
- разработана и рассчитана оптическая система, преобразующая излучения лазера в световую трубку, параметры которой соответствуют исходным данным.

Литература

1. Климков Ю.М. К выбору оптических систем для концентрации излучения ОКГ на сравнительно большом расстоянии. // Светотехника. 1970. №7. С. 25.
2. Климков Ю.М. Основы расчета оптико-электронных приборов с лазерами. М.: Сов. радио, 1978.
3. Русинов М.М. Габаритные расчеты оптических систем. М.: Геодезиздат, 1959.
4. Слюсарев Г.Г. Методы расчета оптических систем. М.: Машиностроение, 1969.
5. Турыгин И.А. Прикладная оптика. / Учебное пособие для вузов. М.: Машиностроение, 1966.

РАЗРАБОТКА АЛГОРИТМА И ПРОГРАММЫ ДЛЯ РАСШИРЕНИЯ ВОЗМОЖНОСТЕЙ АППАРАТУРЫ ОЦЕНКИ КАЧЕСТВА ИЗОБРАЖЕНИЯ ОПТИЧЕСКИХ СИСТЕМ

Ле Зуй Туан

Научный руководитель – д.т.н., профессор В.К. Кирилловский

Для устранения недостатков метода прямого определения ФРТ, основанного на применении ПЗС-камеры, разработаны компьютерные изофотометрические методы экспериментального анализа пятна рассеяния, сформированного при работе изготовленной оптической системы. Эти методы позволяют повысить диапазон регистрации освещенности камеры до 10^3 – 10^4 при устранении погрешностей от нелинейности световой характеристики ПЗС-матрицы.

Введение

Развитие оптико-электронных приборов и комплексов, например, приборов ночного видения, тепловизионных систем, цифровых камер, медицинских приборов, сопровождается применением высококачественных оптических систем. Для получения требуемого качества необходимо измерять основные параметры, характеризующие качество изготовленных оптических систем, с минимальной погрешностью.

Наиболее полной оценкой качества изображения, создаваемого объективом, является распределение освещенности в плоскости изображения точки и в близких к этой плоскости сечениях светового пучка, т.е. определении функции рассеяния точки (ФРТ). Экспериментально найденная ФРТ, характеризующая качество системы, позволяет учитывать все особенности волновой поверхности, сформированной системой, в том числе и характер микрорельефа оптических поверхностей, дефекты оптических материалов, блики, отклонение пропускания (или отражения) на зрачке, сдвиги изображения, возникшие при работе прибора, и т.д. С помощью ФРТ в осредненной и мгновенной форме можно учитывать флюктуации оптических характеристик воздушного тракта, непрерывно изменяющиеся во времени по сложному случайному закону. ФРТ позволяет надежно аттестовать качество изображения системы с несколькими зрачками. ФРТ, являясь первичной характеристикой качества изображения, позволяет перейти к другим требуемым характеристикам качества изображения, таким как ФРЛ, ФПМ (ЧКХ), краевая функция, функция концентрации энергии и другие. Все перечисленные факторы свидетельствуют в пользу методов анализа ФРТ при их цифровой обработке. Эти функции несут информацию сразу обо всех пространственных частотах и могут быть получены за один прием, что существенно сокращает время измерений.

В настоящее время в области оптических измерений широко используют ПЗС-камеры в качестве приемника. С помощью ПЗС-камеры, установленной в фокальную плоскость исследуемого объектива, пятно рассеяния этого объектива сразу регистрируется в виде цифровой картины, представленной значениями ФРТ. Однако ПЗС представляет собой матрицы потенциальных ям, каждая из них работает как отдельный фотоприемник, поэтому напряжения видеосигнала зависят от формы световой характеристики фотоприемника. За пределами этого линейного участка напряжение сигнала не пропорционально уровню освещенности, причем отклонение от пропорциональности изменяется по нелинейному закону. В табл. 1 видно, что фотометрическая структура пятна, которое дает ПЗС-камера, искажается и сильно отличается от дифракционной теории. Кроме того, на картине появляется только первое кольцо, второе, соответствующее уровню относительной освещенности 0,0045, не зарегистрировано, т.е. диапазон регистрации освещенности камеры составляет не более чем 200 раз.

Способ повышения возможности ПЗС-камеры

Для расширения диапазона регистрируемых освещенностей и устранения нелинейной световой характеристики ПЗС-матрицы было разработано алгоритмическое и программное обеспечение, которое основано на регистрации серии фотоснимков пятна рассеяния с переменным временем экспозиции на одном и тоже фотоприемнике, т.е. получении совокупности фотометрических сечений, соответствующих различным уровням равной освещенности. Из каждого фотометрического сечения формируется фигура с четким контуром, которая называется изофотой. Из системы изофот можно получить карту распределения освещенности в пятне рассеяния (рис. 1). Такой метод назван методом изофотометрии с изменяющимся временем накопления [5]. Диапазон регистрации освещенности по данному методу не зависит от протяженности рабочего участка световой характеристики фотоприемника, а зависит от диапазона изменения времени накопления, который свыше 10^3 . В экспериментальной установке была использована матричная видеокамера VAC-135 с диапазоном относительного времени накопления составляет свыше трех порядков ($1 \div 2047$). При этом распределение света регистрируется с диапазоном выше 10 раз по сравнению с возможностью камеры.

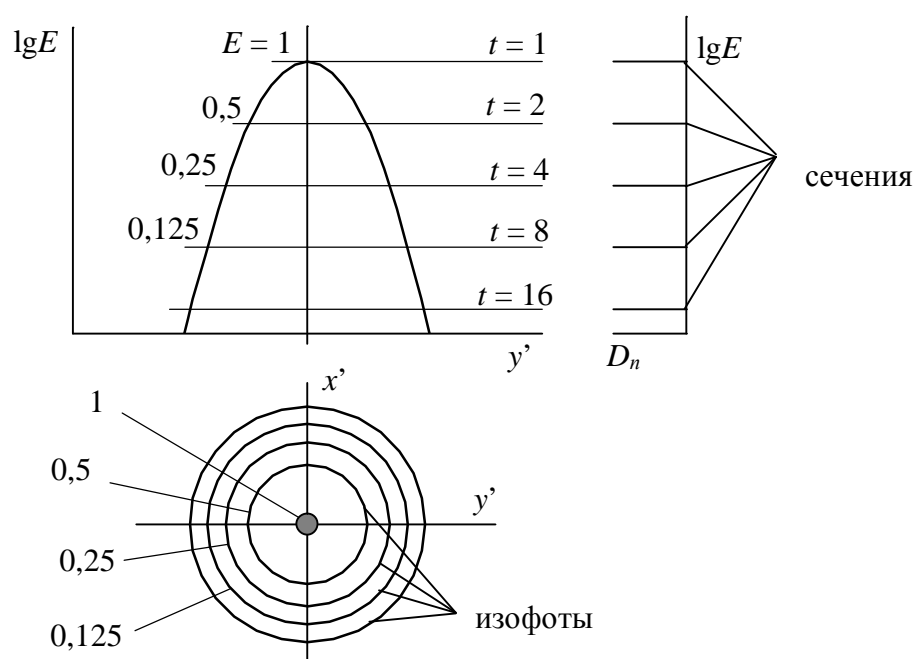
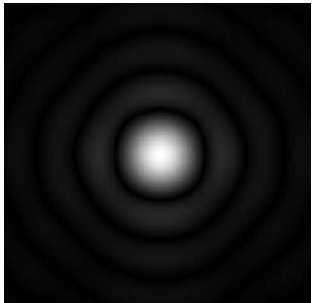

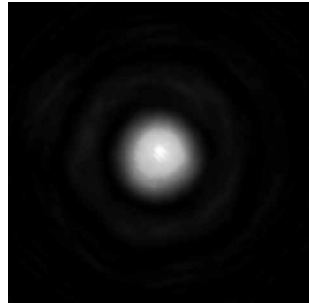
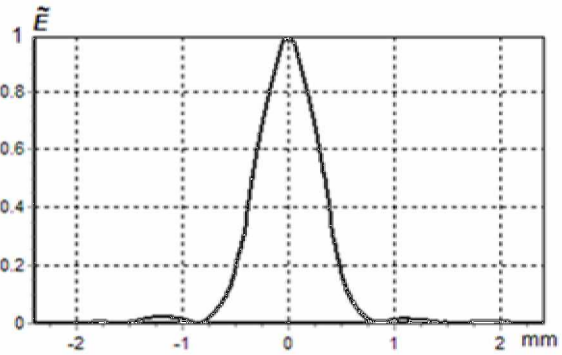
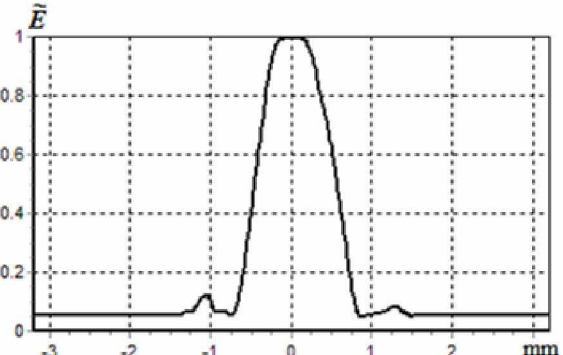
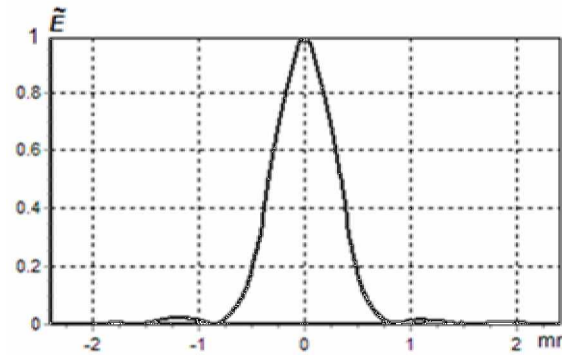
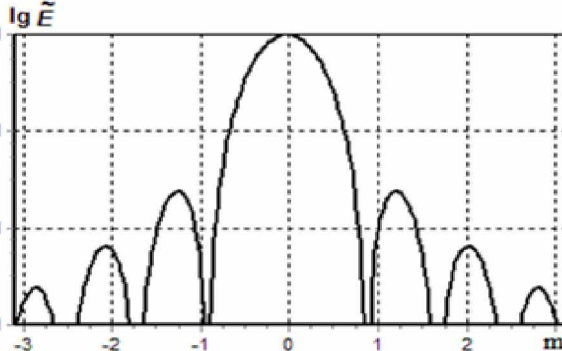
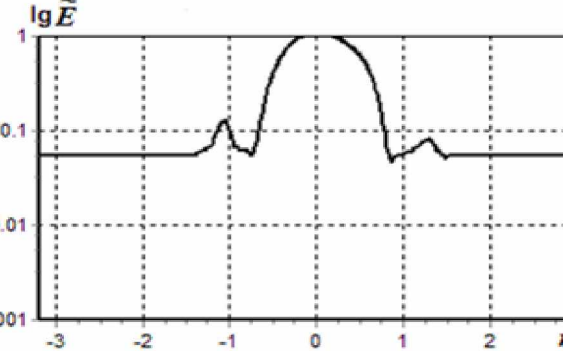
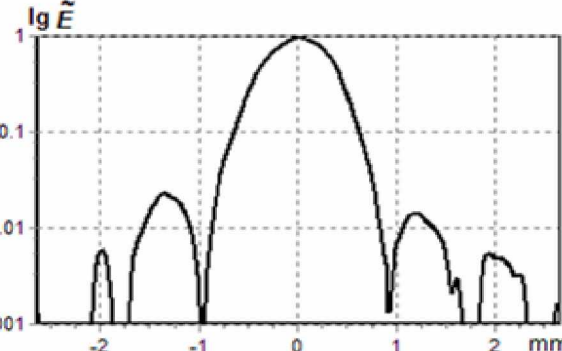


Рис. 1. Схема метода изофотометрии с переменным временем накопления и вид изофотограммы

На рис. 2 изображена принципиальная схема устройства. Установка содержит осветительную систему, в которую входят лазер 1, двухлинзовый объектив осветительной системы 3, формирующий точечный источник света. В осветительную систему введен перемещающийся (вращающийся) фотометрический клин 2, плавно меняющий световой поток. Сформированный точечный источник света посылает расходящийся пучок через объектив коллиматора 5 в зрачок исследуемой системы 6, которая строит пятно рассеяния 7. Это изображение переносится с увеличением при помощи микроскопа 8 на чувствительную площадку приемника изображения 9 (ПЗС-камеры). Камера превращает изображения пятна рассеяния в видеосигнал и посылает в компьютер через порт USB. При помощи разработанной программы "Videoprocess" можно наблюдать пятно рассеяния на мониторе и выполнять съемку последовательных фотокладов в цифровом виде. Полученные цифровые изображения вводят в модуль выделения изофот.

Таблица 1. Дифракционная картина на отверстие диаметром $d = 0.2$ мм, зарегистрирована на расстоянии $l = 230$ мм по разным методам

Пятно, рассчитанное по дифракционной теории	Пятно, которое дает камера EVC-135	Пятно, полученное по методу изофотометрия
		
		
		

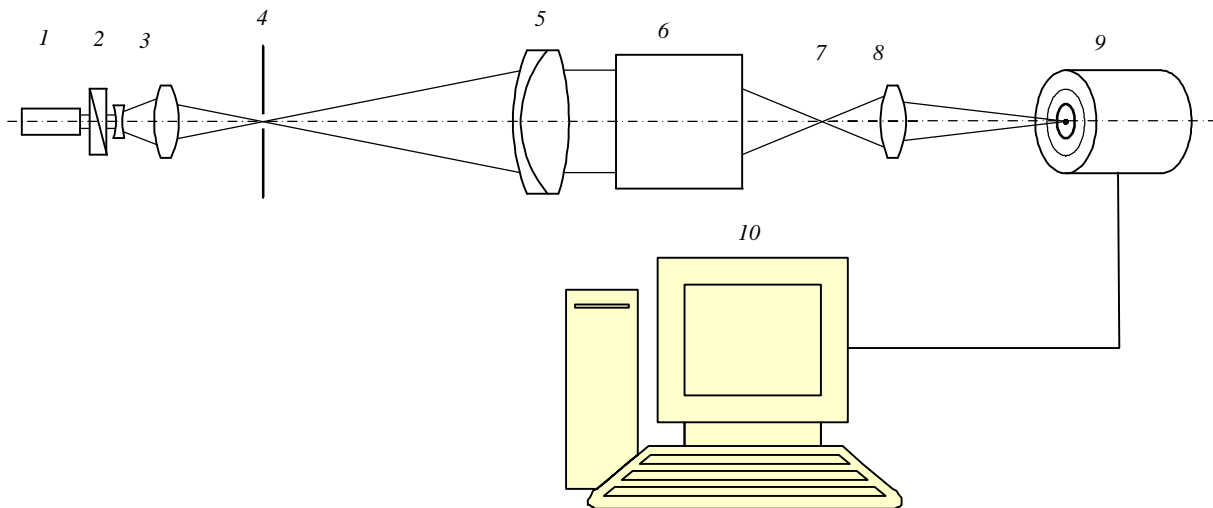


Рис. 2. Схема установки для измерения ФРТ фотообъектива по методу изофотометрии с изменяющимся временем накопления при помощи камеры: 1 – лазер; 2 – фотометрический клин; 3 – система для фокусировки лазерного пучка; 4 – точечная диафрагма; 5 – объектив коллиматора; 6 – исследуемый объектив; 7 – изображение точечного тест-объекта (пятно рассеяния); 8 – проекционный микро-объектив; 9 – видеокамера (приемник изображения); 10 – компьютер

На рис. 3 показано алгоритмическое решение компьютерной версии изофотометрии с переменным накоплением. В исходном состоянии параметр «exposure» камеры установлен на минимальное значение t_1 , а фотометрический клин введен так, что световой поток в схеме контроля формирует изображение, дающее изофоту в виде точки (первая изофота). Эта изофота соответствует максимуму освещенности $E_{омн1}$, которое принимаем равным единице.

В процессе измерения, после установки увеличенного времени экспозиции $t_2 = k \cdot t_1$, программа выполняет регистрацию фотографии пятна рассеяния и вводит ее в модуль выделения изофоты. Эта изофота является контуром фотометрического сечения, соответствующего новому уровню относительной освещенности $E_{омн2} = t_1 / t_2$.

Таким образом, $E_{омн n} = 1 / t_{омн n} = t_1 / t_n$, поэтому помещаем изофоту в систему координат на уровень $E_{омн n}$ и записываем его в память. Когда t превышает максимальное значение t_{max} , работа программы заканчивается, и на мониторе получается графическое отображение ФРТ.

В результате запоминания ряда фотометрических сечений и построения в пространственных координатах их исходного взаимного расположения формируется объемная диаграмма ФРТ, т.е. функция распределения освещенности в изображении точечного источника, построенном исследуемой оптической системой.

Программа и результат

На основе созданного алгоритма разработано программное обеспечение «КИЗО». Таким образом, можно констатировать создание нового метода – компьютерной изофотометрии с изменяющимся временем накопления. С помощью матричной камеры EVC-135 и с применением созданного программного обеспечения метод позволяет определить ФРТ в диапазоне изменения освещенности свыше 10^3 и обработать результаты измерения для получения основных количественных и качественных характеристик изображения исследуемого объектива.



Рис. 3. Алгоритмическое решение компьютерной версии изофотометрии с переменным накоплением

Для экспериментальной иллюстрации всех этапов работы программы склеенная линза служила исследуемым объективом. Ее структурные параметры представлены в табл. 1. На рис. 4 показаны графики ФРТ и ЧКХ, полученные путем расчетов и экспериментов по методу изофотометрии.

Параметры			Материал
r	d	$D_{св}$	Стекло
128,680	10.00	50	К8
-90,860	4.00	50	ТФ1
-275,720		50	

Таблица 2: r – радиусы кривизны поверхностей, d – расстояние между двумя последующими поверхностями, $D_{св}$ – световой диаметр поверхности

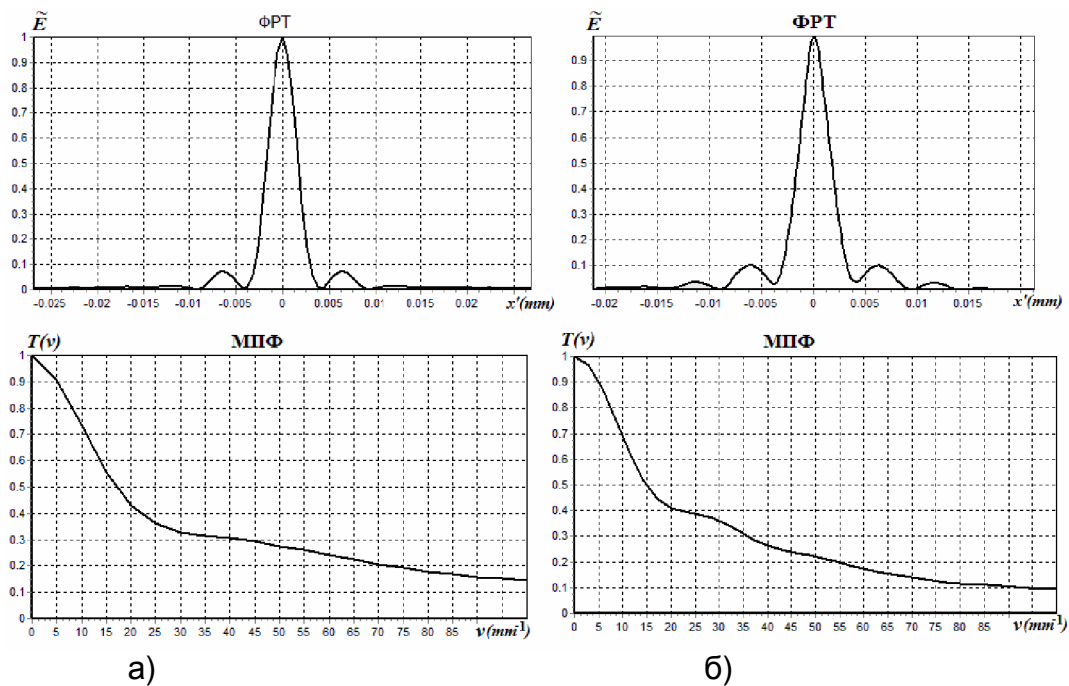


Рис. 4. ФРТ и МПФ склеенного объектива для осевого пучка: а) теоретические результаты; б) экспериментальные результаты по методу изофотометрии

На рис. 4 видно, что теоретическая и экспериментальная МПФ хорошо сходятся. Различие между ними можно объяснить следующим:

- имеются погрешности изготовления оптических поверхностей линз;
- положение плоскости приемника не совпадает с плоскостью, в котором рассчитана МПФ по теории;
- в аппаратуре контроля присутствуют остаточные блики и рассеянный свет.

Заключение

Для устранения недостатков традиционных методов прямого определения ФРТ, основанных на сканировании, и модернизации разработанных изофотометрических методов экспериментального широкодиапазонного исследования характеристик качества изображения оптических систем осуществлена настоящая работа, посвященная развитию данных методов на основе их компьютеризации.

Благодаря разработке алгоритмического и программного обеспечения для установок контроля качества изображения создан новый метод компьютерной изофотометрии с изменяющимся временем накопления световой энергии. Метод позволяет на новом, современном уровне реализовать уникальные возможности изофотометрии с изменяющимся временем накопления, объединяя достоинства широкодиапазонного анализа структуры изображения тест-объекта, высокой скорости и широкого объема компьютерной обработки экспериментальных данных. Создана возможность оперативного определения полного набора характеристик качества изображения, таких как широкодиапазонный двумерный массив значений ФРТ, модуляционная передаточная функция, функция концентрации энергии, функция рассеяния линии и пограничная кривая.

Литература

1. Кирилловский В.К. Оптические измерения Часть 4. Оценка качества оптического изображения и измерение его характеристик. СПб: СПбГУ ИТМО, 2005.
2. Креопалова Г.В., Лазарева Н.Л., Пуряев Д.Т. Оптические измерения. М.: Машиностроение, 1987.
3. Еськова Л.М., Гаврилин. Д.В. Компьютерные методы контроля оптики. Учебно-методическое пособие к лабораторному практикуму. СПб: СПбГУ ИТМО, 2004.
4. Оптический производственный контроль. / Под ред. Д. Малакары. М.: Машиностроение, 1985.
5. Иванова Т.А., Кирилловский В.К. Проектирование и контроль оптики микроскопов. Л.: Машиностроение, 1984.

РАЗРАБОТКА И ИССЛЕДОВАНИЕ КОМПЬЮТЕРИЗИРОВАННОГО ИНТЕРФЕРОМЕТРА С ДИФРАКЦИОННОЙ РЕШЕТКОЙ НА ОСНОВЕ СХЕМЫ РОНКИ И ПРОГРАММНОГО ОБЕСПЕЧЕНИЯ ДЛЯ РАСШИФРОВКИ ИНТЕРФЕРОГРАММ

Ле Зуй Туан, А.А. Лишкевич

Научный руководитель – д.т.н., профессор В.К. Кирилловский

Целью данной работы стали модернизация и компьютеризация интерферометра с дифракционной решеткой на основе схемы Ронки. Достоинства интерферометра: 1) простота и надежность; 2) нечувствительность к вибрациям; 3) использование нелазерного источника с освещением белым светом или светом с выделением любой длины волны (при контроле хроматизма); 4) отсутствие эталонной оптической поверхности. Для компьютерной обработки интерферограмм, полученных на дифракционном интерферометре с решеткой, разработана программа «Tiger», которая позволяет автоматически расшифровывать интерференционные полосы, рассчитывать деформации волнового фронта и характеристики качества изображения оптической системы, такие как ФТР, ФРЛ, ЧКХ.

Введение

В последнее время наряду с классическими интерферометрами, построенными по принципу деления амплитуды световой волны, ведутся работы по созданию и применению интерферометра, построенного по принципу сдвига раздвоенного по амплитуде рабочего волнового фронта либо его деления и последующего сдвига. Внимание исследователей к этим приборам привлечено, главным образом, благодаря их высокой устойчивости к вибрациям. Кроме того, приборы по таким схемам не требуют применения эталонных оптических поверхностей и позволяют вести контроль при освещении белым светом или с выделением любой длины волны.

Для контроля оптических систем микроскопов усовершенствованы схемы интерферометров сдвига на базе применения низкочастотной дифракционной решетки. Такие схемы были предложены в начале прошлого века итальянским ученым Васко Ронки для контроля астрономических систем.

Наблюдая ронкиграмму, легко получить качественное представление о деформации волнового фронта и соответственно о погрешностях контролируемой поверхности. На рис. 1, 2 показаны примеры ронкиграмм и соответствующие им погрешности поверхностей.

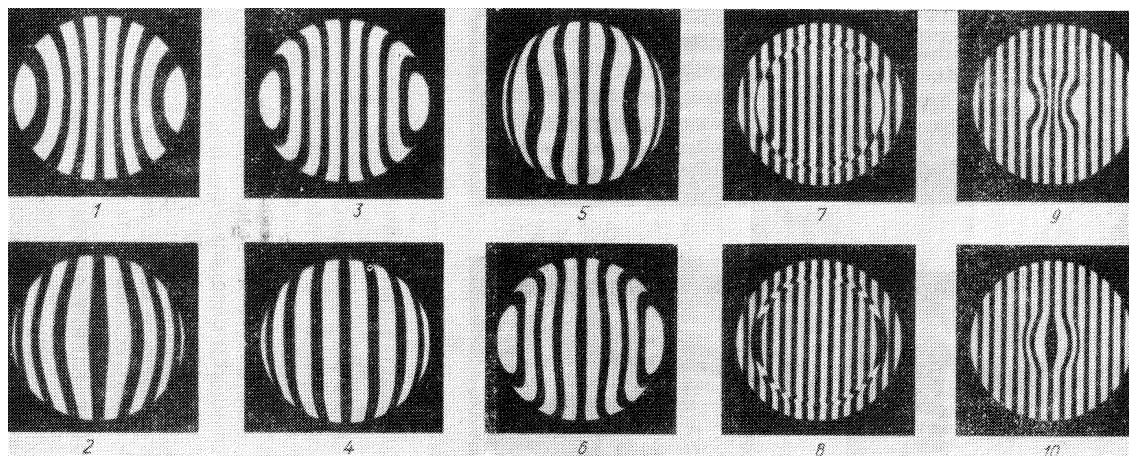


Рис. 1. Ронкиграммы типичных погрешностей поверхности

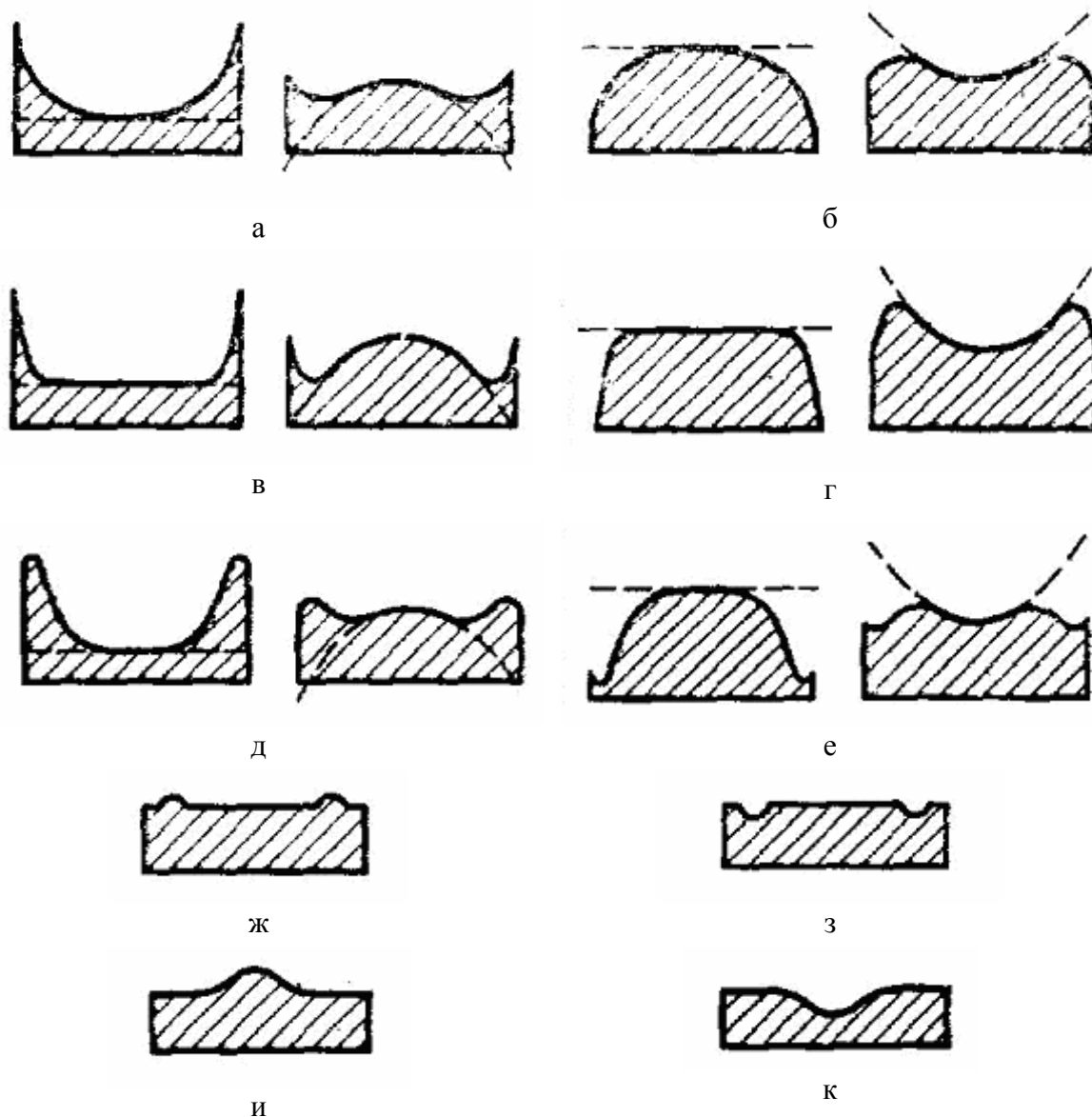


Рис. 2. Погрешности поверхности, соответствующие ронкиграммам: а – недостаточная коррекция третьего порядка; б – избыточная коррекция третьего порядка; в – недостаточная коррекция девятого порядка; г – избыточная коррекция девятого порядка; д – недостаточная коррекция третьего порядка, избыточная коррекция девятого порядка; е – избыточная коррекция третьего порядка, недостаточная коррекция девятого порядка; ж – промежуточная приподнятая зона; з – промежуточная опущенная зона; и – центральный бугор; к – центральная яма

Принцип метода

Принцип действия интерферометра [2, 3] состоит в наложении и интерференции двух взаимно когерентных сферических волновых фронтов в режиме интерференции сдвига (рис. 3).

Сферическая волна, сформированная исследуемой оптической системой, имеющая центр кривизны в точке P , падает на линейную решетку R , установленную нормально к оптической оси вблизи плоскости фокусировки тест-объекта. Решетка имеет пространственную частоту m и отстоит от центра кривизны сферической волны на расстояние y . Это расстояние считается положительным, если отсчет производится в на-

правлении, обратном направлению распространению света. При этом по обеим сторонам от точки P образуются дифракционные спектры различных порядков, имеющие для монохроматического света вид точек. Расстояние между спектрами для малых углов дифракции равно $d = y \cdot \beta$, где $\beta = m\lambda$ – угол дифракции. Отсюда $d = ym\lambda$.

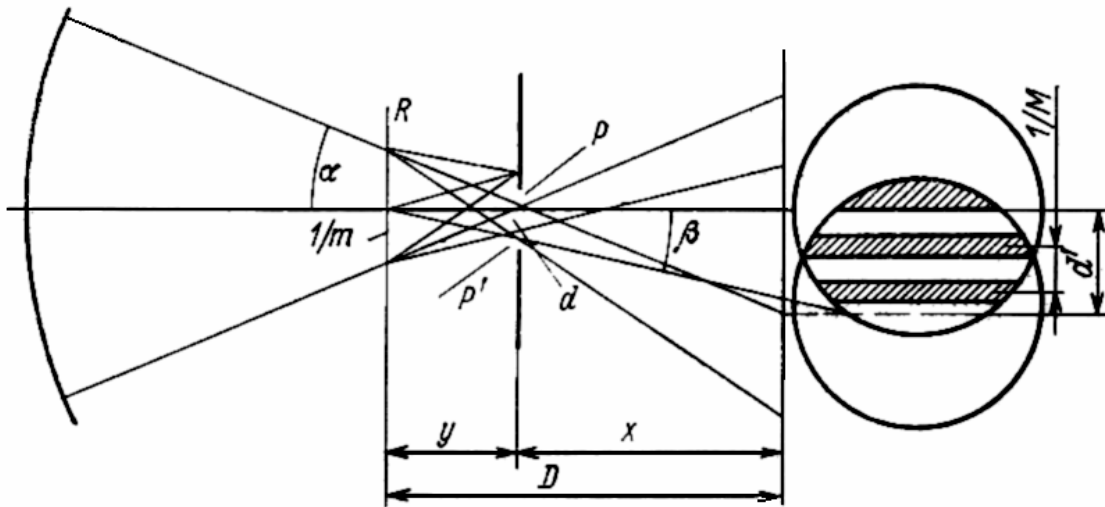


Рис. 3. Схема интерферометра Ронки

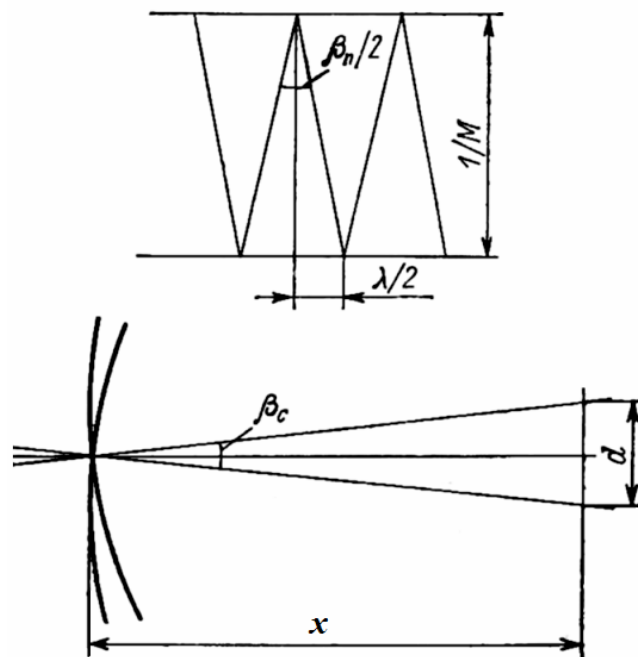


Рис. 4. Принцип образования полос в интерферометре с решеткой, определение шага интерференционных полос

Итак, решетка порождает ряд дифрагированных волн. Диафрагмой выделим из всей совокупности дифрагированных фронтов только два соседних фронта, повернутых на угол β_c и смещенных на расстояние d один относительно другого (рис. 4).

После концентрации в точках P и P' сферические фронты расходятся, образуя в той области пространства, где они перекрываются с заклоном на угол β' , интерференционные полосы (рис.6а). (Под углом β' наблюдаются дифракционные спектры P и P' из центра зрачка Z исследуемой оптической системы (рис. 6б).

$$\beta' = d / x.$$

Расстояние между интерференционными полосами (шаг интерференционных полос) составляет $1/M = \lambda/\beta' = \lambda x/d$. Имея в виду, что $d = ym\lambda$, находим

$$\frac{1}{M} = \frac{\lambda x}{ym\lambda} = \frac{x}{my}.$$

Отсюда получаем соотношение:

$$\frac{M}{y} = \frac{m}{x}.$$

Физический смысл сокращения величины λ очень интересен. Он состоит в том, что работа данного интерферометра сочетает явления как интерференции, так и дифракции. При дифракции увеличение длины волны λ влечет за собой увеличение расстояния d между спектрами, как видно из формулы. Это соответствует увеличению угла взаимного заклона дифрагированных фронтов β' и должно было бы сопровождаться увеличением частоты интерференционных полос.

Однако в процессе интерференции дифрагированных фронтов, наложенных под углом β' , то же увеличение λ влечет за собой, как видно из выражения, пропорциональное уменьшение частоты полос, соответствующих той же величине деформации (в данном случае заклона) волнового фронта. В итоге частота интерференционных полос в интерферометре Ронки (для ахроматической исследуемой системы) не зависит от длины волны излучения, а связана лишь с частотой решетки и геометрическими соотношениями в схеме, что, при небольших апертурах исследуемых систем, позволяет описывать механизм формирования полос Ронки просто теневым отображением штрихов решетки (рис. 5). Это же позволяет вести исследования интерферометре данного типа с применением источника белого света.

При наличии продольного хроматизма или сферохроматизма в исследуемой системе наблюдается окрашивание интерференционных полос, полученных в белом свете. Это объясняется изменением длины отрезка y для разных длин волн при указанных aberrациях. Это свойство данного интерферометра позволяет использовать его при исследованиях хроматизма.

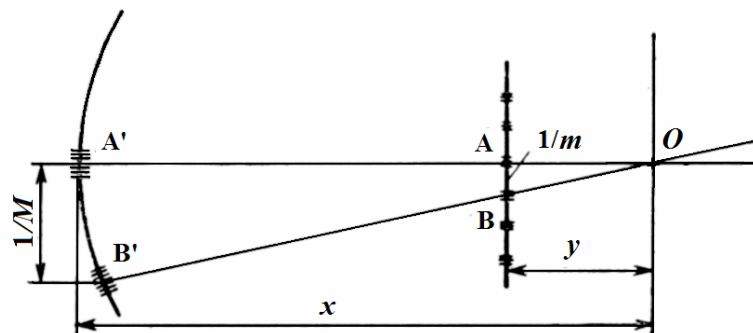


Рис. 5. Формирование полос при малых углах дифракции

Итак, частота интерференционных полос связана с частотой решетки соотношением $M = m \cdot y/x$.

При отсутствии деформаций наложенных фронтов эти полосы прямолинейны, параллельны и отстоят друг от друга на равные промежутки.

Таким образом, возникающая в этой схеме интерференционная картина есть результат сложения и интерференции сдвига рабочей волны, раздвоенной в результате дифракции на решетке и несущей информацию об aberrациях исследуемой оптической системы.

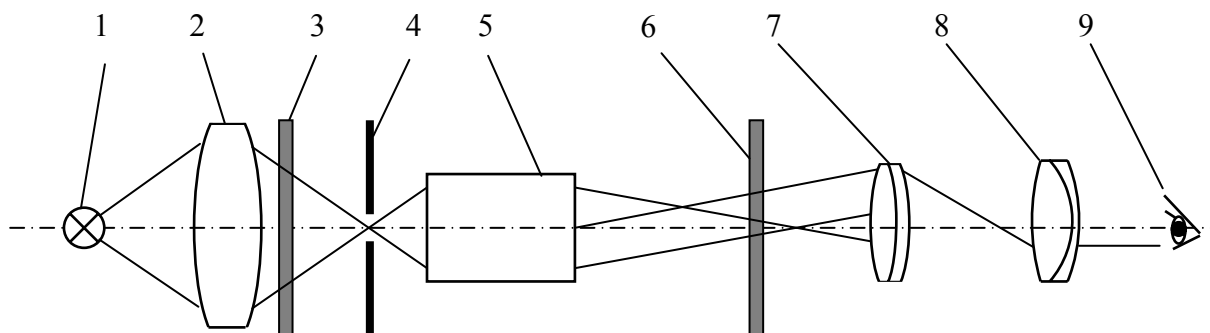


Рис. 6. Принципиальная схема интерферометра: 1 – лампа; 2 – конденсор; 3 – светофильтр; 4 – точечная или щелевая диафрагма; 5 – исследуемый микрообъектив; 6 – дифракционная решетка; 7 – объектив наблюдательной системы; 8 – окуляр наблюдательной системы; 9 – глаз наблюдателя

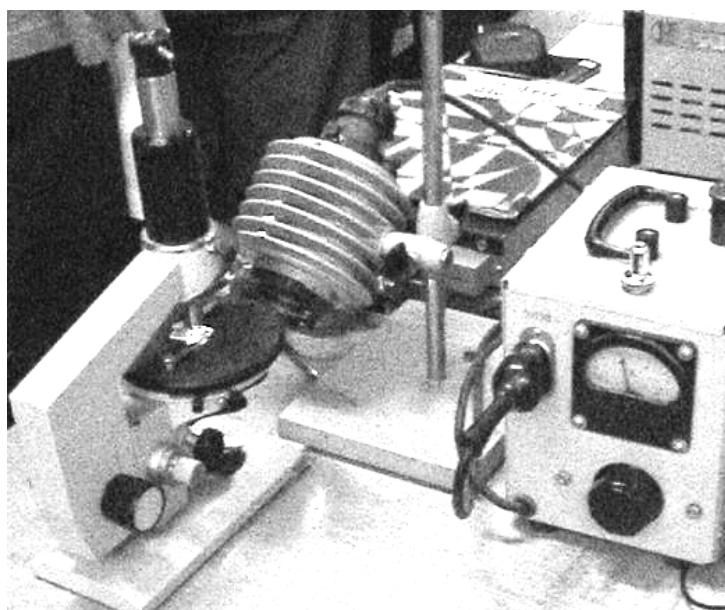


Рис. 7. Интерферометр для контроля оптики микроскопа

Данная работа была направлена на решение такой актуальной задачи, как модернизация интерферометра на базе схемы Ронки, что способствовало бы расширению возможностей данного прибора и упрощению обработки результатов.

Традиционный интерферометр на базе схемы Ронки рассчитан на визуальное восприятие интерферограмм. В этом случае можно судить о максимальных ошибках поверхности по стрелкам прогиба полос. Такого анализа не всегда бывает достаточно, и возникает необходимость прибегнуть к специальным компьютерным программам, предназначенных для более полного и точного анализа.

Проблему компьютеризации интерферометра на базе схемы Ронки можно решить, оснастив его цифровой фотокамерой.

Цифровая фотокамера в схеме прибора заменяет глаз наблюдателя, что позволяет получать фотографии интерференционных картин. Таким образом, создается возможность оперативного ввода информации в компьютер для последующей обработки.

Следует отметить, что компьютеризация интерферометра на базе схемы Ронки предоставляет и другие возможности. Так, например, можно использовать графический редактор для улучшения качества полученных изображений интерферограмм. На фотографиях интерферограмм можно инструментально измерить ошибки волнового фронта

по прогибам полос, в том числе на персональном компьютере. Возможно хранение и систематизация данных.

В данной работе использовалась фотокамера Olympus Camedia C-370. Указанная цифровая камера оснащена зум-объективом ($f'_{об}$ 5,8–17 мм; 1/2,9–1/5,0; 3,2^x) и имеет переменный угол поля зрения. Таким образом, создается возможность оптимизировать масштаб и качество изображения интерференционных картин для исследуемых оптических систем и поверхностей различных апертур.

Исследования интерферограмм

Модернизированный интерферометр на базе схемы Ронки позволяет проводить полноценные и подробные исследования. В процессе работы был исследован действующий макет установки. Результаты исследований, проведенных с помощью этого макета, приведены в данном разделе.

Были получены качественные фотографии интерферограмм. Фотографии интерферограмм были введены в компьютер и успешно обработаны в программе Adobe Photoshop.

Рассмотрим принцип восстановления волнового фронта при обработке интерферограммы Ронки. Основная формула для геометрической модели метода Ронки [1] имеет вид

$$\frac{\partial w \cos \varphi}{\partial x} - \frac{\partial w \sin \varphi}{\partial y} = -\frac{nd}{r}, \quad (1)$$

где φ – угол между линиями решетки и осью y ; r – радиус кривизны волнового фронта; w – деформация волнового фронта; x и y – координаты, нормальные к оптической оси; n – нумерация полос. При $\varphi = 0$, т.е. когда линии параллельны оси y , интерферограмма по этой оси описывается формулой:

$$\frac{\partial w}{\partial x} = -\frac{nd}{r}. \quad (2)$$

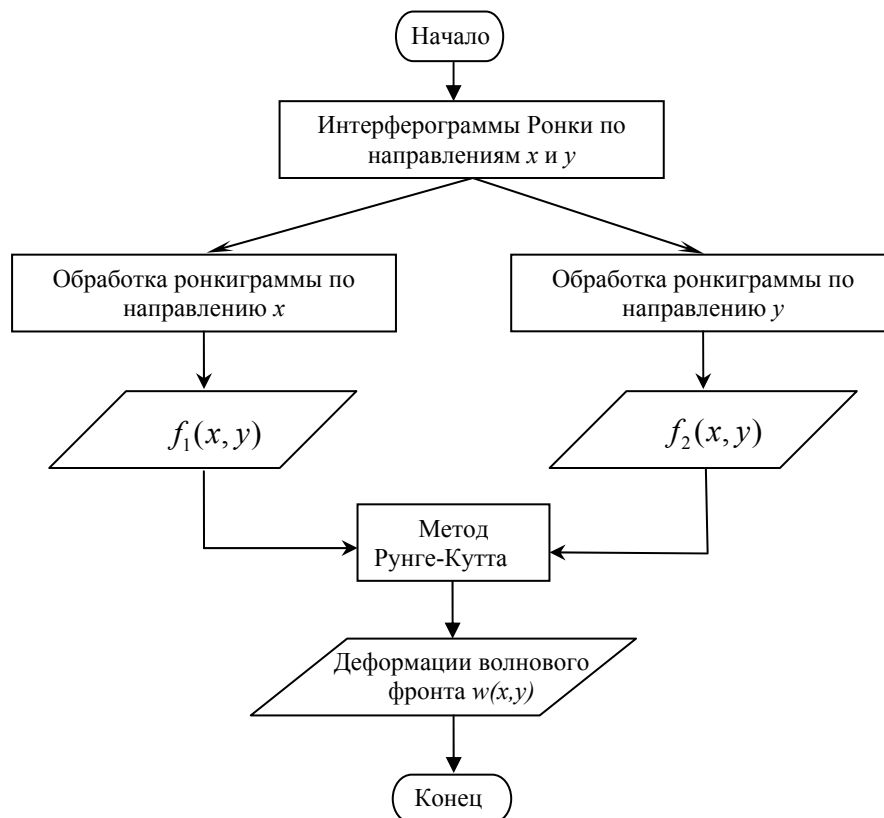


Рис. 8. Алгоритм обработки ронкиграммы

Ронкиграмма в этом случае зависит только от первой производной волнового фронта по направлению x , поэтому при ее обработке получим

$$f_1(x, y) = \partial w / \partial x. \quad (3)$$

Подобным образом, при обработке интерферограммы, полученной, когда линии решетки параллельны оси x ($\varphi = 90^\circ$), получим $f_2(x, y) = \partial w / \partial y$. При обработке интерферограммы по двум направлениям получим дифференциальные уравнения:

$$\begin{cases} \frac{\partial w(x, y)}{\partial x} = f_1(x, y) \\ \frac{\partial w(x, y)}{\partial y} = f_2(x, y) \end{cases}. \quad (4)$$

Решение системы дифференциальных уравнений (4) дает полный волновой фронт $w(x, y)$. Систему уравнений (4) можно решать численным методом Рунге-Кутты. При этом необходимо выбрать начальное условие. Удобно выбрать $W|_{\substack{x=0 \\ y=0}} = 0$.

Схема алгоритма обработки ронкиграмм показана на рис. 8. На основе этого алгоритма разработана программа «Tiger», которая позволяет восстанавливать волновой фронт, обрабатывая две ронкиграммы, полученные при ориентации решетки последовательно по двум направлениям, анализировать результаты для получения характеристик качества оптических систем, таких как, например, ФРТ, ФРЛ и ЧКХ.

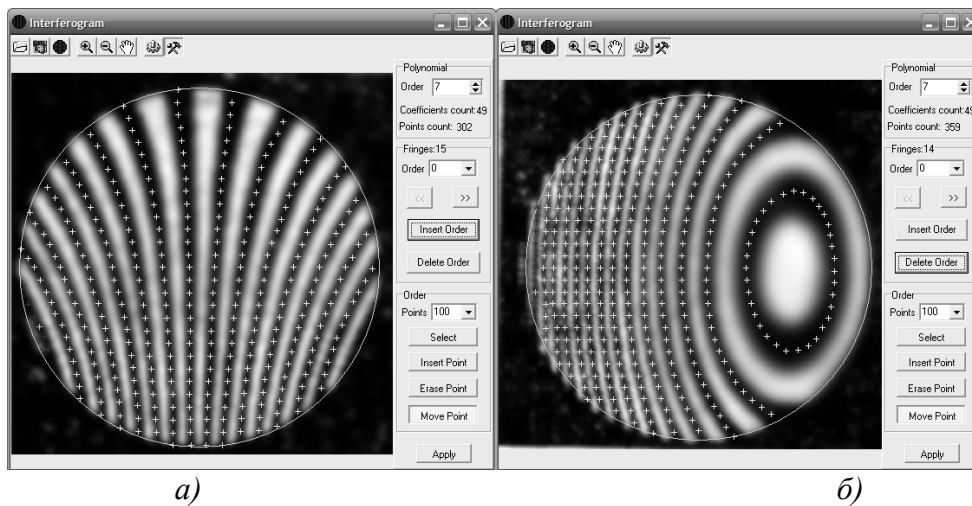


Рис. 9. Ронкиграммы и их обработка в программе «Tiger»: а) по направлению x ; б) по направлению y

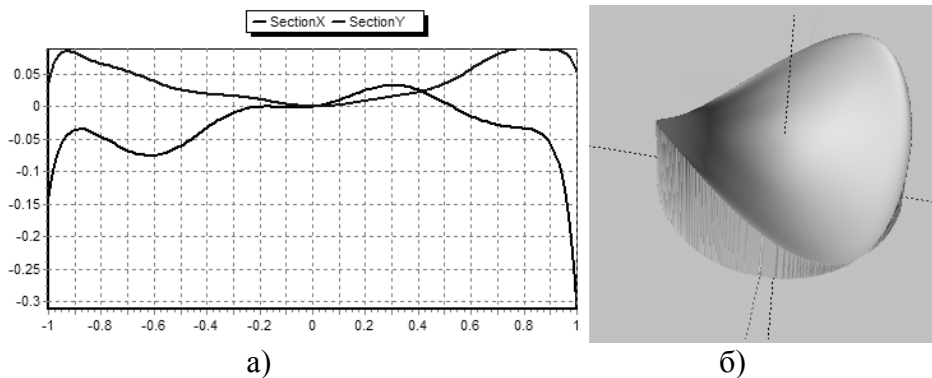


Рис. 10. Результат обработки Ронкиграммы по направлению x : а) кривая $f_1(x, y)$ по двум сечениям; б) объемная фигура $f_1(x, y)$

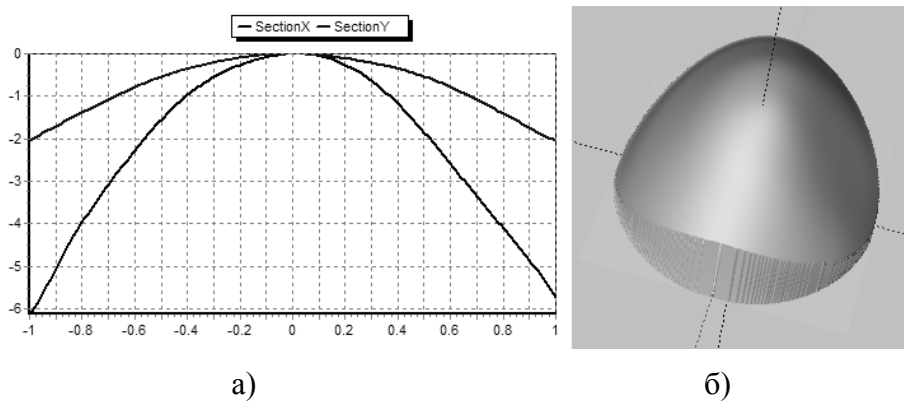


Рис. 11. Результат обработки ронкиграммы по направлению y
 а) кривая $f_2(x, y)$ по двум сечениям; б) объемная фигура $f_2(x, y)$

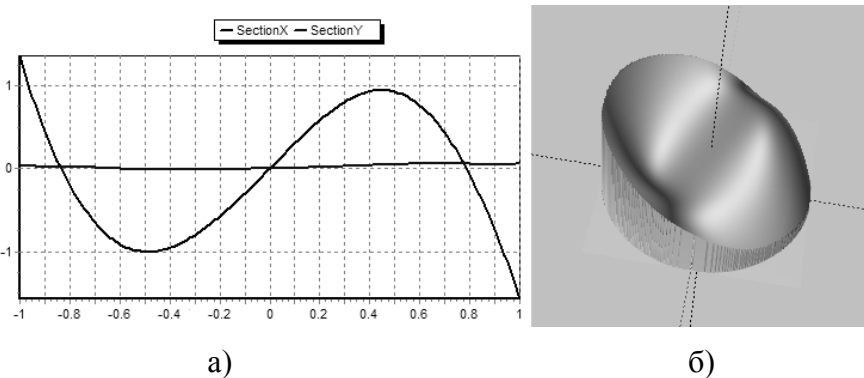


Рис. 12. Результат обработки ронкиграммы после решения системы дифференциальных уравнений по направлениям x и y . Получена полная картина поверхности деформаций волнового фронта

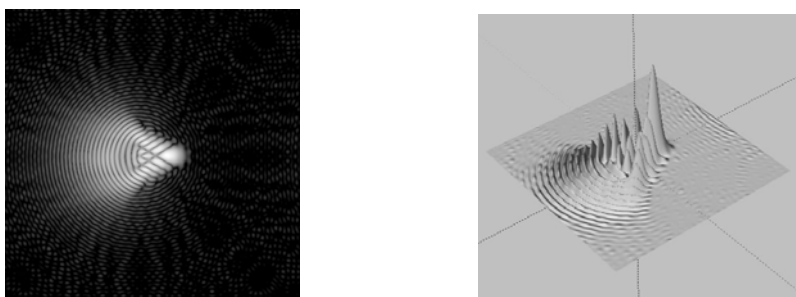


Рис. 13. Изображение точки и трехмерное отображение ФРТ, полученные при преобразовании Фурье поверхности деформаций волнового фронта

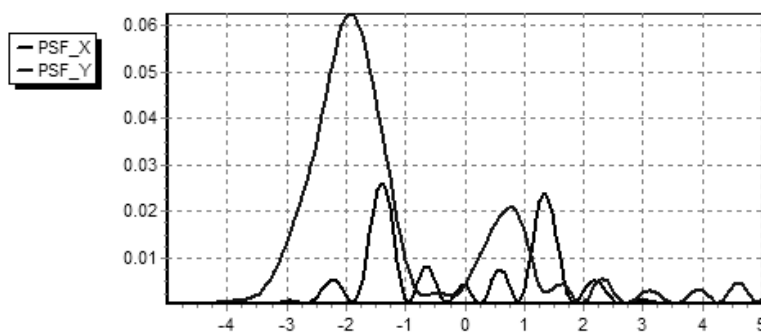


Рис. 14. График функции рассеяния точки в каноническом масштабе пространственной координаты (x, y) . Оцифровка оси ординат выполнена исходя из числа Штреля

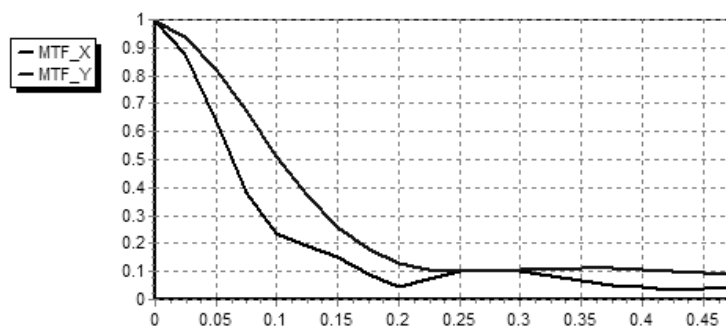


Рис. 15. ЧКХ в каноническом масштабе

Для иллюстрации и проверки работы программы на рис. 9–15 показаны результаты обработки ронкиграмм комы, представленных в [1] (рис. 9.5, стр. 225).

Заключение

Данная работа ставила цель компьютеризацию интерферометра на базе схемы Ронки, путем агрегатирования интерферометра Ронки цифровой фотокамерой. Выполненные работы позволило создать комплекс компьютеризированного интерферометра Ронки с новыми возможностями:

- экранное отображение и измерение координат интерференционных полос с большой точностью;
- возможность регистрации интерференционных картин и компьютерной обработки их структуры, что дает положительный эффект в точности исследования;
- применение компьютера дает возможность автоматической расшифровки интерферограммы путем опознавания координат полос;
- применение компьютерных программ создает возможность всесторонней обработки информации с целью дальнейшего получения характеристик качества исследуемой системы, таких как ФРТ, ФРЛ и ЧКХ;
- компьютерная обработка позволяет отобразить результаты интерферометрии Ронки в форме теневой картины исследуемой системы, что в ряде случаев улучшает условия оценки исследуемого объекта.

Литература

1. Малакара Д. Оптический производственный контроль. М.: Машиностроение, 1985.
2. Иванова Т.А., Кирилловский В.К. Проектирование и контроль оптики микроскопов Л.: Машиностроение, 1984.
3. Кирилловский В.К. Оптические измерения. Часть 5. Изд-во СПб: СПбГУ ИТМО, 2006.

ИССЛЕДОВАНИЕ МУАРОВОГО ЭФФЕКТА, ВОЗНИКАЮЩЕГО ПРИ НАЛОЖЕНИИ ЧАСТОТНЫХ СПЕКТРОВ ОБЪЕКТА

А.Н. Иванов

Научный руководитель – к.т.н., доцент В.Н. Назаров

Рассмотрен дифракционный способ образования муаровых полос с помощью ранее исследованной «зеркальной» апертуры. Реализованы оптические системы, позволяющие получить муаровые картины с произведением и вычитанием дифракционных спектров. Показано, что наличие муаровой картины позволяет увеличить чувствительность дифракционных измерений на порядок. Предложены области применения описанных схем.

Введение

Дифракционные методы контроля (ДМК) геометрических параметров и пространственного положения объектов достаточно часто используются в измерительной технике, когда требуется высокая точность, бесконтактность, малое время и локальность измерений и автоматизация процесса измерения. Препятствием для дальнейшего увеличения точности ДМК является недостаточная чувствительность из-за регистрации интенсивности дифракционной картины квадратичными приемниками. Поэтому дальнейшие исследования в целях повышения точности ДМК и упрощения схем их реализации требуют создания оптической системы, использующей амплитудно-фазовое распределение световых полей за контролируемым объектом [1, 2].

Муаровые полосы, образованные умножением частотных спектров

В работе [3] было приведено описание дифракционной картины от «зеркальной» апертуры, образованной краем объекта и его изображением в плоском зеркале (рис. 1). Особенностью такой дифракционной картины является наличие в ней дополнительной системы поперечных полос (рис. 2), пересекающих дифракционные полосы, если край объекта не параллелен отражающей поверхности.

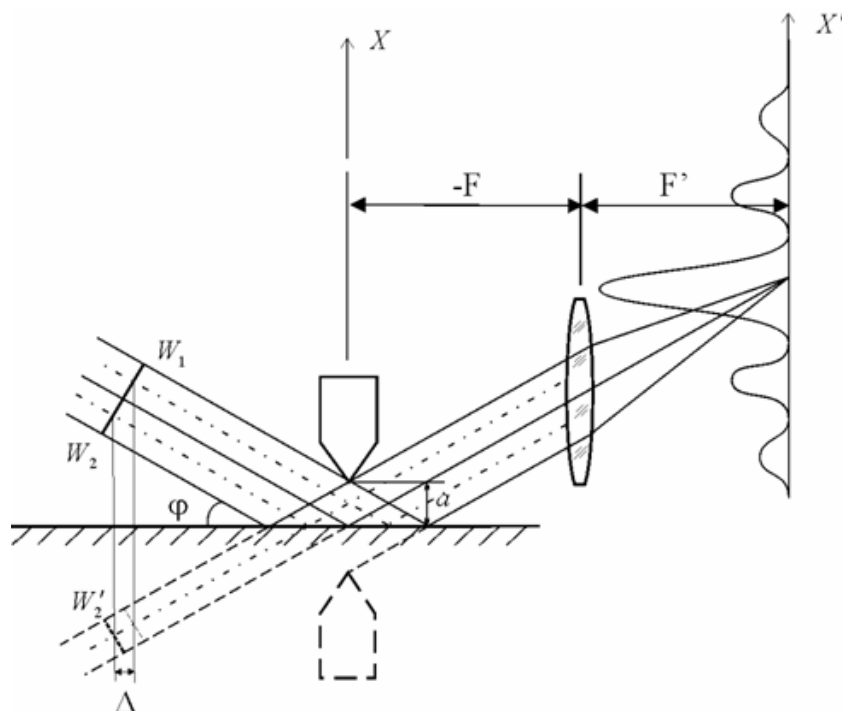


Рис. 1. Схема получения муаровой картины с умножением частотных спектров

Проведенные нами исследования позволили получить аналитическое выражение, описывающее дифракционную картину в дальней области от такой апертуры [4]:

$$I(x', y') = (a(y)^2 / f \lambda) \text{sinc}(\omega(x') a(y) / 2)^2 \cos((\omega(x') a(y) + k a(y) \varphi) / 2)^2, \quad (1)$$

где $\omega(x') = k x' / f$ – пространственная частота, $a(y)$ – функция, описывающая изменение ширины зазора между зеркалом и краем объекта, φ – угол падения волнового фронта на апертуру, $y = y'$. Модель дифракционной картины от «зеркальной» апертуры, ширина которой изменяется по линейному закону, приведена на рис. 3.

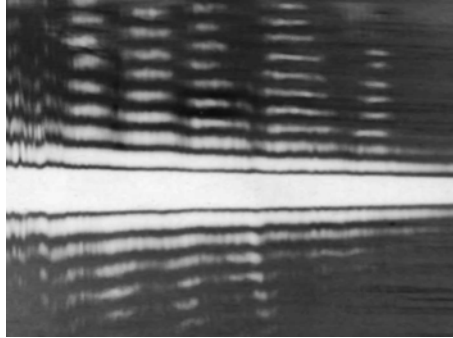


Рис. 2. Дифракционная картина, содержащая муаровые полосы, от апертуры с параметрами $a = 0.05$ мм, $\theta = 4.6 \times 10^{-3}$ рад., $f = 150$ мм

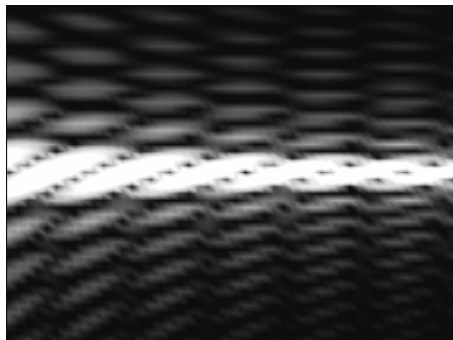


Рис. 3. Модель дифракционной картины от апертуры с параметрами $a = 0.05$ мм, $\theta = 4.6 \times 10^{-3}$ рад., $f = 150$ мм

Эти полосы могут быть объяснены как муаровые. Частотные спектры, описываемые sinc-функцией и cos-функцией, можно считать амплитудно-фазовыми решетками с функциями пропускания:

$$t_1(x', y') = \text{sinc}(\omega(x') a(y) / 2),$$

$$t_2(x', y') = \cos((a(y) \omega(x) - k a(y) \varphi) / 2).$$

Тогда наложение этих решеток приведет к муаровому эффекту. Приравняв $t_1(x', y') = 0$, $t_2(x', y') = 0$, решим эти уравнения относительно порядков их минимумов:

$$m = \frac{a(y) \omega(x)}{2\pi},$$

$$n = \frac{a(y) \omega(x) - k a(y) \varphi - \pi}{2\pi}.$$

Согласно [5], разность полученных параметрических уравнений даст нам параметрическое уравнение муаровых полос

$$p = m - n = \frac{k a(y) \varphi - \pi}{2\pi}. \quad (2)$$

Если функция $a(y)$ является линейной: $a(y) = a_0 + \theta y$, то уравнение муаровых полос будет:

$$y = \left(\frac{(p+1/2)2\pi}{k\varphi} - a_0 \right) / \theta.$$

Ширина муаровых полос будет тогда $S = y(p_2) - y(p_1) = \lambda / \theta \varphi$.

Муаровые полосы, образованные вычитанием частотных спектров

Рассмотренная в предыдущем разделе муаровая картина, полученная умножением частотных спектров, имеет серьезный недостаток – муаровые полосы расширяются с увеличением порядка дифракционной картины. Это связано с тем, что в высоких порядках резко падает крутизна sinc -функции, что ведет к падению контраста [6]. Поэтому требуется получить такую оптическую схему, которая позволяет добиться более равномерного распределения сигнала в области локализации муаровых полос.

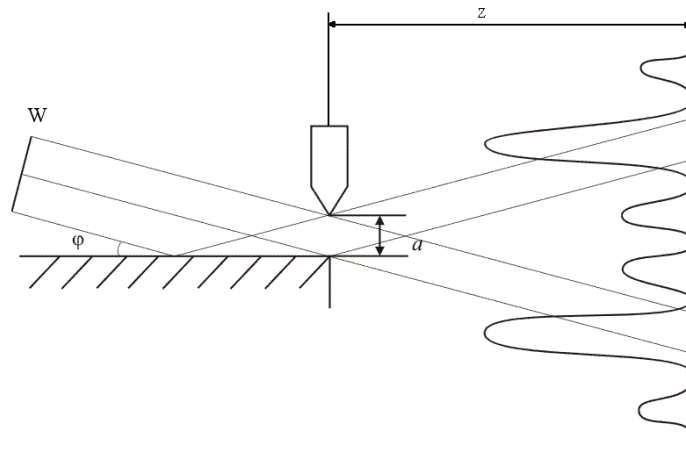


Рис. 4. Схема получения муаровых полос с вычитанием частотных спектров

Для этого было предложено модернизировать «зеркальную» апертуру – расположить край объекта так, чтобы он совпадал с краем зеркала (рис. 4). В этом случае в плоскости регистрации мы будем наблюдать две дифракционные картины, разнесенные на расстояние $l = 2z\varphi$. Распределение амплитуды в дальней области при освещении плоской волной единичной амплитуды будет:

$$U(x') = U1(x') + U2(x'), \quad (3)$$

$$U1(x') = \frac{\exp(ikz)}{\sqrt{i\lambda z}} \int_{-a/2}^{a/2} U(x) \exp(i(\omega(x') + k\varphi)x) dx,$$

$$U2(x') = \frac{-\exp(ikz)\exp(-ika\varphi)}{\sqrt{i\lambda z}} \int_{-a/2}^{a/2} U(x) \exp(i(\omega(x') - k\varphi)x) dx.$$

Взяв интегралы в выражении (3), и умножив полученное выражение на комплексно-сопряженное значение, получим выражение для распределения интенсивности:

$$I(x') = (a(y)^2 / z\lambda) (\text{sinc}(a(\omega(x') + k\varphi)/2) - \text{sinc}(a(\omega(x') - k\varphi)/2))^2. \quad (4)$$

В этом случае произойдет наложение порядков дифракционных картин, что приведет к возникновению системы муаровых полос, эквивалентных муаровым полосам, образованным разностью функций пропускания двух решеток [7].

Рассмотрим случай, когда $a = f(y)$. Тогда дифракционная картина должна содержать систему поперечных полос, так как при определенных значениях $a(y)$, кото-

рым соответствует совпадение максимумов дифракционных картин, будет происходить падение интенсивности. На рис. 5 и рис. 6 приведены экспериментально полученная картина такого муара и модель для случая, когда ширина апертуры меняется по линейному закону $a(y) = a_0 + \theta y$.



Рис. 5. Дифракционная картина, содержащая муаровые полосы, от апертуры с параметрами $a = 0.05$ мм, $\theta = 7 \times 10^{-3}$ рад., $z = 180$ мм

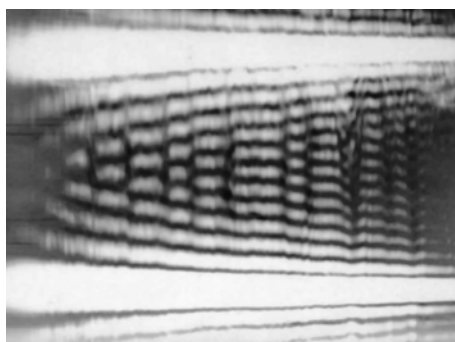


Рис. 6. Модель дифракционной картины от апертуры с параметрами $a = 0.05$ мм, $\theta = 7 \times 10^{-3}$ рад., $z = 180$ мм

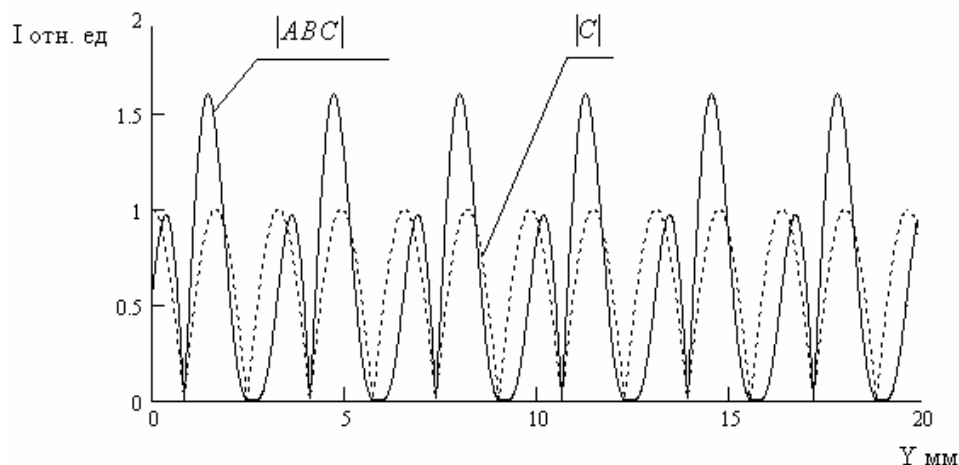


Рис. 7. Образование поперечных муаровых полос в дифракционной картине

Выведем уравнение муаровых полос для такой формы апертуры. Необходимо учесть, что в этом случае апертура имеет вид прямоугольной трапеции, и смещение ее центра вдоль оси y приводит к появлению дополнительных фазовых множителей вида $\exp(-i\theta y(\omega(x') + k\varphi)/2)$ и $\exp(-i\theta y(\omega(x') - k\varphi)/2)$ перед $U1$ и $U2$ соответственно в выражении (3). Поэтому вместо выражения (4) мы получим следующее выражение:

$$I(x', y') = A(x', y')^2 + B(x', y')^2 - 2 A(x', y') B(x', y') C(y'), \quad (5)$$

где $A = \text{sinc}(a(y)(\omega(x') + k\varphi)/2)$, $B = \text{sinc}(a(y)(\omega(x') - k\varphi)/2)$, $C = \cos(k\theta y\varphi)$, $y = y'$.

Анализ выражения (5) показывает, что темные поперечные полосы возникают в том случае, если функция ABC обращается в нуль. Этим нулевым значениям соответствуют решения уравнения $C = \cos(k\theta y\varphi) = 0$ (рис. 7). Отсюда можно определить уравнение муаровых полос

$$y = \frac{\lambda(2p+1)}{4\theta\varphi}. \quad (6)$$

Ширина муаровых полос будет тогда $S = y(p_2) - y(p_1) = \lambda/2\theta\varphi$.

Заключение

В работе рассмотрены способы увеличения точности ДМК путем использования фазовой информации о форме контролируемого объекта, которая содержится в его частотном спектре. Это позволило получить муаровые картины умножения и вычитания, которые обладают высокой чувствительностью к изменению геометрических параметров контролируемых объектов. Проведенные теоретические и экспериментальные исследования показывают, что в линейной мере чувствительность может достигать 0.1 мкм. Это позволяет применять предложенные схемы для контроля формы кромок объектов, например лекальных линеек и угольников, ножей спектральных щелей, деформации объектов в режиме реального времени, контроля вибраций и биений.

Подобная схема может быть также использована для контроля пространственного положения объектов – например в качестве автоколлиматора, так как из выражений (2) и (6) очевидно, что ширина муаровых полос зависит от угла падения волнового фронта. Проведенные расчеты показывают, что чувствительность в этом случае может достигать 1".

Литература

1. Назаров В.Н., Линьков А.Е. Дифракционные методы контроля геометрических параметров и пространственного положения объектов. // Оптический журнал. 2002. Т. 69. № 2. С. 76–81.
2. Diaz – Uribe R., Jimenez – Hernandez A. Phase measurement for segmented optics with 1D diffraction patterns. // Optics Express, 2004. Vol. 12, № 7. P. 1192–1204.
3. Pryor T.R., Hageniers O.L., North W.P.T. Diffractographic dimensional measurement. Part 1: Displacement measurement. // Appl. Opt. 1972. Vol. 11. № 2. P. 308–313.
4. Иванов А.Н., Назаров В.Н. Дифракционный метод контроля геометрических параметров и пространственного положения объектов с помощью «зеркальной» апертуры. Международный оптический конгресс «Оптика – XXI век», VII конференция «Прикладная оптика» // В сборнике трудов конференции. СПб. ГОИ – 2006, т. 1, С. 97–101.
5. Дюрелли А., Паркс В. Анализ деформаций с использованием муара. М.: Мир, 1974. 360 с.
6. Роберт Эрф. Голографические неразрушающие исследования. М.: Машиностроение, 1979. 448 с.
7. J. Der Hovanessian, Y.Y. Hung. Moire contour – sum contour – difference, and vibration analysis of arbitrary objects. // Appl. Opt. 1971. Vol. 10. № 12. P. 2734–2738.

ОДНОКАМЕРНАЯ СКАНИРУЮЩАЯ СТЕРЕОСКОПИЧЕСКАЯ СИСТЕМА ДЛЯ РЕКОНСТРУКЦИИ ТРЕХМЕРНОЙ СТРУКТУРЫ ОБЪЕКТОВ

А.С. Мачихин

(Московский государственный технический университет им. Н.Э. Баумана)

Научные руководители –

д.т.н., профессор В.Я. Колючкин, к.т.н., доцент Л.Н. Тимашова

(Московский государственный технический университет им. Н.Э. Баумана)

Исследована проблема повышения точности реконструкции трехмерной структуры объектов с помощью пассивной стереоскопической системы. Предложено использовать однокамерную сканирующую систему. Рассмотрены основные этапы получения трехмерных образов объектов. Обоснованы преимущества использования предлагаемой системы.

Введение

Восстановление трехмерной структуры объектов (ТСО) является важной задачей технического зрения и машинной графики. Известно несколько подходов к ее решению, возможности применения каждого из которых зависят от предъявляемых к системе требований. В настоящее время наибольший интерес представляют пассивные системы реконструкции ТСО. Отсутствие оптического сканирования и подсветки объекта позволяет значительно расширить возможности применения систем получения и обработки объемных образов в медицинских, охранных, робототехнических и других приложениях.

Оптико-локационный метод, основанный на зондировании пространства лазерным излучением, и параллаксный метод, использующий структурированную подсветку объектов, неприменимы при условии обеспечения пассивности регистрации. Поэтому наиболее широкое распространение получил стереоскопический метод реконструкции ТСО, составляющий основу классической фотограмметрии [1]. Изученность и богатый опыт его практического применения позволили сформировать математический аппарат, дающий возможность успешно решать задачу получения трехмерных образов.

Стереоскопический метод получения информации о дальности до объектов основан на регистрации нескольких изображений сцены под различными ракурсами. В этом случае относительное положение изображений точки сцены на снимках зависит от дальности до точки. Сопоставляя эти изображения, можно при определенных условиях реконструировать трехмерную структуру сцены.

Классическая стереоскопическая система состоит из двух или более жестко закрепленных камер. Точность восстановления ТСО определяется погрешностями определения параметров системы регистрации (СР) и координат сопряженных точек на полученных с разных ракурсов изображениях. К недостаткам стереоскопических систем традиционно относят необходимость высококачественной калибровки СР, сложность алгоритмов обработки изображений для решения задачи стереоотождествления и зависимость точности восстановления ТСО от характера сцены и условий регистрации.

Альтернативой традиционным стереоскопическим системам, состоящим из нескольких СР, могут стать однокамерные сканирующие стереоскопические системы (ОССС). Данная работа посвящена исследованию преимуществ их применения. Предпринята попытка с их помощью повысить точность получаемой ТСО.

Оптическая схема однокамерной сканирующей стереоскопической системы

Для получения стереоэффекта при регистрации изображения одной видеокамерой оптическая система (ОС) должна работать как минимум при двух положениях зрачков

диаметром D_d , симметрично расположенных относительно оптической оси системы на расстоянии b , называемом базой системы (рис. 1). Такая ОС реализуется, например, с помощью линейно перемещающейся вдоль оси x диафрагмы 1 в плоскости входного зрачка объектива 2. Регистрация изображений осуществляется с помощью матричного приемника излучения (МПИ) 3.

Данная схема позволяет в процессе работы изменять длину базы и регистрировать несколько изображений сцен с различных ракурсов.

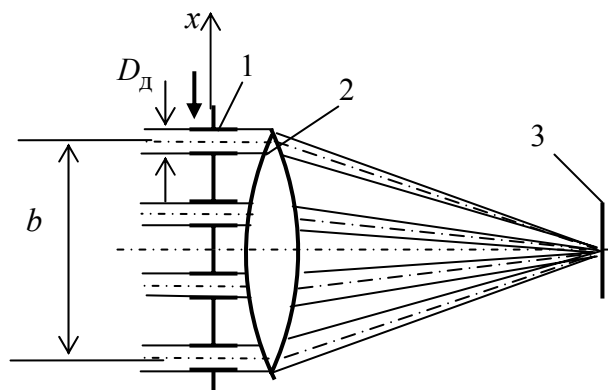


Рис. 1. Оптическая схема однокамерной сканирующей стереоскопической системы

Погрешность δL измерения дальности определяется фокусным расстоянием объектива $f'_{об}$, базой системы b , дальностью до объекта L и размером элемента матрицы $a_{\mathcal{E}}$:

$$\delta L = \frac{a_{\mathcal{E}} \cdot L^2}{b \cdot f'_{об}}. \quad (1)$$

Зная размеры объекта $a_x \times a_y$ и дальность L , определим его угловые размеры

$$\operatorname{tg} W_x = \frac{a_x}{L}, \quad \operatorname{tg} W_y = \frac{a_y}{L}. \quad (2)$$

Зная размеры матрицы $a_{Mx} \times a_{My}$, найдем требуемое фокусное расстояние объектива

$$f'_{об} = \frac{a_{Mx}}{2 \operatorname{tg} W_x} = \frac{a_{My}}{2 \operatorname{tg} W_y}. \quad (3)$$

С учетом (2) и (3) выражение (1) принимает вид

$$\delta L = \frac{a_{\mathcal{E}} \cdot L^2 \cdot 2 \operatorname{tg} W_x}{b \cdot a_{Mx}}. \quad (4)$$

Из (4) можно определить дальность до объекта, при которой для заданных размеров объекта, базе и параметрах матрицы обеспечивается требуемая погрешность измерения дальности. Однако следует отметить, что эта формула не учитывает погрешность определения координат сопряженных точек.

При использовании однокамерной, как и любой другой стереоскопической системы, для оценивания трехмерных координат точки по полученным снимкам необходимо: 1) знать внутренние параметры СР (задача калибровки); 2) знать параметры взаимного ориентирования снимков (задача взаимного ориентирования); 3) знать координаты сопряженных точек, соответствующих одной и той же точке пространства, на снимках (задача поиска сопряженных точек) [2]. Рассмотрим особенности решения каждой из указанных задач при использовании предлагаемой однокамерной системы.

Оценка внутренних параметров

Данная процедура заключается в высокоточном определении параметров СР, состоящей из оптической системы и МПИ: фокусного расстояния, координат главной точки, коэффициентов радиальной дисторсии и двух масштабных коэффициентов. Внутренние параметры СР определены ее конструкцией и могут считаться неизменными. Поэтому их оценка выполняется один раз до начала непосредственного использования системы в лабораторных условиях с применением специальных калибровочных стендов, кодированных меток и т.д. [4, 5]. В настоящее время разработаны методики высококачественного проведения калибровочных экспериментов и автоматической обработки их результатов. Поэтому на практике в подавляющем большинстве случаев удается получить удовлетворительные с точки зрения точности значения внутренних параметров СР.

Основной проблемой на этом этапе являются значительные временные затраты, связанные, в частности, с необходимостью калибровки каждой СР в отдельности. Преимущество использования ОССС заключается в необходимости калибровки лишь одной СР. Это позволяет уменьшить временные и материальные издержки.

Оценка параметров взаимного ориентирования

Для классической стереоскопической системы данная процедура заключается в определении матрицы поворота и вектора трансляции, определяющих относительное положение получаемых изображений. Искомые параметры взаимного ориентирования определяются на основе решения системы уравнений, составленных для набора контрольных точек, координаты которых известны с высокой точностью [1–3].

В некоторых случаях данная задача может быть решена в лабораторных условиях. При этом используются структурированная подсветка, маркирование объекта и другие приемы, позволяющие автоматизировать определение искомых параметров. Однако в большинстве практически важных случаев взаимное положение камер может изменяться в процессе регистрации сцены. Поэтому необходимо его уточнять в ходе работы без использования тестовой сцены и указанных приемов. В таких условиях погрешность нахождения параметров взаимного ориентирования определяется погрешностью рассматриваемого ниже нахождения сопряженных точек.

При использовании ОССС задача взаимного ориентирования сводится к исследованию движения сканирующего элемента. Использование шагового двигателя позволяет производить регистрацию при заранее известных положениях диафрагмы. Поэтому в лабораторных условиях можно измерить необходимые параметры, а в процессе регистрации их уточнить корреляционным алгоритмом с использованием нескольких изображений, полученных при разных положениях диафрагмы.

Поиск сопряженных точек

Данная процедура заключается в определении положения проекции точки трехмерного пространства на одном изображении при заданной проекции той же точки на другом изображении стереопары.

Автоматический поиск сопряженных точек является центральной задачей технического стереозрения. Несмотря на эпиполярное ограничение и ограничение упорядоченности, позволяющие существенно сократить размер зоны поиска, сопоставление окрестностей сопряженных точек не поддается строгой формализации. Возникающие при съемке с разных ракурсов проективные и яркостные искажения зачастую приводят к значительным отличиям изображений одних и тех же фрагментов сцены. Кроме того, в

случае регистрации сцены со слабо выраженной текстурой эффективность работы корреляционных, статистических и других алгоритмов поиска сопряженных точек значительно снижается. При разработке пассивной системы решение задачи стереоотождествления дополнительно осложняется невозможностью использования структурированной подсветки, маркирования и других приемов, позволяющих преодолеть указанные проблемы.

Преимущество ОССС заключается в возможности получения нескольких изображений сцены с разных ракурсов. После задания точки на одном изображении поиск ей сопряженных точек проводится на всех остальных снимках. При этом между собой все точки, являющиеся проекцией одной и той же точки сцены, должны удовлетворять эпиполярным ограничениям. Такой контроль в процессе поиска сопряженной точки позволяет существенно снизить вероятность ложной идентификации фрагментов.

Макетирование однокамерной стереоскопической системы

ОС реализована на базе объектива с фокусным расстоянием $f'_{об} = 200$ мм и диаметром зрачка $D_{вх.зр.} = 100$ мм. Таким образом, база системы примерно равна $b = 100$ мм. Размеры объекта составляют $a_x \times a_y = 200 \times 300$ мм. Дальность до объекта – $L = 1000$ мм. Угловые размеры объекта при этом равны $2W_x \times 2W_y \cong 12^0 \times 16^0$. Стандартная матрица $\frac{1}{2}$ " имеет размеры $a_{Mx} \times a_{My} = 4,8 \times 6,4$ мм². При $a_{э} = 0,003$ мм число элементов равно 1500×2000 .

При таких условиях изображение объекта не вписывается в матрицу СР. Для согласования размеров используется телескопическая система с угловым увеличением $\Gamma = 10^x$. Эквивалентное фокусное расстояние ОС при этом будет равно $f'_{эКВ} = 20$ мм.

Диаметр сканирующей диафрагмы определяется, исходя из требуемого качества изображения: размер пятна рассеяния должен быть не больше размера элемента МПИ. Проведенный абберационный расчет показал, что это условие удовлетворяется при диаметре диафрагмы $D_d = 6$ мм.

Диафрагма установлена на столике с автоматизированным однокоординатным перемещением. Управление перемещением осуществляется с помощью компьютера.

Данная схема согласно (1) обеспечивает погрешность измерения расстояния до объекта $\delta L = 1,5$ мм.

На рис. 2 представлены два снимка объекта, полученные при крайних положениях диафрагмы. Объект содержит ступеньку высотой 50 мм. Крестиками обозначены сопряженные точки объекта.

На рис. 3 приведена реконструированная трехмерная структура этого объекта. По осям абсцисс и ординат отложены номера сопряженных точек, по оси аппликат – вычисленные значения дальности L . Погрешность измерения высоты ступеньки не превышает 3 мм. Отличие погрешности от расчетного значения связано с неточностью установки сканирующей диафрагмы.

В работе были исследованы возможности современных программных средств для оценки внутренних параметров СР и параметров взаимного ориентирования стереоизображений. Можно утверждать, что в ближайшем будущем эти задачи будут успешно решены применительно к ОССС.

В настоящее время ведутся работы по созданию программного обеспечения для решения задачи стереоотождествления с учетом описанных выше особенностей ОССС.

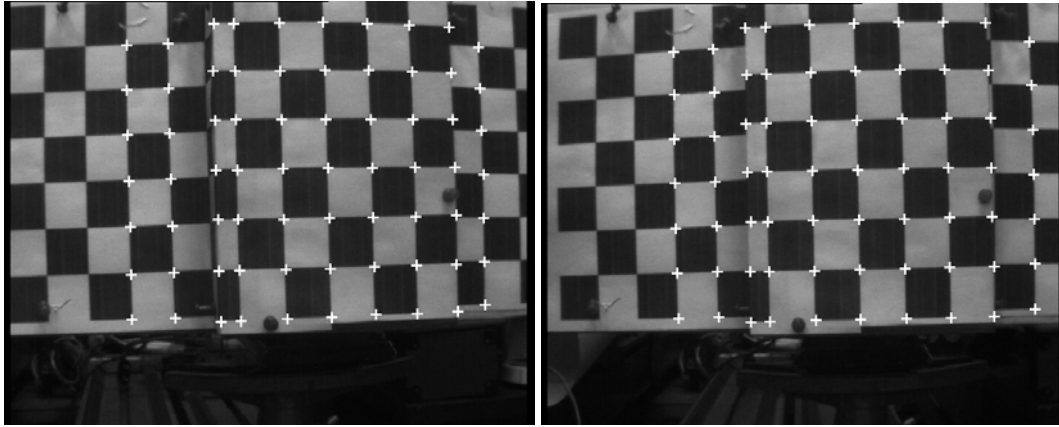


Рис. 2. Стереоснимки, полученные с помощью макета ОССС

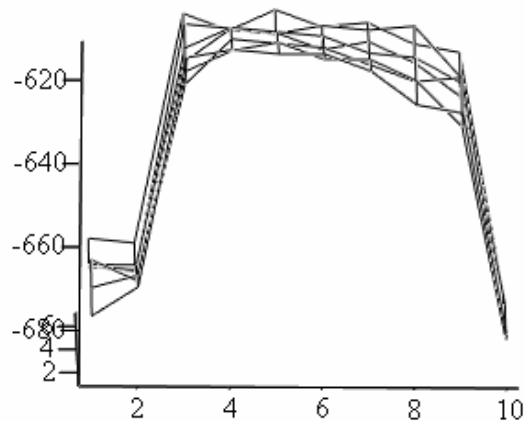


Рис. 3. Реконструированная ТСО

Заключение

На основе проведенных теоретических и экспериментальных исследований предложен вариант стереосистемы, в которой используется одна СР со сканированием в пределах выходного зрачка ОС. По сравнению с классическими пассивными стереоскопическими системами ОССС обладают рядом важных преимуществ, в том числе:

- меньшие габариты;
- упрощенная процедура калибровки;
- более высокая точность восстановления ТСО за счет получения нескольких изображений сцены с разных ракурсов.

Литература

1. Лобанов А.Н. Аэрофототопография. / М.: Недра, 1971. 560 с.
2. Грузман И.С., Киричук В.С. и др. Цифровая обработка изображений в информационных системах. Новосибирск: Изд-во НГТУ, 2002. 352 с.
3. Форсайт Д., Понс Ж. Компьютерное зрение. Современный подход. М: Вильямс, 2004. 928 с.
4. Шапиро Л., Стокман Д. Компьютерное зрение. М: Бином. Лаборатория знаний, 2006. 752 с.
5. Князь В.А., Желтов С.Ю., Салтыкова Е.Б. Автоматическая трехмерная реконструкция наблюдаемой сцены по стереовидеопоследовательностям. // Вестник компьютерных и информационных технологий. 2006. №3. С. 39–43.

ИССЛЕДОВАНИЕ КОНТРОЛЬНОГО ЭЛЕМЕНТА АВТОКОЛЛИМАЦИОННОГО УГЛОМЕРА С ГИПЕРБОЛИЧЕСКОЙ ОТРАЖАЮЩЕЙ ГРАНЬЮ

Син Сянмин

Научный руководитель – д.т.н., профессор И.А. Коняхин

Рассматривается новый вид отражающей системы в виде специального зеркального триэдра, одна из отражающих граней которого выполнена в виде фрагмента гиперболической поверхности. Использование специального триэдра в качестве контрольного элемента при определении углов поворота объектов методом автоколлимации позволяет увеличить диапазон измеряемых углов при упрощении общего алгоритма обработки изображения в плоскости анализа углоизмерительной системы.

Введение

Использование неплоских зеркал для формирования контрольных элементов оптико-электронных автоколлиматоров позволяет реализовать углоизмерительные системы с расширенными метрологическими свойствами в плане повышения точности и увеличения диапазона измерения угловых координат [1].

Структура оптико-электронной углоизмерительной системы (ОЭУС) включает оптико-электронный автоколлиматор 1, установленный на некотором базовом объекте, а также отражающую оптическую систему – контрольный элемент 2, связанный с объектом, повороты которого контролируются (на рис. 1).

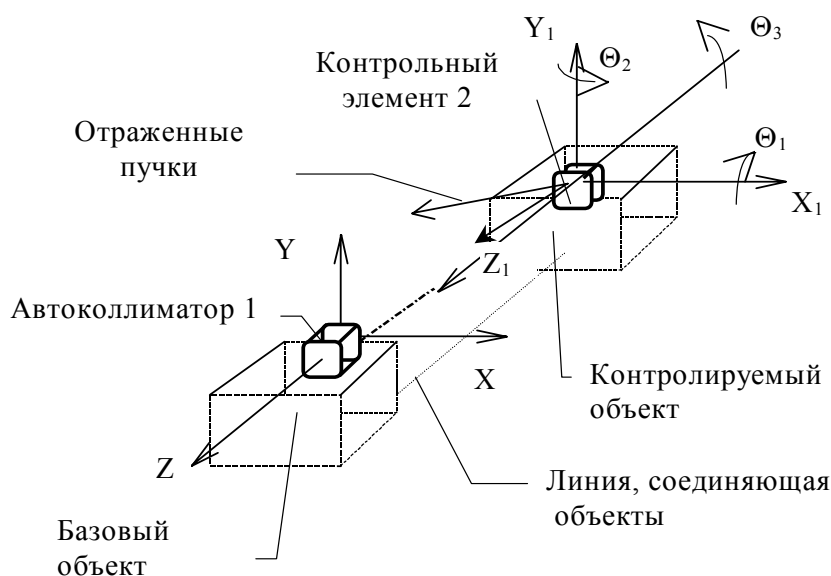


Рис. 1. Структура автоколлимационной углоизмерительной оптико-электронной системы

Автоколлиматор 1 включает излучающий канал, формирующий пучок, падающий на контрольный элемент 2, и приемный канал, регистрирующий отраженный пучок. Контрольный элемент (КЭ) 2 представляет собой отражающую оптическую систему, которая формирует отраженный пучок с требуемыми свойствами.

Если с базовым объектом связать неподвижную систему координат XYZ , а с контролируемым объектом – систему $X_1Y_1Z_1$, угловое пространственное положение контролируемого объекта определяется тремя угловыми координатами – углами Θ_1 , Θ_2 , Θ_3 поворота системы координат $X_1Y_1Z_1$ относительно осей системы координат XYZ .

Рассмотрим задачу измерения так называемых коллимационных угловых коорди-

нат – угла Θ_1 или Θ_2 . Типичным КЭ для автоколлимационных измерений является плоское зеркало с коэффициентом преобразования «угол поворота объекта – угол отклонения отраженного пучка», равным $K = 2$. Однако при его использовании рабочая дистанция угломера не превышает 2–3 метров вследствие значительного виньетирования приемным объективом автоколлиматора отраженного пучка даже при малом (единицы угловых минут) диапазоне измерения [2].

При использовании КЭ в виде зеркального триэдра с неплоской отражающей гранью автоколлимационные измерения реализуются на значительных дистанциях без существенного виньетирования, что позволяет увеличить диапазон измерения.

Принципы построения и параметры триэдрической отражающей системы

Рассмотрим КЭ в виде зеркального триэдра, одна из отражающих граней которого – например, грань 3 выполнена в виде фрагмента поверхности второго порядка, ограниченного линиями пересечения этой поверхности с гранями 1 и 2. Две другие грани – 1, 2 – плоские и совпадают с плоскостями координатного трёхгранника $X_0Y_0Z_0$ исходной системы координат (на рис. 2). Начало исходной системы координат совпадает с точкой пересечения трех граней.

Пусть предварительно грань выполнена в виде фрагмента цилиндрической поверхности, например, кругового цилиндра с малой кривизной. Ориентация цилиндрической поверхности такова, что ее образующая перпендикулярна ребру двугранного угла, образованного двумя плоскими гранями 1, 2 и, соответственно оси OZ_0 исходной системы координат. Ось исходного цилиндра и координатная ось OZ_0 имеют точку пересечения. В этом случае орт нормали N_3 к цилиндрической грани определяется двумя параметрами: углом α , задающим положение плоскости направляющей поверхности и углом β , в этой плоскости между осью OZ_0 и ортом нормали, который изменяется в интервале $0 - \beta_{\max}$, где β_{\max} – малый угол.

Изображение, формируемое пучком, отраженным от рассмотренного отражателя в плоскости анализа приемного канала автоколлиматора представляет собой часть эллиптической дуги, описываемой уравнениями [1]:

$$X_0(\varepsilon, \beta) = \frac{\sqrt{6}}{3} \sin(\varepsilon) \cdot \cos(\varepsilon) (1 - \cos(2\beta)) - \frac{\sqrt{3}}{3} \cos(\varepsilon) \sin(2\beta) \quad (1)$$

$$Y_0(\varepsilon, \beta) = \frac{-\sqrt{2}}{3} \cos(\varepsilon)^2 (1 - \cos(2\beta)) - \sin(\varepsilon) \sin(2\beta), \quad (2)$$

где угол ε определяется как $\varepsilon = \frac{\pi}{4} + \alpha$. При повороте отражателя на коллимационный угол Θ_2 относительно оси OY_1 приборной системы координат (на рис. 3) эллиптическая дуга поворачивается на угол Ξ относительно центра плоскости анализа. Измеряемый угол Θ_2 определяется по алгоритму:

$$\Theta_2 = \frac{\Xi}{K}, \quad (3)$$

где K – коэффициент преобразования, равный

$$K = \frac{\sqrt{2} \cdot \cos(\varepsilon)}{\sqrt{2 \cdot \sin(\varepsilon)^2 + 1}}. \quad (4)$$

Поскольку изображение поворачивается без линейного смещения центра, по-

грешность вследствие виньетирования пучка отсутствует. Однако экспериментальные исследования описанного КЭ элемента выявили принципиальную сложность алгоритма обработки изображения. В частности, для определения угла поворота Ξ дуги необходимо предварительно измерить координат нескольких точек дуги с последующей итерационной аппроксимацией «ближайшим» эллипсом.

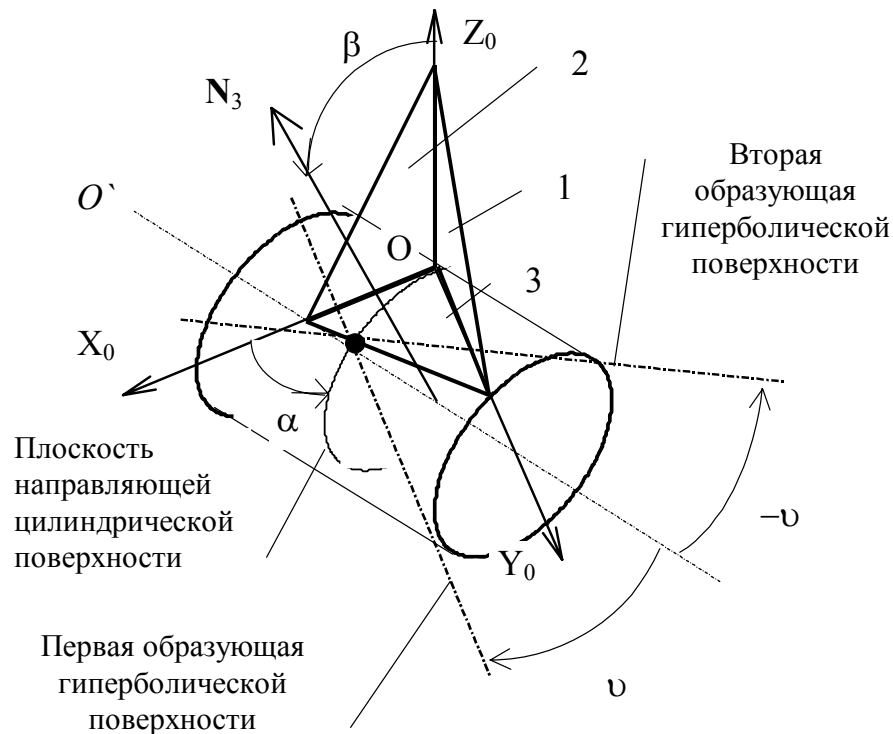


Рис. 2. Конфигурация триэдрического КЭ

Для упрощения алгоритма измерения исходный цилиндр предлагается «деформировать» до однополостного гиперboloида. При этом каждой точке грани 3 будет соответствовать две образующие, составляющие с образующей исходного цилиндра углы $\pm \nu$ [3]. В случае малости ν можно полагать, что отражающие свойства гиперболической поверхности аналогичны цилиндрической, каждой точке которой соответствует две направляющие, плоскости которых составляют с осью OX_0 углы $\alpha \pm \vartheta$, соответственно. Тогда выражения (1)–(4), описывающие действие КЭ, будут справедливы при угле ε , изменяющемся в интервале:

$$\varepsilon \in \left[\frac{\pi}{4} + \alpha - \nu, \frac{\pi}{4} + \alpha + \nu \right]. \quad (5)$$

Анализ показывает, что изображение в плоскости анализа автоколлиматора при использовании гиперболической грани приобретает тонкую структуру, соответствующую расщепленной эллиптической дуге, формируемой триэдром с цилиндрической гранью. Изображение в плоскости анализа для значения параметра $\alpha = \frac{\pi}{4}$, построенное по выражениям (1)–(2) с учетом (5), приведено на рис. 3. Из выражения (4) следует, что коэффициент преобразования K для граничных областей изображения при малой гиперболичности равен $K = \pm \nu$. Тогда, в соответствии с выражением (3), границы изображения при повороте КЭ будут разворачиваться в противоположных направлениях, что приведет к утолщению (сужению) толщины исходной дуги.

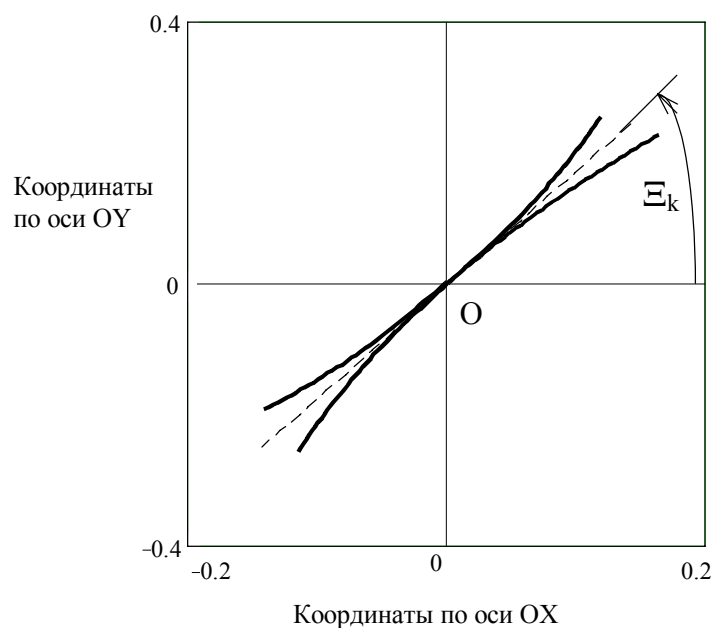


Рис. 3. Результат моделирования изображения, формируемого триэдрическим КЭ

Следовательно, для измерения угла поворота достаточно выполнить измерение поперечной ширины дуги, что реализуется нетрудоемким алгоритмом.

Заключение

Рассмотрены принципы построения контрольного элемента для автоколлимационных измерений, использование которого позволяет при несложном алгоритме обработки изображения определять величину коллимационного угла поворота объекта на дистанциях до десятков метров.

Указанные метрологические свойства реализуются при трансформации исходной цилиндрической формы грани исходного зеркального триэдра в гиперболическую.

Моделирование алгоритма работы КЭ показывает, что использование гиперболической грани позволяет значительно упростить алгоритм автоколлимационных угловых измерений сравнительно с применением цилиндрического контрольного элемента. Это следует из того, что для определения углов поворота отражателя (и, соответственно, скрепленного с ним контролируемого объекта) достаточно измерить разницу в поперечном размере эллиптических дуг изображения без определения параметров формы каждого эллипса.

Литература

1. Джабиев А.Н., Коняхин И.А., Панков Э.Д. Автоколлимационные углоизмерительные средства мониторинга деформаций. СПб: СПбГИТМО (ТУ), 2000. 197 с.
2. Аникст Д.А., Константинович К.М., Меськин И.В., Панков Э.Д. Высокоточные угловые измерения / Под ред. Ю.Г. Якушенкова, М.: Машиностроение, 1987. 480 с.
3. Корн Г., Корн Т., Справочник по математике для научных работников. Пер. с англ. М.: Наука, 1984.

АДАПТИВНЫЙ МНОГОПРОХОДНЫЙ АЛГОРИТМ ОПРЕДЕЛЕНИЯ КООРДИНАТ ИЗОБРАЖЕНИЙ ТОЧЕЧНЫХ ИЗЛУЧАТЕЛЕЙ

Д.В. Жуков, А.А. Усик

Научный руководитель – д.т.н., профессор И.А. Коняхин

Предлагается многопроходный алгоритм определения координат изображения точечного излучателя, основанный на последовательном уточнении области локализации изображения. Проводится исследование алгоритма на математической модели. Выявляется адаптивность алгоритма к размеру изображения, а также независимость от первоначальной оценки области локализации изображения.

Введение

Для реализации обширного круга измерительных задач в области научной и промышленной техники все чаще используются оптико-электронные системы (ОЭС), основанные на матричных приемниках оптического излучения (ПОИ). Системы определения пространственных координат объектов, их перемещения и деформации могут быть крайне эффективны, при реализации в них наиболее оптимального алгоритма обработки изображения, снимаемого с фоточувствительных элементов матрицы. Одним из важных факторов повышения точности таких систем при регистрации как визирных целей – светодиодов, так и, например, естественных источников излучения – звезд, является получение оценки определения координат изображений источников с субпиксельной точностью.

Ранее были получены результаты [2], показавшие сильную зависимость точности определения координат центра изображения точечного излучателя (ИТИ) по алгоритмам взвешенного суммирования (ВС) от первоначального определения геометрических параметров области локализации изображения (окна): координат центра и полуширины. Определение координат центра окна может давать существенную ошибку в условиях засветки или при отклонении распределения облученности в плоскости изображения от гауссовой формы.

Нами была поставлена задача поиска алгоритма, в котором устранялись бы вышеупомянутые зависимости. Мы предъявляем к нему жесткие требования по ресурсоемкости (времени исполнения) с целью возможности использования в широком классе ОЭС.

Описание алгоритма

Для решения поставленной выше задачи обратимся к алгоритмам ВС, но при этом попытаемся как можно точнее оценить параметры окна.

Положим, что изображение, формируемое на приемнике, имеет распределение в форме функции Гаусса:

$$Q(x, y) = A \exp\left(-\frac{(x-a)^2 + (y-b)^2}{2\sigma^2}\right).$$

Тогда суммарную облученность можно найти как интеграл:

$$M = \iint_{x, y \in (-\infty; \infty)} Q(x, y) dx dy = 2\pi\sigma^2 A.$$

Отсюда можно выразить параметр σ , характеризующий ширину изображения:

$$\sigma = \sqrt{\frac{M}{2\pi A}}.$$

Основная «масса» изображения содержится в круге с центром в (a, b) и радиусом $R = 3\sigma$. Этот факт позволяет нам точнее определить размер окна в алгоритме ВС и тем самым уменьшить ошибку, вносимую фоновым шумом.

Зададимся некоторым окном, которое, предположительно, должно содержать изображение:

$$W_k = \{(i, j) \in F \mid \tilde{a}_k - \tilde{R}_k \leq x_i \leq \tilde{a}_k + \tilde{R}_k, \tilde{b}_k - \tilde{R}_k \leq y_j \leq \tilde{b}_k + \tilde{R}_k\},$$

где F – множество пикселей кадра, (x_i, y_j) – координаты пикселя (i, j) на матричном приемнике, $(\tilde{a}_k, \tilde{b}_k)$ – координаты центра окна, \tilde{R}_k – полуширина окна.

В данном окне можно провести оценки параметров изображения σ , a и b :

- $\tilde{\sigma}_{k+1} = \sqrt{\frac{\tilde{M}_{k+1}}{2\pi\tilde{A}_{k+1}}}$, где $\tilde{A}_{k+1} = \max_{(i,j) \in W_k} Q_{ij}$ и $\tilde{M}_{k+1} = \sum_{(i,j) \in W_k} Q_{ij}$;
- $\tilde{a}_{k+1} = \frac{\sum_{(i,j) \in W_k} Q_{ij}^2 x_i}{\sum_{(i,j) \in W_k} Q_{ij}^2}$;
- $\tilde{b}_{k+1} = \frac{\sum_{(i,j) \in W_0} Q_{ij}^2 y_i}{\sum_{(i,j) \in W_k} Q_{ij}^2}$.

Из полученных значений можно сформировать новое окно W_{k+1} с параметрами \tilde{a}_{k+1} , \tilde{b}_{k+1} и $\tilde{R}_{k+1} = 3\tilde{\sigma}_{k+1}$. Задав параметры первоначального окна W_0 (\tilde{a}_0 , \tilde{b}_0 , \tilde{R}_0) и, применив к нему вышеописанную рекурсивную процедуру, получим искомые оценки координат центра изображения: $\tilde{a} = \tilde{a}_n$, $\tilde{b} = \tilde{b}_n$. Здесь мы применяем алгоритм взвешенного квадратичного суммирования (ВКС) для определения центра изображения в данном окне, так как в большинстве случаев он дает лучший результат, чем классическое взвешенное суммирование [1–3].

Структура математической модели

Модель реализована для размеров изображения порядка 6–48 пикселей при отношении сигнал/шум (ОСШ) в пределах 10–500. Под ОСШ понимается отношение квадрата максимального значения сигнала к дисперсии шума. Задача первоначального определения области локализации изображения полагается решенной. Полуширина окна \tilde{R}_0 полагалась равной 15 пикселей. В моделировании зависимости ошибки определения центра от ОСШ центр изображения полагался смещенным относительно центра окна на 0,3 пикселя по горизонтали и вертикали (рис. 1). Также полагалось, что распределение облученности в изображении имеет форму функции Гаусса с наложенным на него аддитивным гауссовым шумом.

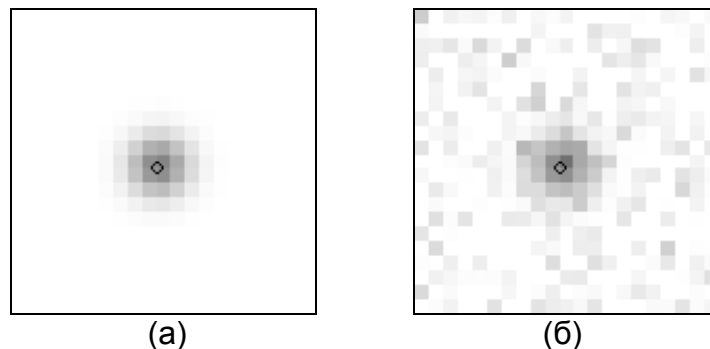


Рис. 1. Изображение точечного излучателя (а) в отсутствие шума, (б) в присутствии шума

Для построения модели использовалось программирование на языке С.

Результаты

Для выяснения эффективности работы алгоритма были проведены 2 серии экспериментов.

1. Зависимость точности определения координат центра изображения от размера изображения при фиксированном ОСШ равном 200 *отн.ед.*
2. Зависимость точности определения координат центра изображения от ОСШ для изображения с параметром σ равным 1,5.

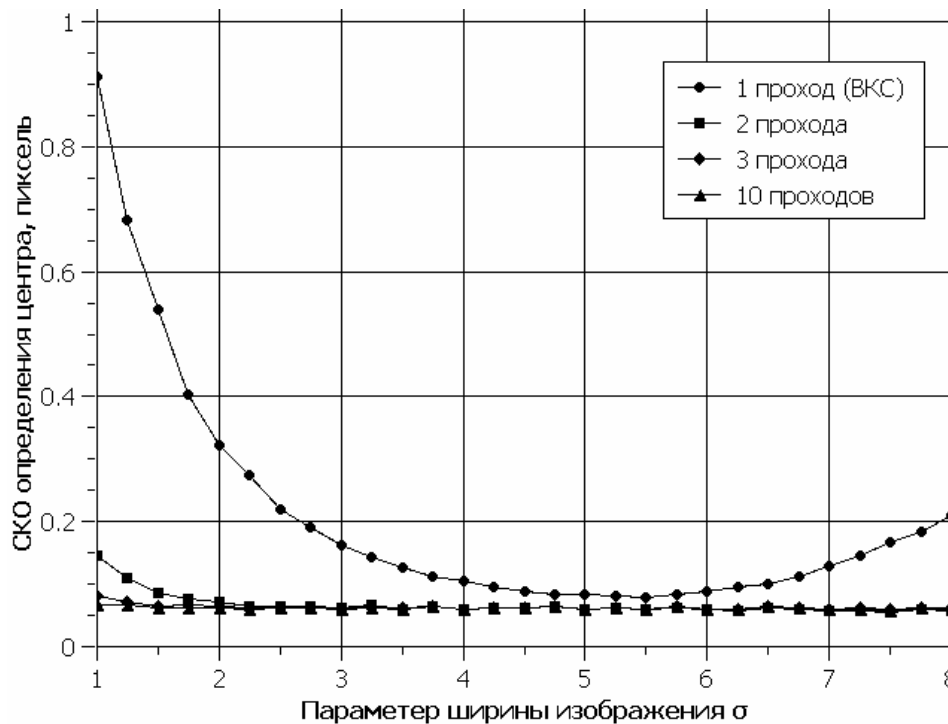


Рис. 2. Эмпирические зависимости точности определения координат центра от размера изображения при ОСШ = 200

Каждый эксперимент повторялся 250 раз при различном количестве проходов алгоритма, после чего вычислялось среднее квадратическое значение отклонения координат. По полученным данным были построены графики (рис. 2–3).

Математическое моделирование показало, что предложенный нами итерационный алгоритм быстро сходится (в большинстве случаев за 3 итерации), что находит свое отражение в построенных графиках. Уже после небольшого количества проходов точность определения координат достигает своего предела. Из рис. 2 видно, что первый проход алгоритма, представляющий собой фактически алгоритм взвешенного квадратического суммирования (ВКС), имеет сильную зависимость от размеров изображения и достигает максимальной точности только при полуширине окна, равной эффективно-му радиусу изображения, но уже второй проход алгоритма практически устраняет зависимость точности определения координат от размера изображения. Очень важным результатом стал факт адаптации алгоритма под размер изображения даже при задании начального окна с полушириной меньшей, чем эффективный радиус изображения. Более того, моделирование показало, что нецелесообразно задавать размеры исходного окна заведомо больше, чем размер изображения. Наиболее эффективным будет принять размер окна как среднее всех возможных значений размеров изображений. Зависимость ошибки определения центра от положения первоначального окна имеет тот же характер, что и зависимость от размеров. Трех итераций в данном случае достаточно для практически полного устранения данной зависимости. Шумовая зависимость (рис. 3)

показывает, что дополнительные проходы алгоритма способны существенно уточнить измеряемые координаты на всем диапазоне значений ОСШ. Здесь, как и следовало ожидать, точность практически перестает расти после 3-ей итерации.

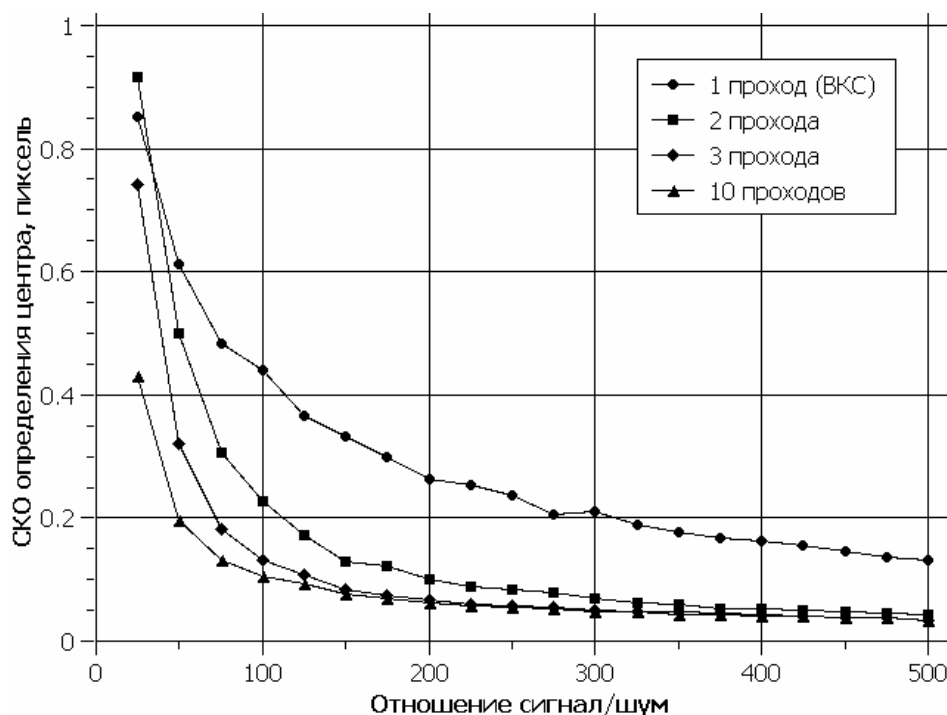


Рис. 3. Эмпирические зависимости точности определения координат центра от уровня шума при $\sigma = 1,5$

Заключение

1. Численные исследования показывают существенную эффективность представленного алгоритма по сравнению с ВКС, при этом позволяя варьировать его ресурсоемкость и эффективность в зависимости от числа проходов.
2. Адаптивный многопроходный алгоритм на основе взвешенного квадратичного суммирования практически не зависит от первоначальной оценки области локализации изображения на кадре и способен адаптивно подстраиваться под размер изображения.
3. Являясь алгоритмом на основе ВС, он гораздо менее ресурсоемок, чем аппроксимация функцией Гаусса по методу наименьших квадратов [2], и может быть легко внедрен в оптико-электронные измерительные системы, с минимальной их модификацией.

Литература

1. Данилов Д.В., Пашков В.С. Оценка координат изображений точечных излучателей. / Оптико-электронные приборы и системы. Сб. научных статей. Вып. 96./ Под ред. Э.Д. Панкова. СПб, 1996. С. 29–33.
2. Жуков Д.В., Усик А.А., Коняхин И.А. Аналитический обзор способов определения координат изображений точечных источников. // XXXVI научная и учебно-методическая конференция профессорско-преподавательского и научного состава. 30 января – 2 февраля 2007 года: Программа. СПбГУ ИТМО, 2007.
3. Пантюшин А.В., Жуков Д.В., Шомрина М.А. Оценка нелинейности статической характеристики при определении смещений с помощью матричных фотоприемников на основе КМОП структур. // XXXVI научная и учебно-методическая конференция

профессорско-преподавательского и научного состава. 30 января – 2 февраля 2007 года: Программа. СПбГУ ИТМО, 2007.

СРАВНЕНИЕ НЕЙРОСЕТЕВЫХ АЛГОРИТМОВ ПРИ ОПТИЧЕСКИХ МЕТОДАХ КОРРЕКЦИИ СЕГМЕНТОВ ОТРАЖАЮЩИХ ПОВЕРХНОСТЕЙ РАДИОТЕЛЕСКОПОВ

П.В. Кужаков

Научный руководитель – к.т.н., доцент А.Л. Андреев

Проводится сравнительный анализ работы алгоритмов на базе нейронных сетей в оптико-электронных системах. Отмечена возможность использования алгоритма обучения, позволяющего улучшить быстродействие системы в рабочем режиме за счет математической обработки результатов измерений.

Введение

Типовой канал измерения пространственных координат предназначен для контроля деформаций элементов радиотелескопа для миллиметрового диапазона длин волн. Радиотелескопы данного класса являются новейшими научными инструментами и требуют высокоточного контроля формы поверхности зеркала. Основное зеркало создаваемого Российской АН радиотелескопа РТ-70 (Суффа) имеет диаметр 70 метров, форма зеркала – трехмерная парабола. Поверхность зеркала состоит из 2000 плоских металлических секций. Работа в миллиметровом диапазоне длин волн требует минимального (не больше чем 0,1 мм) отклонения поверхности от теоретической трехмерной параболы.

На радиотелескоп оказываются различные воздействия, среди которых климатические, что вызывает смещение статически неопределимой конструкции большого количества ЩОП. Для решения задачи восстановления системы возможно использование нейронных сетей для различных задач, где требуется обеспечение управлением коррекции сложной поверхности. Нейронные сети усваивают и реализуют сложные алгоритмы управления (которые нельзя получить аналитически). Во многих случаях системы, управляемые нейронными сетями, могут быть построены при единственном допущении, что для структуры, выполняющей управление, существует алгоритм эффективного управления.

Основная часть

В настоящее время имеется большое разнообразие в построении зеркальных систем (ЗС) радиотелескопов, анализ и опыт использования которых для наблюдения космических источников радионаблюдения (КИР) миллиметровых длин волн (ММДВ) показывает, что наиболее перспективно использование больших полноповоротных наземных трех зеркальных радиотелескопов с третьим подвижным перископическим зеркалом (ПЗ).

Для повышения точности и эффективности в системе наведения такой зеркальной системы (ЗС) на КИР может использоваться сканирование принимаемого сигнала через точечный приемник с помощью перископического зеркала и адаптация поверхности основного зеркала с помощью управляемого перемещения щитов отражающей поверхности (ЩОП). Вместо сканирования принимаемого сигнала перископическим зеркалом можно использовать управляемый подвижный приемник (УПП), устанавливаемый вблизи исходного вторичного фокуса ЗС. Тогда приемник будет следить за уходом вторичного фокуса ЗС. При этом значительного снижения требований к точности наведения можно получить при использовании матричного приемника. На этапе проектирования расчетным и экспериментальным путями определяются оптимальные положения ПЗ или УПП при различных сочетаниях ошибок наведения и деформаций ЗС. Эти данные записываются в память вычислительно-управляющего комплекса (ВУК) и исполь-

зуются в процессе наведения ПЗ или УПП для формирования задающих сигналов на их приводы, на основании текущих значений измеряемого вектора состояния РТ. При таком способе приема сигналов нет необходимости требовать от приводов наведения точности соизмеримой с шириной диаграммы направленности антенны (ДНА). Требования к приводам наведения по точности могут быть существенно снижены, а разрешающая способность РТ повышена.

Одним из возможных способов решения задачи обеспечения управлением и коррекцией сложной поверхности, состоящей из статически неопределимой конструкции большого количества ЩОП, является применение алгоритма обучения по схеме приведенной на рис. 1.

Схема обучения, приведенная на рис. 1, позволяет обойти необходимость калибровки системы после каждого снятия отсчета координат объекта, и использовать алгоритм обучения [1–5].

Суммарная погрешность определения пространственного положения точек на поверхности каждого ЩОП не должна превышать 0,25 мм при относительно малом времени измерения. По причине сложности реализации системы измерения с требуемым быстродействием предлагается измерение положения вершины и оси ОЗ разделить на два этапа – предварительный статический этап обучения (зенитальный) и динамический этап (непосредственная оценка).

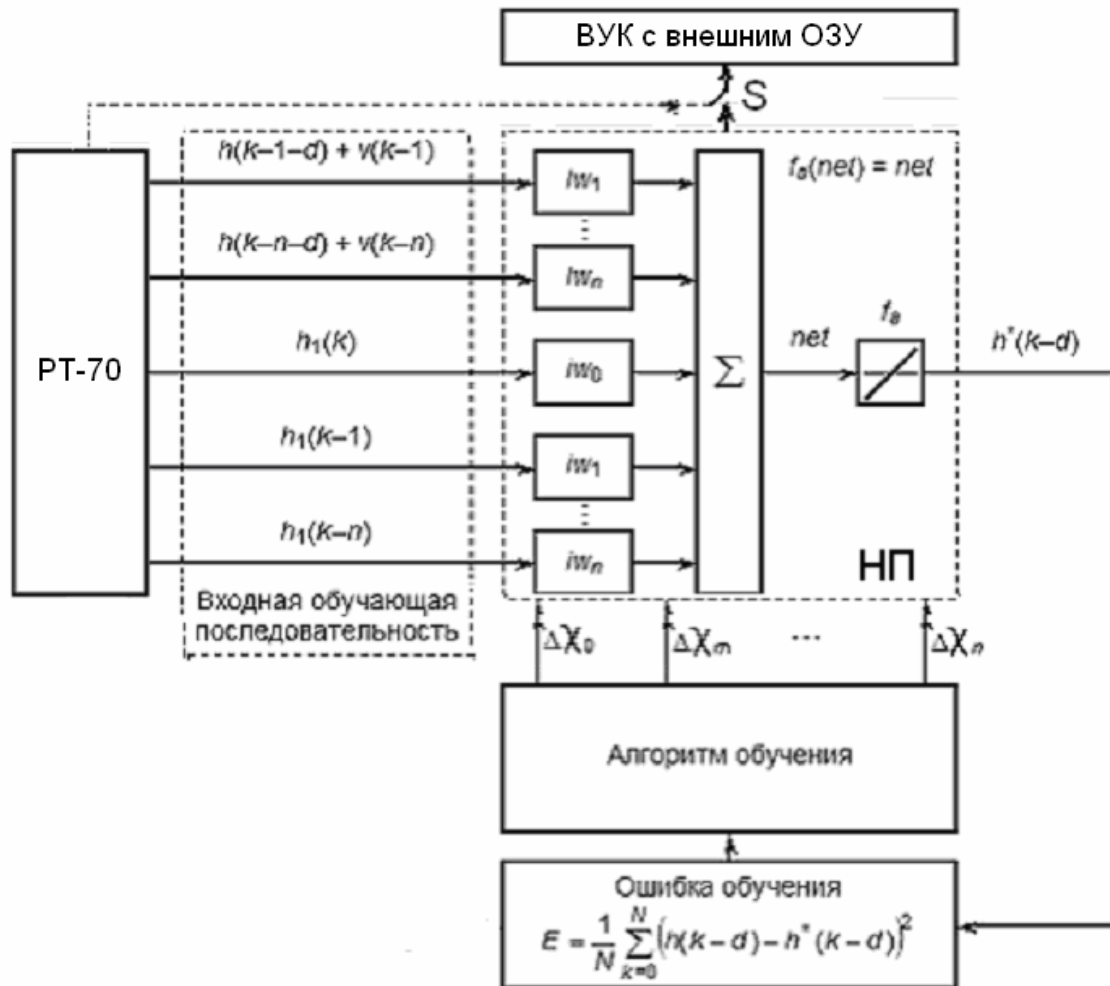


Рис. 1. Схема обучения системы в статическом режиме

На первом этапе обучения ОЗ переводится в зенитальное положение, и в течение длительного времени выполняется измерение положения ЩОП сначала с помощью универсальных измерительных средств – автоматических тахеометров – программным

наведением на произвольно выбранную и контролируруемую точку со средней погрешностью 0,25 мм. На рис. 2 в качестве примера приведена статическая характеристика канала измерения положения актуаторов χ одного из отражающих щитов. По оси абсцисс отложены истинные значения определённого значения положения актуаторов χ , а по оси ординат – поля числа отклонений $\tilde{\chi}$. Пунктиром показана идеальная статическая характеристика системы. Затем с помощью контактных измерительных средств, например, оптических мерных жезлов, положение уточняется с конечной погрешностью порядка 0,05 мм. Полученные данные о положении ЩОП заносятся в компьютерную модель, выполняющую расчет параметров ОЗ.

На втором этапе ОЗ находится в рабочем положении. С помощью специализированной измерительной системы измеряются пространственные координаты ограниченного числа (30–40) контрольных точек $\chi_m, \chi_n, \dots, \chi_q$ на поверхности зеркала с требуемой точностью и быстродействием 0,02 с (при разработке специальной системы снятия сигнала – 0,002 с). При этом погрешность установки ЩОП должна быть значительно меньше величины систематической погрешности измерения.

По измеренным данным выполняется коррекция расчетной модели параметров ОЗ для текущего азимутального и угломестного положения. Для контроля точности установки углов можно воспользоваться тахеометрами.

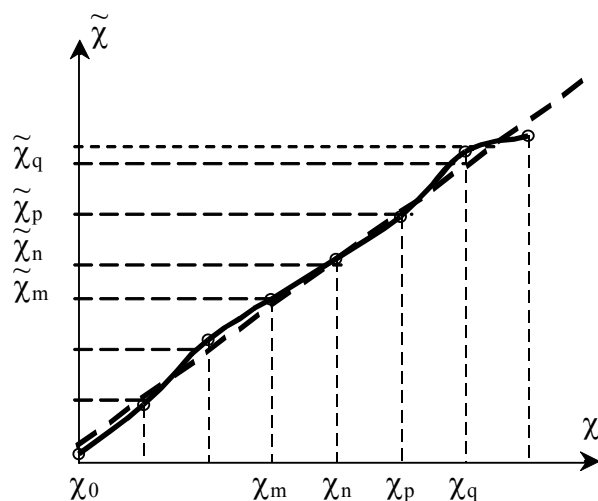


Рис. 2. Статическая характеристика канала измерения положения актуаторов χ

На этапе калибровки переключатель S , должен быть переведен в нижнее положение.

Предлагается следующая схема управления системой коррекции сегментов отражающих поверхностей радиотелескопов:

1. каждый из трех измерителей находит расстояния до $N \sim 1200$ точек на поверхности зеркала и до других измерителей;
2. на ВУК вычисляются координаты этих N точек в системе, связанной с измерителями;
3. в памяти ВУК находятся 6 параметров предыдущих параметров параболоида – «старого» идеального параболоида (ИП) от предыдущей коррекции. Т.е для проведения самой первой коррекции параметры «старого» параболоида принимаются равными параметрам, померенным радиоголографическим методом;
4. в памяти ВУК находятся также старые ($N \sim 1200$) значения положения δ_i ($i=1, \dots, 1200$) актуаторов, которые они имели после окончания предыдущего цикла коррекции;
5. ВУК вычисляет поле ($N \sim 1200$) чисел – отклонений Δi , $i=1, \dots, 1200$ измеренных точек зеркала вдоль направления оси старого (ИП) до поверхности старого параболоида (ИП);
6. если каждое из отклонений мало, $\Delta < 50$ мкм, то коррекция не производится за ненадобностью;

7. если отклонения по совокупности не могут считаться малыми, тогда ВУК, которые вычисляет 6 параметров нового ИП и 1200 чисел – перемещений актуаторов, необходимых для корректировки каждой из $N \sim 1200$ деформаций выведения $N \sim 1200$ точек поверхности зеркала в положение нового ИП;
8. дается команда на актуаторы;
9. дается команда на двигатели контррефлектора для перевода его в положение нового фокуса и приемника для помещения его на новое направление оптической оси;
10. запоминаются в памяти ВУК 6 новых параметров ИП, которые для последующей корректировки будут являться параметрами старого ИП;
11. запоминаются в памяти ВУК новые (текущие) положения актуаторов;
12. через $t=10^m$ цикл коррекции повторить.

При задании определённого значения положения актуаторов χ_i на выходе системы появляется двоичный параллельный код соответствующей оценки χ_i^* , который посредством устройства ввода/вывода (УВВ) ВУК заносится в его собственную память. Для уменьшения на этапе калибровки влияния случайной составляющей погрешности измерения, при каждом положении актуаторов в ВУК обеспечивается ввод не однократного значения измеренных точек зеркала χ_i^* , а целый массив значений, состоящий из 1200 значений. Этот массив в дальнейшем обрабатывается ВУК с целью вычисления поля числа отклонений $\bar{\chi}$. Значения $\bar{\chi}$ используется в дальнейшем как адрес той ячейки памяти, в которую должно быть записано истинное значение измеряемого угла установленное ВУК. Это же значение угла отображается на ВУК, и затем записывается во внешнее ОЗУ по адресу поля числа отклонений.

Аналогичным образом формируются все адреса поля числа отклонений и данные измеряемых точек зеркала записываемых в другие ячейки ОЗУ (рис. 3).

На следующем этапе ВУК осуществляет интерпретацию значений новых параметров ИП, которые для последующей корректировки будут являться параметрами старого ИП. Все рассчитанные значения $\chi_1, \chi_2, \dots, \chi_{m-1}$ и др. записываются в промежуточные, пока еще свободные ячейки внешнего ОЗУ.

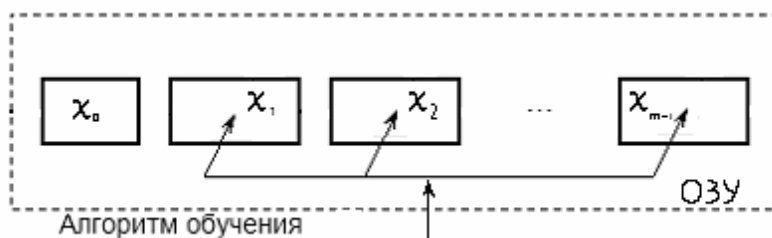


Рис. 3. Схема формирования массива данных во внешнем ОЗУ на этапе обучения

Нелинейная статическая характеристика рис. 2 аппроксимируется кусочно-линейной функцией с различной крутизной наклона на различных участках. Эта функция в виде таблицы-массива хранится во внешнем ОЗУ измерительной системы. Число калибруемых точек определяется степенью нелинейности исходной статической характеристики системы и требуемой точностью коррекции систематической погрешности.

После завершения цикла обучения ключ S переводится в верхнее положение, и система подготавливается к режиму измерения. Теперь двоичный код каждого однократного значения измеренных точек зеркала χ_i^* , формируемый системой, поступает непосредственно на вход дешифратора адреса внешнего ОЗУ, и из соответствующей ячейки памяти сразу считывается ошибка знания положения коррекции актуаторов, заранее вычисленная на этапе калибровки.

Для иллюстрации возможностей предложенной модели и алгоритма было проведено цифровое моделирование в среде *MATLAB*. В качестве примера был рассмотрена система, имеющая передаточную функцию следующего вида:

$$W(s)=1/(s+1)^a. \quad (1)$$

На входе системы описываемая передаточная функция (1) с коэффициентом ней-роуправления постоянной времени $T = 0,002$ с и коэффициентом демпфирования $a = 2$ и комплексной переменной s . Обучение ОЭС проводилось согласно схеме, изображенной на рис. 1. В качестве алгоритма обучения ОЭС в статическом режиме был использован алгоритм обратного распространения ошибки – метод Левенберга-Марквардта [6].

Для обучения нейронных сетей использовалась одна из разновидностей алгоритма обратного распространения ошибки – метод Левенберга-Марквардта (LM) [7]. Этот метод относится к методам Ньютона, которые являются примером способа быстрой оптимизации обучения. Основным выражением методов Ньютона является выражение для шага:

$$x_{k+1} = x_k - H_k^{-1} g_k, \quad (2)$$

где H_k – матрица Гессе вторых частных производных функционала ошибки по настраиваемым параметрам; g_k – вектор градиента функционала ошибки. Процедура минимизации на основе метода Ньютона, как правило, сводятся быстрее, чем те же процедуры на основе метода сопряженных градиентов. Суть этих методов заключается в том, что на каждом шаге приращения вычисляется аппроксимирующая матрица Гессе, и само приращение вычисляется как функция градиента. Этот метод более быстрый, чем метод градиентного спуска (табл. 1). Однако для случая нейронных сетей прямого распространения он требует больших вычислений. Главный недостаток алгоритма LM – в том, что требуется значительный объем памяти для хранения матриц больших размеров. Например, матрица Якоби составляет $Q \times n$, где Q – число обучаемых наборов и n – число параметров сети. Это означает, что при оценке гессиана потребуются значительные ресурсы для ее хранения. Поэтому матрица разбивается на подматрицы меньших размеров. Причем в процессе формирования матрицы Гессе использованные подматрицы могут быть удалены из оперативной памяти. Параметр *mem_reduc* определяет, на какое количество подматриц разбивается исходная матрица. Если параметр *mem_reduc* равен 1, подматрицы делятся по строке на две части и сначала обрабатывается одна половина, а затем вторая. Это экономит половину объема памяти, требуемой для вычисления полного якобиана. И если нам доступна достаточная оперативная память, то надо установить параметр *mem_reduc* равным 1. Если память оказалась исчерпанной, следует назначить параметр *mem_reduc* равным 2 и выполнить расчеты заново. Если при этом памяти не будет достаточно, следует увеличить значение этого параметра.

Алгоритм обучения	Значение фиксированной максимальной точности (ошибки), E					
	E=10 ⁻¹		E=10 ⁻³		E=10 ⁻⁵	
	Время обучения, мс	Число циклов обучения	Время обучения, мс	Число циклов обучения	Время обучения, мс	Число циклов обучения
Градиентного спуска (ГС)	42,52	80	данная точность не была установлена	данная точность не была установлена	данная точность не была установлена	данная точность не была установлена
Левенберга-Марквардта	5,1	4	10,63	9	14,27	13

Таблица 1. Сравнительная характеристика алгоритмов обучения ИП

Критерий выбора параметров ИП зависит от компромисса между различными требованиями:

- 1) экономии хода актуатора (ресурса работы двигателей);
- 2) экономии времени настройки (быстрота работы);

- 3) быстрого времени вычисления параметров ИП (ВУК);
- 4) максимального использования площади зеркала (время наблюдения на телескопе источников, используя длину волны 1 мм).

Более оптимальным из вышеперечисленных методов решений поставленной задачи оказывается следующий метод.

1. Получим из расстояний координаты от датчиков до каждого сегмента (выбрав систему отсчета относительно опорного кольца (ОК)).
2. Ищем параболу, идеальную по форме, но с неизвестным положением, наиболее близкую к измеренным точкам.
3. При нахождении точки под параболой двигаем актуатор вверх, в противном случае – вниз.

$\text{Min } |y_{\text{ид}} - y_{\text{д}}|$ – оптимизируем наихудший случай (при оптимизации форма дана жестко, а оптимизируется (меняется) положение параболы).

$\text{Min } \Sigma (y_{\text{ид}} - y_{\text{д}})^2$ – среднее квадратическое отклонение,

$\text{Min } \Sigma |y_{\text{ид}} - y_{\text{д}}|$ – минимизируем энергию ($y_{\text{ид}}$ – данные идеального параболоида в памяти ВУК, $y_{\text{д}}$ – данные, полученные после внешнего воздействия на систему).

Подстройка отражающей поверхности к положению ИП осуществляется путем перемещения большого числа $N \sim 1200$ отдельных точек в направлении, перпендикулярном зеркалу, при помощи актуаторов. Во время обучения нейросети на вход подается совокупность столбцов, каждый из которых содержит одну из возможных комбинаций параметров. Тогда длина обучающего вектора для нейросети будет равна Zm , где m – шесть параметров ИП, задаваемые в модели. На вход нейросети предъявляется последовательность, полученная решением уравнений на шаге интегрирования при идеальных значениях показаний системы (параметр управления подстройки равен 0). Таким образом, при работе нейросети на вход будут подаваться шесть искомых параметров ИП (с параметром управления подстройки равным 1), что не приводит к математическим усложнениям. Необходимо лишь запомнить значение суммы элементарных перемещений каждого актуатора при каждом акте.

Заключение

Результаты моделирования приведены на рис. 4 в виде графика сравнения алгоритмов. За счет быстрых алгоритмов математической обработки результатов измерений, время обучения уменьшилось более чем в 4 раза относительно использования менее быстрых нейросетевых алгоритмов.

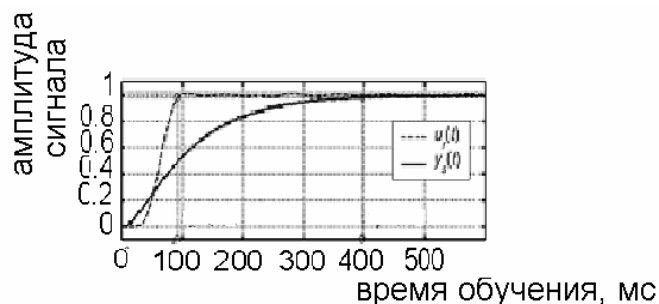


Рис. 4. Результат моделирования с алгоритмом обучения (показан пунктиром)

На рис. 5 приведен график изменения ошибки обучения в зависимости от числа выполненных циклов обучения. В результате получили 3 цикла обучения, это показывает, что алгоритм Левенберга-Марквардта – самый быстродействующий и пригоден для обучения ОЭС с несколькими сотнями настраиваемых параметров.

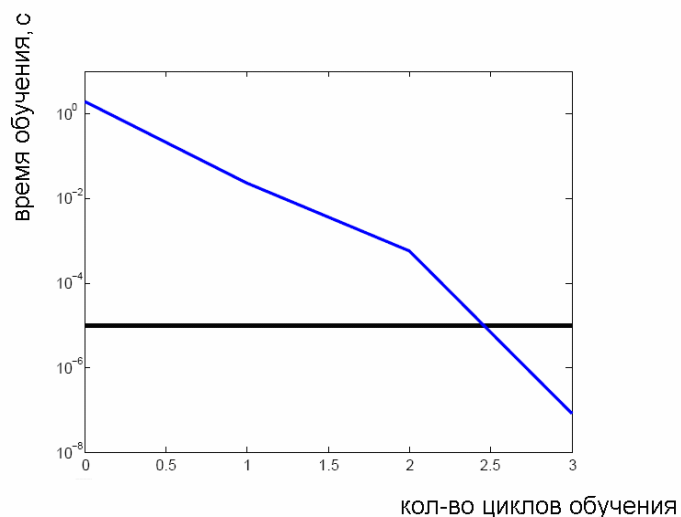


Рис. 5. График моделирования

Литература

1. Андреев А.Л., Кужаков П.В. Исследование оптико-электронной системы для высокоточного измерения пространственного положения объекта. // Международный оптический конгресс «Оптика –XXI век», конференция «Прикладная оптика-2006» 16-20 октября 2006 г., СПб.
2. Терехов В.А. Нейросетевые системы управления: Учеб. пособие для вузов. М.: Высш. шк., 2002. 183 с.
3. Городецкий А.Е., Козлов В.В., Артеменко Ю.Н., Тарасова И.Л., Вычисления в системах управления: Учеб. пособие. СПб: СПбГПУ, 2006. 463 с.
4. Обобщенный метод наименьших квадратов. Теория и применение в астрометрии. СПб: Наука, 1997. 318.
5. Андреев А.Л. Автоматизированные телевизионные системы наблюдения. Часть 2. Арифметико-логические основы и алгоритмы. Учебное пособие для курсового и дипломного проектирования. СПб: СПбГУИТМО, 2005. С. 56–83.
6. Сергиенко А.Б. Цифровая обработка сигналов. СПб: Питер, 2003. 608 с.
7. Медведев В.С. Нейронные сети. MATLAB 6. М.:Диалог-МИФИ, 2002. 496 с.

БЕЗАПЕРТУРНЫЙ РАМАНОВСКИЙ НАНОСПЕКТРОМЕТР БЛИЖНЕГО ПОЛЯ: ДЕМОНСТРАЦИЯ ЭФФЕКТА УСИЛЕНИЯ ОПТИЧЕСКОГО ПОЛЯ ПРИ РЕГИСТРАЦИИ РАМАНОВСКИХ СИГНАЛОВ

А.А. Ким

Научный руководитель – д.ф.-м.н., с.н.с. А.В. Баранов

В работе рассматриваются принцип действия и технические аспекты конструкции безапертурного рамановского наноспектрометра ближнего поля, использующего локализацию и усиление световых полей вблизи острия металлического (серебряного и золотого) зондов, приводятся результаты демонстрации эффекта усиления оптического поля при регистрации рамановских сигналов от модельных объектов – монокристаллических пластинок кремния и германия.

Введение

Важнейшей проблемой наноиндустрии (современных интегральных схем и интегральных схем ближайшего будущего) является развитие неразрушающих методов контроля параметров наноструктур, таких как химический состав, внутренняя структура и наличие локальных механических напряжений с пространственным разрешением нанометрового масштаба. Этим условиям удовлетворяет современная ближнепольная сканирующая оптическая спектроскопия (рамановская и люминесцентная), использующая в качестве светоподводящих элементов сверхтонкие оптические или металлические зонды, которые позволяют локализовать возбуждающее оптическое поле в область с размером в несколько десятков нанометров и получить необходимое нанометровое пространственное разрешение.

Теоретическое обоснование

Для определения формы, размеров и взаимного пространственного расположения нанообъектов в настоящее время широко используются методы атомно-силовой (АСМ) и туннельной микроскопии с пространственным разрешением до единиц ангстрем. Химический состав, внутренняя структура и наличие локальных механических напряжений с пространственным разрешением нанометрового масштаба могут быть определены с использованием техники просвечивающей электронной микроскопии и дифракции рентгеновских лучей. Однако эти методы требуют вакуумной техники, долговременной подготовки измерений и часто вызывают разрушение объектов исследования. Аналогичную информацию можно получить при использовании оптической спектроскопии рамановского рассеяния, однако существующие методы микро-рамановской техники позволяют получить пространственное разрешение, ограниченное дифракционным пределом ($\sim \lambda/2$). Потому возникла необходимость разработок новых оптических методов исследования наноструктур с наномасштабным разрешением. Такой новой разработкой оказалась ближнепольная микроскопия, использующая факт наличия в структуре электромагнитного поля осциллирующего диполя ближнепольной компоненты, локализованной в пространственном объеме $\sim R^3$, где R – характерный размер колеблющегося диполя (рис. 1) [1].

Экспериментальное подтверждение существования ближнего поля осциллирующего диполя, как физического явления, получено еще в 1928 году в экспериментах с электромагнитными полями радиодиапазона. В 1972 году Эшемом и Николем было предложено использовать пространственную локализацию ближнего поля для конструирования сканирующего микроскопа, построена действующая модель микроскопа ближнего поля в микроволновом диапазоне электромагнитного излучения и разработа-

ны общие принципы конструкции ближнепольного сканирующего оптического микроскопа (БСОМ).

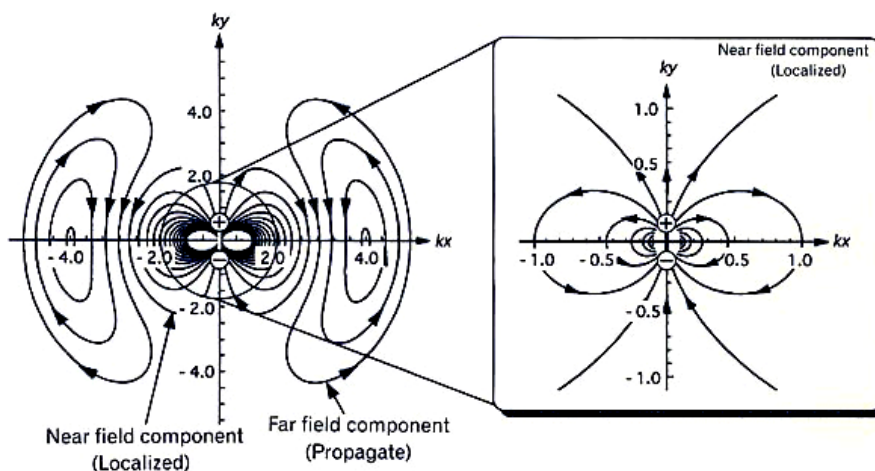


Рис. 1. Структура ближнего и дальнего электромагнитных полей диполя

Требуемой для работы БСОМ локализации оптического поля можно достичь с помощью двух известных на сегодняшний день методов: апертурного и безапертурного методов.

В апертурном методе лазерное излучение доставляется или собирается от исследуемого объекта с помощью оптического зонда с апертурой выходного зрачка излучения меньшей дифракционного предела Аббе ($d \ll \lambda$). Оптический зонд представляет собой световедущее оптическое волокно, конец которого заострен и покрыт металлической пленкой. Покрытие отсутствует на самом конце острия, формируя апертуру для оптического излучения. Разработаны технологии для управления размером выходной/входной апертуры волокна с помощью лазерной абляции и одновременного растяжения вдоль оси волокна, травления торца волокна в плавиковой кислоте. Апертурный метод, однако, имеет ряд принципиальных недостатков. Первым недостатком является малая пропускная способность такой наноапертуры ($T < 10^{-4}$), а вторым – разрушение металлического напыления на острие под действием лазерного излучения с мощностью выше 1–2 мВт. Два этих фактора делают практически невозможным получения спектров рамановского рассеяния от нанообъектов.

Безапертурный метод лишен такого порядка ограничений. Этот метод основан на эффекте локального усиления оптического электромагнитного поля вблизи поверхности металлического острия с малым радиусом кривизны. В данном методе зонд представляет собой металлическое Ag, Au, Cu острие, радиус кривизны которого должен составлять 10–50 нм. Физика процесса заключается в резонансном возбуждении локальных поверхностных плазмонов в металлическом зонде. Плазмоны – коллективные колебания электронов в металле, энергия (частота) которых зависит от типа металла. Поверхностные плазмоны могут быть возбуждены только излучением, поляризация которого параллельно поверхности металла. Коэффициент усиления зависит от условия резонансного возбуждения, которое определяется сортом металла (Au, Ag, Cu), размером и формой зонда, длиной волны и поляризацией лазерного излучения [2].

Благодаря резонансному возбуждению поверхностных плазмонов на вершине зонда, усиление поля может достигать 20 и 100 для Au и Ag зондов, соответственно. На рис. 2 для иллюстрации представлен расчет распределения напряженности оптического электромагнитного поля вблизи острия золотого зонда с радиусом кривизны ~ 5 нм.

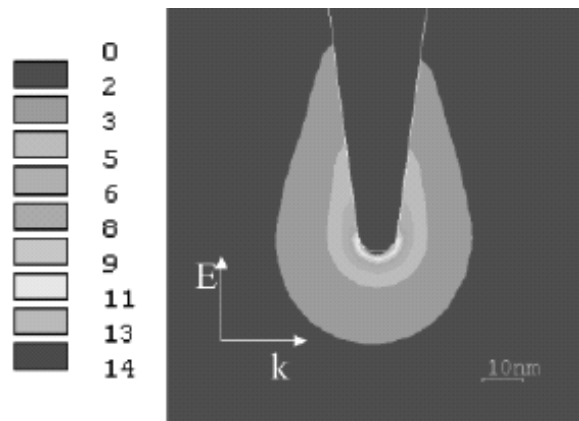


Рис. 2. Поле усиления вблизи Au зонда АСМ в лазерном освещении

Конструкция такого безапертурного БСОМ представляет собой комбинацию атомно-силового или туннельного сканирующего микроскопа с оптическим микроскопом, обычно использующимся при микро-люминесцентных или микро-рамановских измерениях. Такая установка, одна из схем которой показана на рис. 3, позволяет сканировать рельеф образца с помощью атомно-силового микроскопа и получать усиленный сигнал рамановского рассеяния за счет введенного в поле лазерного излучения острия АСМ.

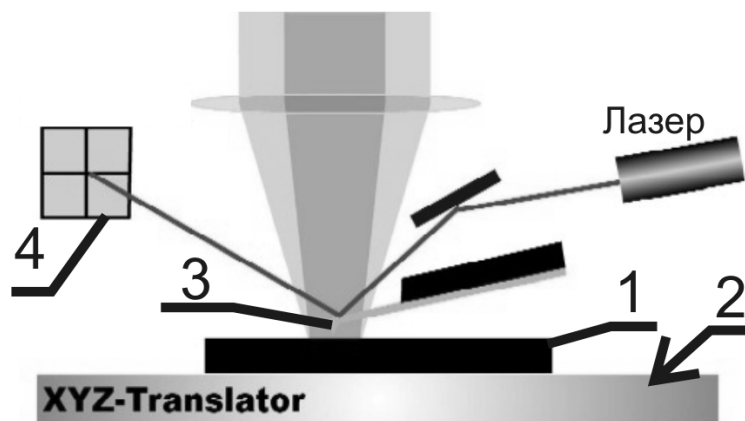


Рис. 3. Принципиальная схема работы безапертурного БСОМ: 1 – непрозрачный объект, 2 – координатный стол, 3 – металлический зонд, 4 – фотодетектор [3]

По сравнению с апертурной ближнепольной микроскопией безапертурная технология предоставляет более широкие возможности получения усиленного сигнала рамановского рассеяния и построения на его основе карт химического состава и механических напряжений, которые совместно с атомно-силовой микроскопией, дают о структурах нанометрового масштаба.

Конструкция макета наноспектрометра

Изложенный выше подход использован для разработки макета безапертурного рамановского наноспектрометра ближнего поля в Центре информационных оптических технологий ИТМО. Оптическая схема наноспектрометра приведена на рис. 4. Макет может работать в двух режимах: микро-рамановский спектрометр с дифракционным пространственным разрешением и наноспектрометр ближнего поля с пространственным разрешением в ~ 50 нм. Переключение между двумя режимами осуществляется посредством переключения поворотного зеркала (30). В обоих режимах оптическая схема

работает на отражение, что необходимо для характеристики наноструктур на непрозрачных подложках.

В режиме микроспектрометра качество получаемого сигнала зависит непосредственно от точности фокусировки лазерного пятна в плоскости предмета, а также чистоты лазерного фронта. Чистоту лазерного фронта можно повысить за счет расширения лазерного пучка телескопом (7) и отсеканием части фронта диафрагмой (8). Достаточно большое количество зеркал (9, 10, 12, 17, 18) может вызвать потери сигнала, поэтому в данной установке в процессе модернизации были заменены на зеркала с серебряным напылением без покрытия, коэффициент отражения которых составил 99 %.

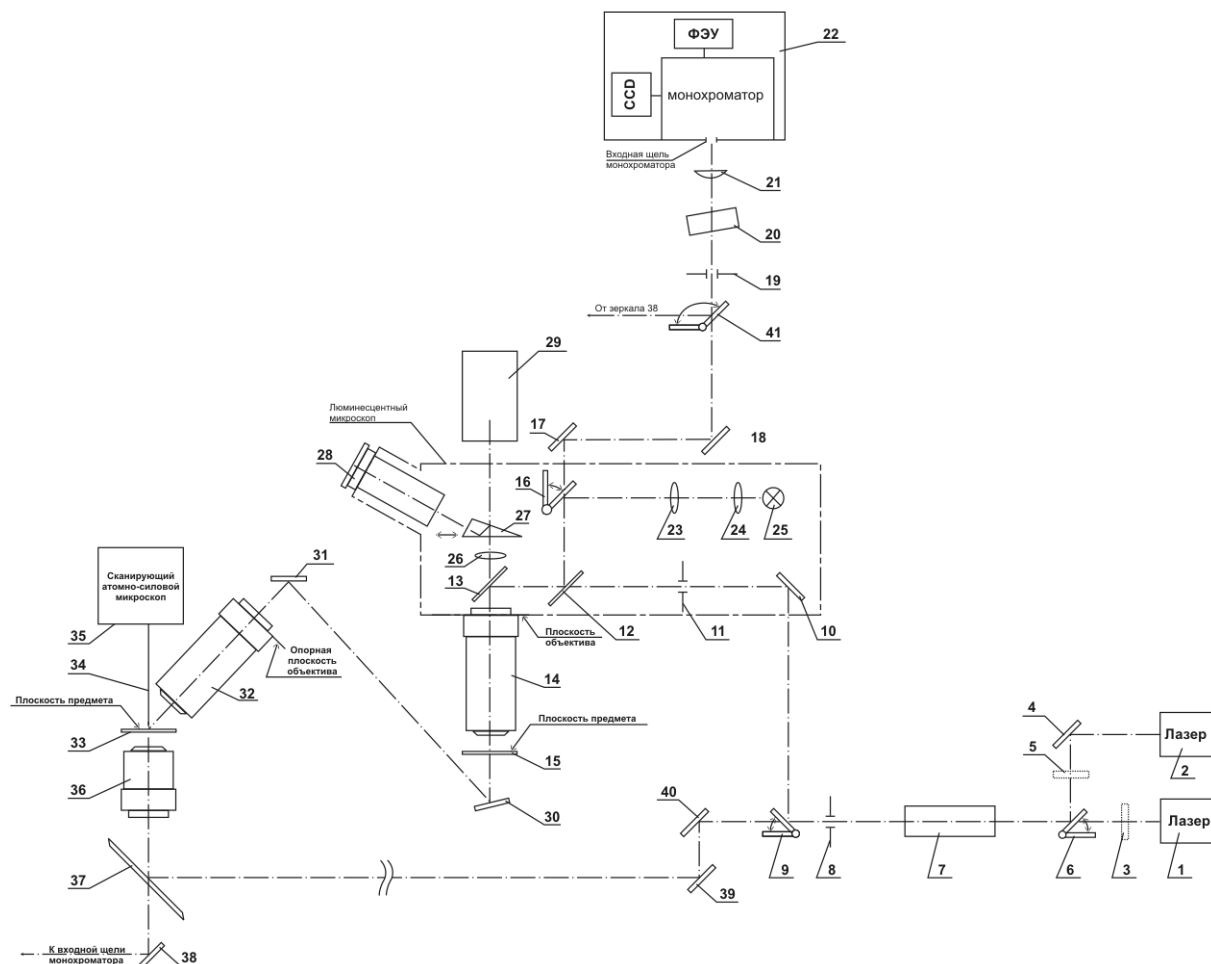


Рис. 4. Оптическая схема наноспектрометра

При работе прибора в режиме наноспектрометра объектив (14) удаляется, и возбуждающее излучение с помощью зеркал (30, 31) направляется на микрообъектив (32) с большим рабочим расстоянием, который фокусирует лазерное излучение на вершину зонда (34) и исследуемый образец в пятно с размером ~ 3 мкм. Вторичное излучение от образца (рамановское рассеяние) собирается тем же объективом и с помощью зеркал (31, 30, 13, 12, 17, 18) подается на входную щель спектрографа (22), оборудованного охлаждаемой CCD камерой, для спектрального анализа. Схема геометрии возбуждения/регистрации спектров КР приведена на рис. 5.

Поз.	Наименование	Поз.	Наименование
1	Лазер гелий-неоновый однододовый ГН-15-1	22	Спектрометр/спектрограф
2	Лазер твердотельный однододовый LCS-DTL-317	23	Линза
3	Светофильтр интерференционный	24	Коллектор
4	Зеркало поворотное	25	Лампа накаливания КГМ 9x70
5	Светофильтр интерференционный	26	Линза тубусная
6	Зеркало откидное	27	Призма
7	Расширитель (телескоп)	28	Окуляр
8	Диафрагма ирисовая	29	Видеокамера
9	Зеркало откидное	30	Зеркало поворотное
10	Зеркало поворотное	31	Зеркало поворотное
11	Диафрагма ирисовая	32	Микрообъектив
12	Пластика свето делительная	33	Препарат
13	Пластика свето делительная	34	Зонд
14	Микрообъектив	35	Микроскоп атомно-силовой сканирующий
15	Препарат	36	Эпиобъектив
16	Зеркало откидное	37	Зеркало кольцевое
17/18	Зеркало поворотное	38	Зеркало поворотное
19	Диафрагма ирисовая	39	Зеркало поворотное
20	Светофильтр голографический узкополосный	40	Зеркало поворотное
21	Линза	41	Зеркало откидное

Таблица. Элементы оптической схемы наноспектрометра, показанной на рис. 4

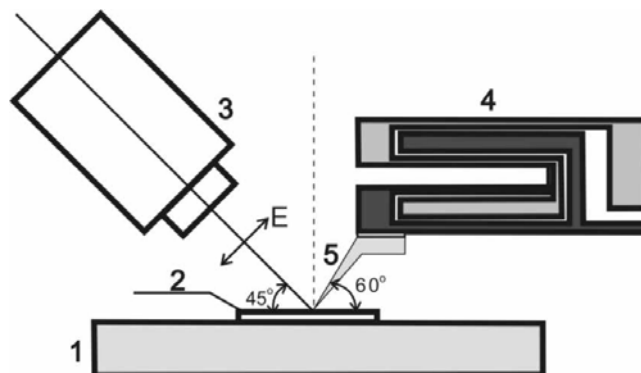


Рис. 5. Схема геометрии возбуждения/регистрации спектров КР нанообъектов на непрозрачной подложке с использованием наноспектрометра: 1 – предметный столик, 2 – образец, 3 – микрообъектив, 4 – пьезосканер, 5 – металлический зонд

В дальнейшем планируется установка в плоскости предмета (33) трехкоординатного пьезостолика «3D Flatscanner NIS-30» (фирмы Nanonics, Израиль) с точностью перемещения образца менее 1 нм. Это позволит позиционировать в трех измерениях исследуемый объект относительно зонда с помощью пьезостолика, в то время как зонд относительно лазерного луча может быть позиционирован с помощью головки сканирующего микроскопа.

Демонстрация эффекта усиления световых электромагнитных полей вблизи острия металлического зонда

Наличие эффекта усиления было продемонстрировано путем сравнения интенсивностей сигналов комбинационного (рамановского) рассеяния (КР), полученных в геометрии отражения от модельных объектов – монокристаллических пластинок Si и Ge, помещенных на предметный столик наноспектрометра. Измерения проводились в отсутствие металлического зонда у поверхности образцов и в случае, когда острие зонда находится на расстоянии ~ 2 нм от поверхности. Первый случай соответствует обычно дальнеполюному возбуждению КР (микро-КР) с интенсивностью $I_{\text{КР}}$. Во втором случае КР возбуждается как дальнеполюной, так и ближнеполюной (локализованной) компонентами оптического электромагнитного поля, и его интенсивность равна $I_{\text{КР}} + I_{\text{УКР}}$, где $I_{\text{УКР}}$ – интенсивность усиленного КР (УКР), возбуждаемого усиленным оптическим полем, локализованным вблизи острия зонда.

Лазерный пучок направлен таким образом, что угол между ним и плоскостью пластины (угол падения) составлял 45° . Поскольку зонд наклонен под углом 60° к плоскости пластины, то угол между осью зонда и направлением пучка составляет 75° . Лазерный пучок линейно поляризован в плоскости падения, и угол между его вектором поляризации и осью зонда составляет 15° . Схема геометрии возбуждения/регистрации КР приведена на рис. 5. Лазерный пучок фокусировался на поверхность образца микрообъективом с рабочим расстоянием 10,9 мм и числовой апертурой 0,5 в эллипсоидальное пятно с отношением малой и большой осей эллипса 2: 2,9 мкм. Спектральное разрешение составляло $4,5 \text{ см}^{-1}$.

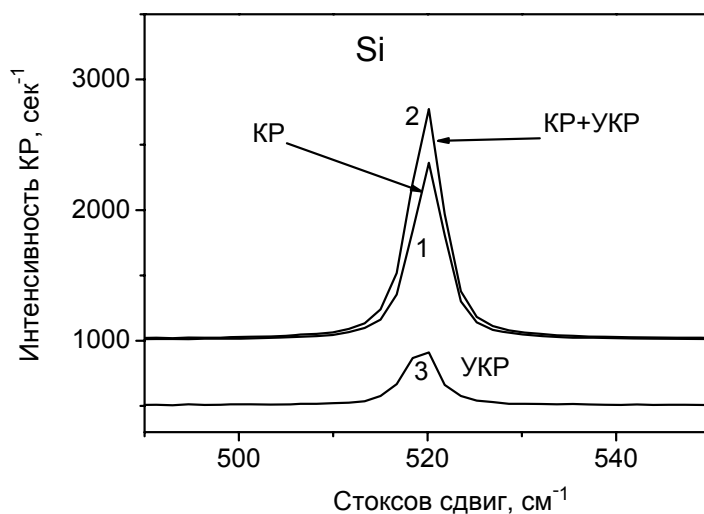


Рис. 6. Спектры КР монокристаллической пластины (001) Si, возбуждаемые излучением 532 нм: (1) – в случае отсутствия зонда вблизи поверхности ($I_{\text{КР}}$), (2) – острие серебряного зонда на расстоянии 2 нм от поверхности ($I_{\text{КР}} + I_{\text{УКР}}$), (3) – дифференциальный спектр $I_{\text{УКР}}$, получаемый при вычитании спектра (1) из спектра (2).

Мощность лазерного излучения на поверхности образца составляла 0,2 мВт.

Время записи спектров – 10 с

На рис. 6–7 показано сравнение спектров линий КР LO-фононами Si и Ge (520 см^{-1} и $300,8 \text{ см}^{-1}$, соответственно), полученных описываемым выше способом, демонстрирующее наличие интенсивной ближнеполюной компоненты оптического поля вблизи металлического острия. На рис. 6 приведены спектры КР монокристаллической пластины (001) Si, возбуждаемые излучением 532 нм, в случае отсутствия зонда вблизи поверхности, $I_{\text{КР}}$ и в присутствии острия серебряного зонда на расстоянии 2 нм от поверхности Si, $I_{\text{КР}} + I_{\text{УКР}}$.

Легко видно увеличение интенсивности КР в присутствии зонда. Наши измерения показывают, что в этом случае $I_{УКР}$ составляет $(0,2-0,30) \times I_{КР}$.

Тот же набор спектров LO-фонона германия $300,8 \text{ см}^{-1}$ приведен на рис. 7, где ясно видно 25–30 % возрастание интенсивности линии в случае нахождения острия серебряного зонда на расстоянии 2 нм от поверхности Ge.

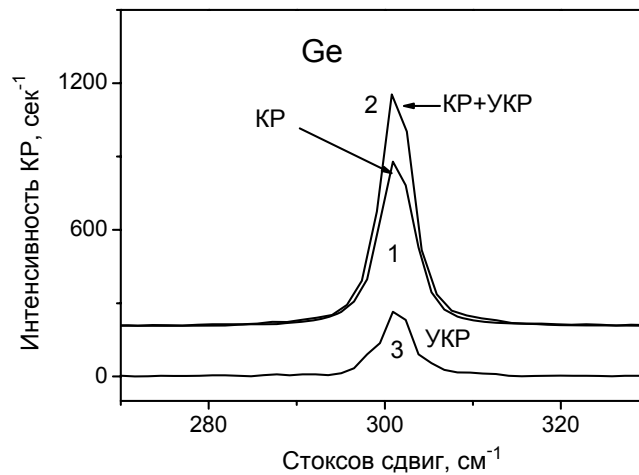


Рис. 7. Спектры КР монокристаллической пластинки (001) Ge, возбуждаемые излучением 532 нм (а) – в случае отсутствия зонда вблизи поверхности ($I_{КР}$). (б) – острие серебряного зонда на расстоянии 2 нм от поверхности ($I_{КР} + I_{УКР}$), (в) – дифференциальный спектр $I_{УКР}$, получаемый при вычитании спектра (а) из спектра (б).

Мощность лазерного излучения на поверхности составляла 0.2 мВт.

Время записи спектров – 10 с

Аналогичные измерения, проведенные на тех же образцах с использованием золотого зонда и возбуждающего излучения с длиной волны 632,8 нм, также показали наличие увеличения интенсивности КР в случае, когда острие зонда находится в 2 нм от поверхности образцов. Измерения показали, что $I_{УКР}$ составляет 0,2 $I_{КР}$. Этот результат является несколько неожиданным, поскольку ожидалось, что усиление локального поля для золотого зонда должно быть существенно меньше, чем для серебряного зонда. Возможно, наблюдаемое несоответствие связано с более оптимальной для реализации плазмонного резонанса формой и размерами изготавливаемых нами золотых зондов по сравнению с серебряными.

В результате было продемонстрировано наличие усиленной ближнепольной компоненты возбуждающего оптического поля вблизи острия серебряного и золотого зондов. Величина усиления рамановского сигнала может быть оценена из следующих простых соображений. Интенсивность рамановского сигнала пропорциональна произведению интенсивности возбуждающего излучения и освещаемого объема материала. Мы оценили, что усиленное поле локализовано в области с диаметром $\sim 50 \text{ нм}$ (объем $1 \times 10^{-4} \text{ мкм}^3$). В то же время весь объем, освещаемый сфокусированным лазерным излучением на поверхности образцов и дающий сигнал обычного рамановского рассеяния, составляет $\sim 3 \text{ мкм}^3$. Из рис. 6 и 7 видно, что интенсивность усиленного рамановского сигнала от локальной области составляет почти 30 % от интенсивности сигнала обычного рамановского рассеяния. Отсюда следует, что усиления рамановского сигнала достигает $\sim 10^4$, что близко по величине к значениям, получаемым другими исследователями [4, 5]. Полученные данные свидетельствуют о возможности использования разработанного макета для детектирования сигналов рамановского рассеяния с пространственным разрешением порядка 50 нм.

Заключение

Полученные результаты позволяют говорить о перспективности развития техники безпертурной ближнепольной рамановской микроскопии, которая позволяет получить информацию об исследуемом объекте с нанометровым пространственным разрешением, определить химический состав, внутреннюю структуру и наличие локальных механических напряжений косвенно определить форму, размеры и взаимное пространственное расположение нанообъектов.

Благодарности

Выражается благодарность сотрудникам Центра информационных оптических технологии Ю.М. Воронину, Г.Н. Виноградовой и П.С. Парфенову за помощь в работе и при подготовке этого материала.

Литература

1. Bohn J.L., Nesbitt D.J., Gallagher A. Field enhancement in apertureless near-field scanning optical microscopy. // J. Opt. Soc. Am. A 2001. Vol.18. No.12. P. 2998–3006.
2. Davis R.C., Williams C.C. Optical dipole model for photodetection in the near field. // J. Opt. Soc. Am. A 2001. Vol. 18. № 7. P. 1543–1551.
3. Sun W.X. and Shen Z.X. Apertureless near-field scanning Raman microscopy using reflection scattering geometry. // Ultramicroscopy, 2003. 94. P. 237–244.
4. Fleischmann M., Hendra P.J., McQuillan A.J. Raman spectra of pyridine adsorbed at a silver electrode. // Chem Phys Let, 1974. 26. 163.
5. L.T. Nieman, G.M. Krampert and R.E. Martinez. An apertureless near-field scanning optical microscope and its application to surface-enhanced Raman spectroscopy and multiphoton fluorescence imaging. // Review of Scientific Instruments. 2001. Vol. 72. № 3. P. 1691–1699.

РАЗРАБОТКА И СОЗДАНИЕ КОНСТРУКЦИИ РЕФЛЕКТИВНОГО ПОЛЯРИЗУЮЩЕГО ФИЛЬТРА ДЛЯ ЖК-ДИСПЛЕЕВ, СПОСОБСТВУЮЩЕГО СНИЖЕНИЮ ПОТЕРЬ СВЕТА

А.В. Волков, Л.А. Масаева

**Научный руководитель – ассистент кафедры оптических технологий
М.В. Погумирский**

В работе предлагается оптимизация конструкции жидко-кристаллических дисплеев путем введения между системой подсветки и первым поляризатором рефлективного поляризационного фильтра, позволяющего снизить потери света на первом поляризаторе.

В настоящее время как крупными компаниями-производителями ЖК-дисплеев, так и университетами и отдельными научно-коммерческими организациями ведутся работы, направленные на уменьшение потребляемой энергии устройствами с ЖК-дисплеями. Способы борьбы за экономию энергопотребления ЖК-дисплеями: использование более экономичных источников излучения в системе подсветки дисплеев (например: светодиоды), уменьшение потерь света в жидких кристаллах за счет синтеза новых химических соединений, уменьшение потерь света в системе подсветки, а именно на диффузных отражателях, рассеивателях и световодах и др. Решениями, позволяющими существенно снизить потери, являются первые два (надо заметить, не взаимозаменяемыми друг друга и предложенное в работе решение). Предлагается снизить потери света на первом поляризаторе за счет введения в конструкцию специально разработанного рефлективного поляризационного фильтра. В работе предлагается качественно новый подход решения проблемы энергосбережения.

Целью работы является разработка конструкции рефлективного поляризационного фильтра, который предназначен для работы в ЖК-дисплеях, с целью уменьшения потерь света в конструкции дисплея, а именно: создание модели рефлективного поляризационного фильтра; выявление зависимости хода лучей света в модели от профиля поверхности разработанной модели и покрытия, нанесенного на ее поверхность.

Дисплей состоит из источника излучения (в состав которого входят люминесцентные лампы и система отражателей); первого поляризатора, пропускающего линейно поляризованный свет, стекла, жидких кристаллов, цветного фильтра, стекла и второго поляризатора. В этой системе поглощают свет: первый поляризатор – 50 %, жидкие кристаллы с цветным фильтром – 60 % дошедшего до них света. Итого теоретически на выход поступает только 16 % падающего потока излучения, а на практике этот показатель еще меньше и составляет 8–10 %. Такие потери света приводят к большим энергозатратам. На данный момент эту проблему решают, используя новые, более экономичные источники света, такие как криптоновые лампы, светодиоды, уменьшая потери света в жидких кристаллах, за счет синтеза новых химических соединений с наиболее подходящими свойствами.

Следует отметить, что на первом поляризаторе происходят колоссальные потери. Обычно в качестве первого поляризатора используется поляроидная пленка, состоящая из упорядоченных кристаллов геропатита. Работа этого поляризатора основана на поглощении одной из компонент излучения. На выходе после этого поляризатора получается линейнополяризованный свет с интенсивностью менее 50 %.

Существующие сейчас поляризационные фильтры, изготовленные на основе интерференционных пленок, способны разделять свет на две компоненты: одну из них отражает, другую пропускает. Если создать условия, при которых отраженная компонента вернется в систему подсветки, а в системе будет присутствовать диффузный отражатель, то часть света, отразившись и изменив направление поляризации при втором взаимодей-

ствии с интерференционным фильтром, пройдет его, а оставшаяся часть другой компоненты вновь отразится. Соответственно, можно реализовать схему многолучевого переотражения. После такого рефлективного поляризационного фильтра можно поставить обычную поляризационную пленку, которая повысит степень поляризации света.

Для уменьшения потерь предлагается разработать конструкцию фильтра, расположенного параллельно плоскости излучения света между первым поляризатором и системой подсветки, пропускающего р-компоненту и отражающего s-компоненту света обратно в систему источника излучения (а точнее – в систему отражателей). Актуальной задачей является разработка конструкции поляризационного фильтра, эффективно разделяющего компоненты, на одной длине волны, для нормального угла падения на плоскость фильтра. В идеале конструкция должна пропускать 100 % р-поляризованного света и отражать 100 % s-поляризованного света на видимом диапазоне длин волн.

Наибольшее отношение величин р- и s-компонент можно видеть при углах падения света на поверхность оптической детали, близких к углу Брюстера и выше. Значит, для получения тонкого плоского поляризующего фильтра, расположенного в плоскости, параллельной плоскости дисплея, необходимо создать поверхность с неким профилем (наподобие линзы Френеля), на грани которого и будет наноситься покрытие.

Были рассмотрены несколько моделей профилей поверхности (рис. 1).

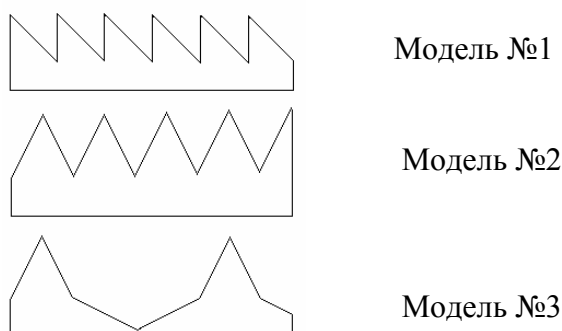


Рис. 1. Модели профилей поверхности

При рассмотрении хода лучей отраженного и прошедшего света были выявлены достоинства и недостатки данных моделей. Основным недостатком первой модели является несимметричность распространения прошедших лучей, что является недопустимым в дисплеях. У второй и третьей модели распространение лучей симметрично. Третья модель (с нанесенным покрытием) в рабочей зоне могла бы дать хорошие результаты, но во вспомогательной зоне она не работает в оптимальном режиме, поскольку отношение рабочей зоны к вспомогательной велико, то задействованными являются только 40 % поверхности. Вторая модель наибольшим образом подходит к дальнейшему рассмотрению. Общими недостатком всех моделей являются острые углы при вершинах профиля поверхности, которые вызывают увеличение рассеивания света.

Для решения поставленной задачи перейдем к выбору угла при вершинах профиля. В решаемой задаче необходимо вернуть свет под тем же углом, под каким он вышел из системы подсветки. Требованиям удовлетворяют два угла: 45° и 90° .

На рис. 2 представлен ход лучей, попавших на одну из граней профилей при нормальном падении излучения на плоскость фильтра. Углы падения света на грани профиля – 45° и $67,5^\circ$, соответственно. Недостатками второй модели относительно первой является большее число переотражений, а также большее количество «ступенек» профиля на единицу площади.

Рассмотрим свойства покрытия при наклонном падении излучения. Пусть плоская волна падает под некоторым углом α_0 на слой, расположенный на границе раздела

двух сред с показателями преломления n_0 и n_m , показатель преломления среды, из которой падает свет – n_0 , показатель преломления слоя – n_1 , показатель преломления среды, в которую свет распространяется – n_m , толщина слоя d_1 , оптическая толщина слоя $n_1 d_1$, фазовая толщина слоя при нормальном падении [1]:

$$\varphi_1 = 2\pi n_1 d_1 (\lambda)^{-1}.$$

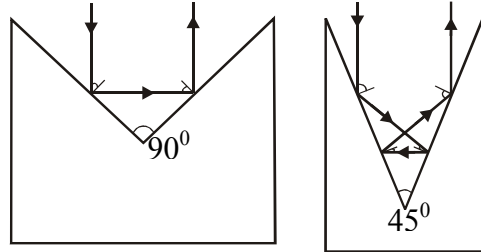


Рис. 2. Ход лучей при нормальном падении на плоскость фильтра

При наклонном падении фазовая толщина слоя становится равной

$$\varphi_{1\alpha} = 2\pi n_1 d_1 (\lambda)^{-1} \cos(\alpha_1),$$

где угол α_1 – угол распространения излучения в слое. Для системы из четвертьволновых слоев $n_1 d_1 = \lambda_0 / 4$, значение фазовой толщины $\varphi_{1\alpha} = \pi \lambda_0 \cos(\alpha_1) / 2\lambda$.

Эффективный показатель преломления κ_i для разных состояний поляризации равен

$$\kappa_i = \begin{cases} n_i \cos(\alpha_i) \\ \frac{\cos(\alpha_i)}{n_i} \end{cases},$$

для s-поляризации и p-поляризации, соответственно, где $i = 0, 1, m$.

Матрица интерференции для рассматриваемой системы:

$$\begin{vmatrix} \cos(\varphi_1) & \frac{1}{n_1} \sin(\varphi_1) \\ n_1 \sin(\varphi_1) & \cos \varphi_1 \end{vmatrix}.$$

Амплитудный коэффициент отражения для такой системы

$$r = \frac{(n_0 m_{11} - n_m m_{22}) + i(n_0 n_m m_{12} - m_{21})}{(n_0 m_{11} + n_m m_{22}) + (n_0 n_m m_{12} + m_{21})}.$$

Для света, поляризованного в разных плоскостях, при наклонном падении

$$\begin{cases} r^p = \frac{(n_o^p m_{11} - n_m^p m_{22}) + i(n_o^p n_m^p m_{12} - m_{21})}{(n_o^p m_{11} + n_m^p m_{22}) + i(n_o^p n_m^p m_{12} + m_{21})} \\ r^s = \frac{(n_o^s m_{11} - n_m^s m_{22}) + i(n_o^s n_m^s m_{12} - m_{21})}{(n_o^s m_{11} + n_m^s m_{22}) + i(n_o^s n_m^s m_{12} + m_{21})} \end{cases}.$$

Энергетический коэффициент отражения

$$R = |r|^2.$$

Конструкцию покрытия выбирается, исходя из требований к ее спектральным и эксплуатационным характеристикам. Поляризующими свойствами обладают многослойные четвертьволновые покрытия из чередующихся слоев с высоким и низким показателем преломления.

Пленкообразующие материалы должны быть прозрачными и иметь минимальное поглощение на длинах волн 350–800 нм. Тугоплавкие окислы (SiO_2 , ZrO_2), а также фторид магния (MgF_2) распространены, легко доступны и удовлетворяют указанным тре-

бованиям. Наносятся такие покрытия электронно-лучевым испарением. Метод электронно-лучевого испарения позволяет получать покрытия, содержащие до 70 слоев. Окислы имеют хорошую адгезию. Чередование их позволяет получить покрытие, обладающее нужными характеристиками. Помимо этого, покрытия, изготовленные из указанных окислов, относятся к 1 группе механической прочности и обладают высокой лучевой стойкостью.

Для расчета покрытия, отвечающего поставленным требованиям, использовалась программа MathCAD. Были рассчитаны две структуры покрытия для угла падения света 45° для длины волны 532 нм.

Покрытие №1 – 35-слойная структура, состоящая из чередующихся четвертьволновых слоев диоксидов циркония и кремния с показателями преломления 1,95 и 1,45, соответственно (рис. 3).

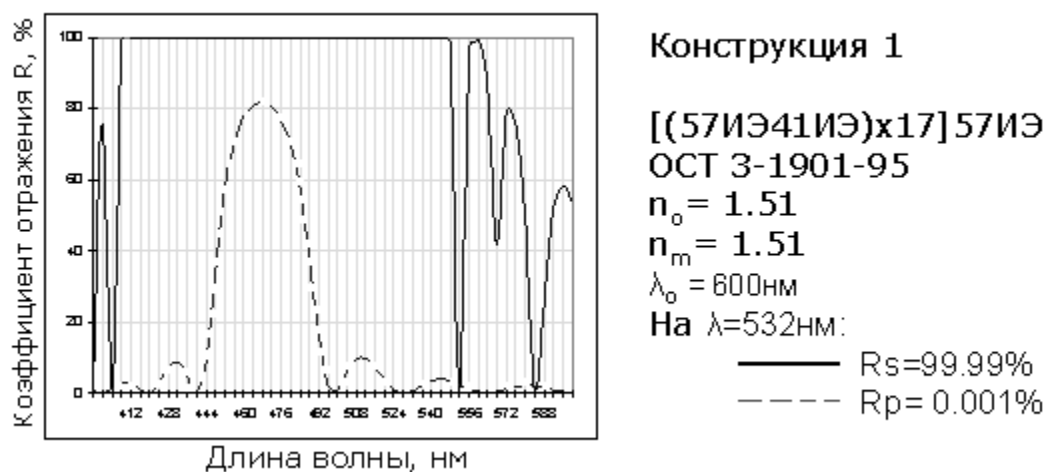


Рис. 3. Спектральная характеристика покрытия №1

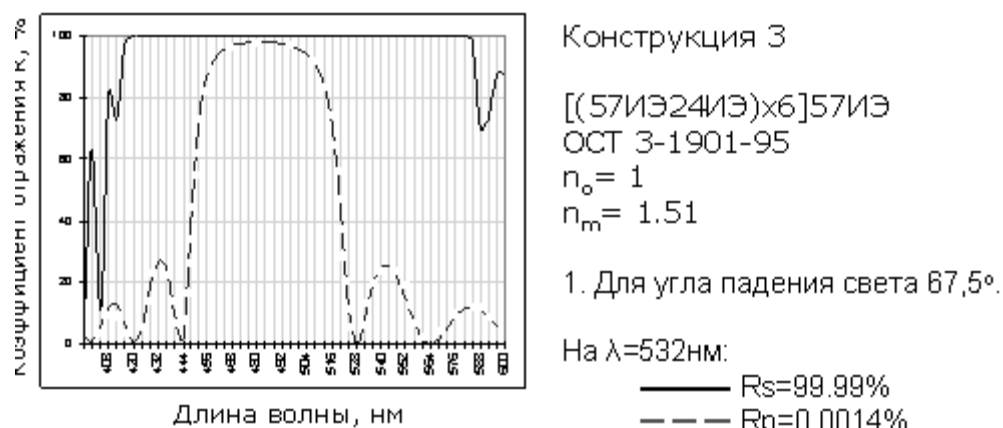


Рис. 4. Спектральная характеристика 31-слойной структуры

Излучение для получения таких спектральных характеристик должен падать из среды с показателем преломления таким же, как у подложки или близким к нему по значению. Такой поляризационный фильтр должен являться склейкой, что для данной конструкции является трудным в изготовлении. Было принято решение синтезировать покрытие для угла падения $67,5^\circ$, спектральная характеристика которого приведена на рис. 4. При отражении от первой грани свет падает на вторую под углом $22,5^\circ$, далее на первую под углом $22,5^\circ$, далее опять на вторую под углом $67,5^\circ$. Отсюда новые требо-

вания: либо покрытие должно эффективно разделять компоненты на обоих углах, либо при падении света под углом $67,5^\circ$ покрытие должно эффективно разделять компоненты, а при угле $22,5^\circ$ отражать обе компоненты света, что отражено на рис. 5.

Далее было синтезировано 31-слойное покрытие из чередующихся четвертьволновых слоев диоксида циркония и фторида магния с показателями преломления 1,95 и 1,37, соответственно. Подложка – стекло К8. Свет падает из среды с показателем преломления, равным 1.

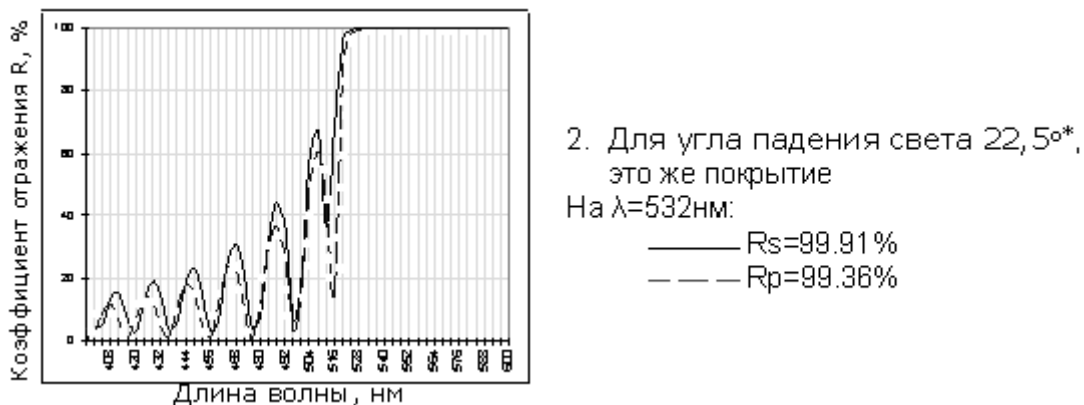


Рис. 5. Спектральная характеристика 31-слойной структуры для угла падения $22,5^\circ$

Рассмотрим рельефную структуру с нанесенным покрытием на длине волны 532 нм и ход лучей, отразившихся от поверхности и прошедших через модель лучей (рис. 6).

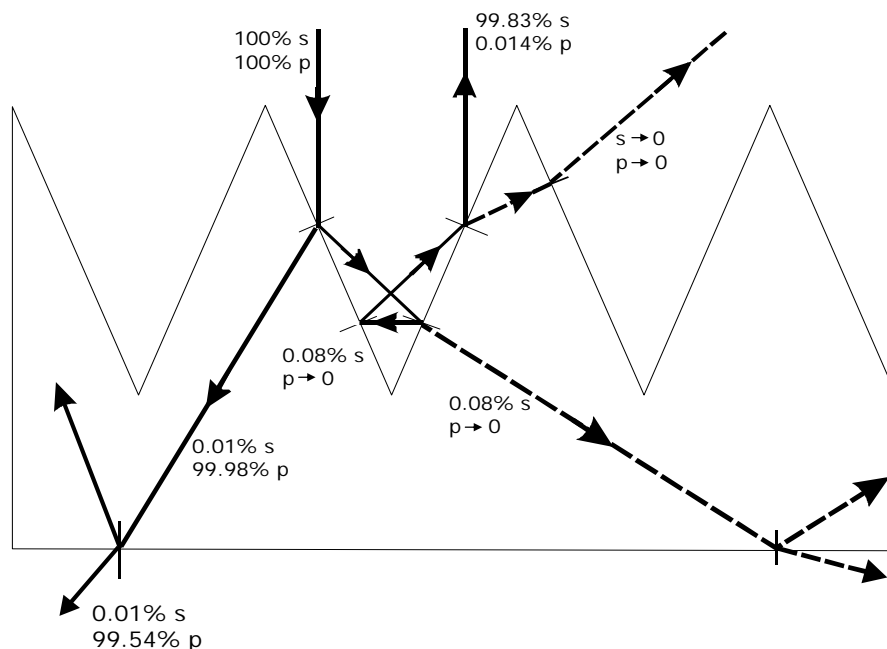


Рис. 6. Ход лучей в РП-фильтре

Луч падает под углом $67,5^\circ$ к поверхности, на которую нанесено покрытие. Как было показано выше, для данного угла произойдет разделение света на две компоненты, одна из которых, а именно р-компонента, пройдет в толщу пластинки и, преломившись, выйдет. Вторая s-компонента, претерпев три отражения, возвратится под тем же углом в систему подсветки. Теоретически можно получить следующее распределение компонент при первом переотражении. Пройдет через рефлексивный поляризующий фильтр 99,54 % р-компоненты и 0,01 % s-компоненты, а отразится 99,83 % s-компоненты и 0,014 % р-компоненты света.

В отражении и разделении света на компоненты участвуют две грани, следовательно, можно изготовить модель, отображенную на рис. 7 и состоящую из двух призм, на боковые наклонные грани которых нанесено поляризующее покрытие.

Покрытие наносилось только на часть грани, для проведения сравнительного анализа хода лучей и их интенсивностей.

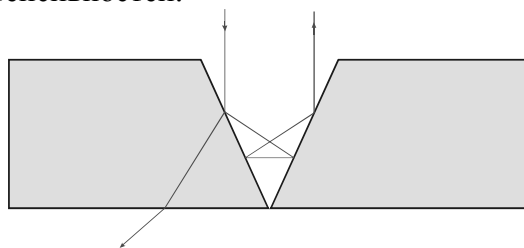


Рис. 7. Масштабная модель фильтра

Измерения интенсивности лучей прошедшего и отраженного пучка света, а также количественное отношение s и p компонент в этих лучах проводились на специально собранной установке, изображенной на рис. 8.

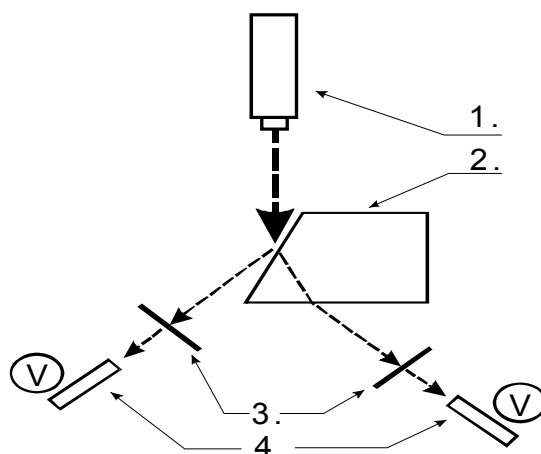


Рис. 8. Схема установки измерения интенсивности излучения: 1 – источник излучения + монохроматор; 2 – модель фильтра; 3 – поляризаторы (аналогичные используемым в ЖК-дисплеях); 4 – фотоприемники, подключенные к вольтметрам

На вольтметрах снималось три значения – без поляризатора 3, с поляризатором при $s=0$, с поляризатором при $p=0$.

Угол падения: $67,5^\circ$ Длина волны=532 нм						
	Отраженный свет			Прошедший свет		
	Без поляризатора, $s+p=$	$p=0, s=$	$s=0, p=$	Без поляризатора, $s+p=$	$p=0, s=$	$s=0, p=$
Утезн.	6	6	6	7	7	7
U1	198	110	18	145	9	49
U2	197	110	17	143	9	49
U3	198	110	18	143	9	48
U4	198	110	18	145	9	50
U5	199	110	19	144	9	49
Уср.	198	110	18	144	9	49
Пересчет данных с учетом ошибки фотоприемников (Утезн.):						
Уср.	198	110	18	197,5	12,3	67,2

Таблица. Измерения прошедших и отразившихся s и p компонент света

Изначально на модель падал свет в составе $p = 50\%$, $s = 50\%$. Для получения долевых соотношений решаем систему уравнений:

$$\begin{cases} s : 110x + 12,3 = 50 \\ p : 18y + 67,2 = 50 \end{cases}$$

$x = 0,41\%$, $y = 0,59\%$.

В итоге получаем:

$$\begin{cases} s_{\text{прошед}} = 12,3 \cdot 0,41 = 5,04\% \\ s_{\text{отраж}} = 110 \cdot 0,41 = 44,9\% \\ p_{\text{прошед}} = 67,2 \cdot 0,59 = 39,6\% \\ p_{\text{отраж}} = 18 \cdot 0,59 = 9,4\% \\ (s + p)_{\text{прошед}} = 5 + 39,6 = 44,6\% \end{cases}$$

Рассмотрим излучение, ушедшее обратно к источнику. Показания вольтметра 60,64 мВ – без поляризатора, 3 мВ – р-компонента, 57 мВ – s-компонента. Составив пропорцию, получаем $s = 37,2\%$, $p = 1,95\%$. Следовательно, через модель рефлективного поляризационного фильтра прошло 44,6 % света со смешанной поляризацией (39,6 % р-компонента). Отразилось к источнику излучения 39,05 % со смешанной поляризацией (37,2 % s-компонента).

При малых углах падения на металлическое зеркало происходит поворот поляризации приблизительно на 45° (проверено экспериментально), т.е. происходит эллиптическая поляризация. В результате после отражения получаем практически естественно поляризованный свет с интенсивностью 39,05 % относительно света, изначально вышедшего из источника излучения.

Если пренебречь потерями в системе переотражателей, можно просчитать, какое количество s-компоненты света пройдет через рефлективный поляризационный фильтр при семикратном переотражении света между фильтром и системой отражателей:

$$R_7 = 39,6 + 39,05 \cdot 36,6 + 39,05^2 \cdot 39,6 + 39,05^3 \cdot 39,6 + 39,05^4 \cdot 39,6 + 39,05^5 \cdot 39,6 + 39,05^6 \cdot 39,6 = 64,88\%$$

Следовательно, яркость увеличится в $k = \frac{64,88\%}{50\%} = 1,29$ раза.

Теоретически рассчитанные показатели модели рефлективного поляризационного фильтра сильно отличаются от практических значений. Выделим возможные причины ошибки.

Измерения и расчеты проводились на длине волны 532 нм. Ширина щели монохроматора, предназначенного для выделения определенной длины волны, была установлена 2 мм (для получения широкого пучка света), а это говорит о том, что был получен расходящийся пучок света на некотором малом диапазоне длин волн. Кроме того, проблему вызвала привязка столика с моделью поляризующего фильтра к источнику излучения, что имеет влияние на угол падения пучка света на поверхность модели.

Заключение

В ходе работы была создана модель рефлективного поляризационного фильтра. Выявлена зависимость хода лучей от профиля поверхности разработанной модели и покрытия, нанесенного на ее поверхность, разработана конструкция рефлективного поляризатора. Получено увеличение эффективности работы системы на длине волны 532 нм при нормальном угле падения света на плоскость фильтра на 29 %, что подтверждено экспериментально на созданной модели.

Требуются дополнительные исследования, направленные на оптимизацию характеристик данной конструкции, а именно, на увеличение рабочего диапазона длин волн и расширения диапазона углов падения света, на которых фильтр будет соответствовать требованиям.

При выполнении работы были решены следующие задачи:

- создана модель рефлексивного поляризационного фильтра, предназначенного для работы в ЖК-дисплеях,
- выявлены зависимости хода лучей света в модели от профиля поверхности и покрытия, нанесенного на ее поверхность,
- создана крупногабаритная модель РП-фильтра.

Литература

1. Ландсберг Г.С. Оптика. М.: Наука, 1976. Т.3. 928 с.
2. Борн М., Вольф Э. Основы оптики. М.: Наука, 1970. 856 с.
3. Губанова Л.А., Путилин Э.С. Интерференционные покрытия, формирующие энергетические и волновые параметры излучения. СПб: СПбГУ ИТМО, 2006. 220 с.
4. Путилин Э.С. Учебное пособие по курсу «Оптические покрытия». СПб: СПбГУ ИТМО, 2005. 199 с.

**ОСОБЕННОСТИ ПОСТРОЕНИЯ ИНКЛИНОМЕТРА
ПО ДИАМЕТРАЛЬНОЙ СХЕМЕ НА ОСНОВЕ ДНГ****Д.А. Соколов****Научный руководитель – к.т.н., доцент Я.И. Биндер**

Проведен обзор существующих схем гироинклинометров на датчиках угловой скорости. Рассмотрены принципы, положенные в основу диаметральной схемы. Выполнен сравнительный анализ диаметральной схемы и схем, разработанных ранее. Представлены конструктивные особенности реализации схемы.

Введение

Инклинометрия – метод определения пространственных координат скважины, позволяющий установить правильность бурения в заданном направлении.

Направленное бурение постепенно становится основным видом бурения, как на суше, так и на море при проходке скважин со стационарных морских платформ. Одновременно существует тенденция повышения требований к точности попадания забоя скважин в заданную точку и к соблюдению проектного профиля скважины. Поэтому необходимо обеспечивать эффективный контроль пространственного положения ствола скважин.

Жизненная необходимость сокращения сроков строительства скважин и увеличения производительности труда в бурении ставит перед создателями инклинометрических приборов и систем задачу повышения не только точности соответствующей измерительной аппаратуры, но и оперативности получения инклинометрической информации, а также сокращения затрат времени при проведении инклинометрических работ. Это привело к разработке новых приборов и информационно-измерительных систем с использованием последних научно-технических достижений. Совершенствуются методы и алгоритмы обработки данных, широко применяется вычислительная техника [1].

Цель бурения состоит в том, чтобы, во-первых, забой скважины достиг проектной точки, во-вторых, был подготовлен ствол скважины такого профиля и такого качества, которые обеспечили бы нормальную дальнейшую эксплуатацию скважины. Соответственно, существуют и требования как к допустимому отклонению забоя от проектной точки, так и к профилю ствола скважины. Для соблюдения этих требований необходимо использовать приборы, дающие информацию о пространственном положении ствола скважины – инклинометры.

Основная цель состоит в том, чтобы добиться универсальности применения бесплатформенного гироинклинометра в непрерывном и точечном режимах при всех возможных траекториях скважин. Достижение этой цели стало возможным благодаря детальному исследованию различных схем компасирования в подземной навигации и предложенной в работе [2] схеме с ориентацией главной оси двухосного датчика угловой скорости в диаметральной плоскости скважины, поэтому данную схему было предложено назвать диаметральной. Итогом разработки диаметральной схемы оказалась возможность не только компасирования произвольно ориентированных скважин, но и проведения калибровки в пуске. При этом для траекторий скважин, не очень близких к плоскости первого вертикала, эти процедуры эффективно совмещаются, позволяя организовать компасирование, инвариантное к пусковой погрешности ДУС – режим, позволяющий минимизировать время цикла в точке измерения.

Обзор существующих схем гироинклинометров на датчиках угловой скорости

Диаметральная схема является результатом эволюционного развития и видоизменения схем, реализованных ранее, последней из которых является поперечная схема, описанная в работе [3].

Обратим внимание на то, что время съемки скважины является самостоятельным и, притом, весьма существенным технико-экономическим фактором, в значительной степени влияющим на выбор используемой при измерениях инклинометрической системы. Поэтому вполне естественно, что наряду с разработками в области точечного компасирования едва ли не большее внимание уделяется гироинклинометрам (ГИ), предназначенным для непрерывной съемки скважин и работающим в режиме гироскопа направления (ГН). Такие ГИ, как правило, имеют так называемую продольную компоновку, позволяющую измерить компоненты абсолютной угловой скорости и после их учета при интегрировании уравнений Пуассона получить искомые параметры ориентации скважины. Поперечная же схема не обеспечивает работу в режиме ГН. Американские авторы, проводившие весьма серьезные теоретические исследования и объективные испытания режима бесплатформенного ГН, считают его по точности вполне соизмеримым с компасированием. В отечественной литературе этому режиму уделялось значительно меньше внимания, тем не менее, необходимо отметить, что именно ГН (даже и реализованный на базе одноосного гиросtabilизатора в кардановом подвесе) положен в основу единственного серийного ГИ российского производства ИГН-73. Принципиальной особенностью этого режима является необходимость начальной выставки. Она, в свою очередь, реализуется либо с помощью внешнего курсоуказания, которое, очевидно, возможно только в устье скважины, либо в режиме гироскопирования самого ГИ (понятно, что в бесплатформенном исполнении это всегда принципиально возможно). Однако в первом случае режим съемки становится уязвимым к срывам питания, отказам и повреждениям каротажного подъемника (в кабельном варианте ГИ) и иным сбойным ситуациям, которыми изобилует геофизическая практика. Второй вариант (с автономным гироскопированием), естественно, предпочтительней – после ликвидации сбоя выставка может быть проведена вновь в той части ствола, где на этот момент оказался ГИ.

Однако точность, да и сама возможность такой выставки будет определяться факторами, детально исследованными при рассмотрении режима компасирования, и, например, так называемая «продольная схема» для бесплатформенных ГН (в частности, на базе ДУС с механическим носителем вектора кинетического момента), представляющаяся единственно возможной, оказывается практически неприменимой в горизонтальных стволах, проходящих вблизи линии «восток–запад». Но и при вполне благоприятной ориентации ствола в точке проведения повторной выставки ее точность в продольной схеме, определяемая пусковыми погрешностями ДУС, будет ограниченной. Для снижения погрешности выставки ее стараются проводить на участках траектории с зенитными углами (θ), достаточно малыми для того, чтобы воспользоваться эффектом практического обнуления «гироскопической» составляющей азимутальной ошибки, но не настолько, чтобы оказалась слишком велика доля составляющей от ошибки горизонтирования. Однако такой «параметрический» вариант начальной выставки, очевидно, противоречит «декларации о технологичности» режима ГН и во многом обесценивает само его использование.

Следует признать, что непрерывный режим бесплатформенного ГН вторичен по отношению к точечному, возможность его эффективного использования и верхний предел точности определяются соответствующими характеристиками гироскопирования. Поэтому, если задаться целью создания универсального, т.е. способного работать как в компасном, так и в непрерывном режиме при любых траекториях стволов,

ГИ, его схеме должны быть присущи некоторые комплексные признаки, удовлетворяющие требованиям обоих режимов.

Принципиально не связанная зависимостью от параметров траектории ствола, поперечная схема для бесплатформенных ГН, как уже указывалось, вряд ли перспективна. При этом наряду с «неуниверсальностью» существуют и другие доводы, стимулирующие поиски направления дальнейшего развития схем инвариантного компасирования, прежде всего в части кинематики.

Сравнительный анализ диаметальной схемы и схем, разработанных ранее

В таблице приведены данные сравнительного анализа, проводимого, исходя из критериев универсальности, определяющих конструктивный облик и функциональное назначение системы подземной навигации. В строках 1–3 кратко охарактеризованы исследованные в цитировавшихся выше работах кинематические схемы ГИ.

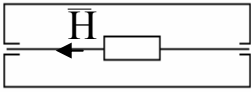
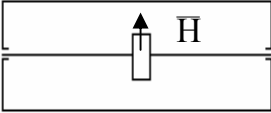
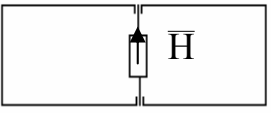
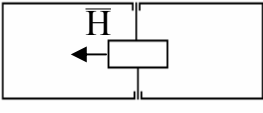
Характеристики Кинематическая схема	Адаптивность к траектории	Возможность инвариантного компасирования	Возможность режима ГН	Эффективность дополнительного поворота	
				Адаптивность	Инвариантное компасирование
1 	-	-	+	-	-
2 	+	+	-	-	-
3 	-	-	-	+	-
4 	+	+	+	+	+

Таблица «Сравнение кинематических схем гироинклинометров»

Продольная схема в компасном режиме имеет пределы применимости, освещенные в работе [1] и, естественно, не может претендовать на адаптивность к траектории, не говоря уже об инвариантном компасировании. В то же время такая схема, обеспечивая минимально возможный диаметр скважинного прибора и реализацию режима ГН, надо полагать, останется так или иначе востребованной.

Приведенные во второй и третьей строках таблицы варианты поперечной схемы используют только одну ось вращения корпуса гироскопа. Схема с двумя поворотами корпуса, образующая совмещением двух одноосных, фактически реализует двухосный карданов подвес, который может быть выполнен достаточно простым и техноло-

гичным, чтобы существенно не сказаться на ценовых показателях, но при этом неизбежны увеличение диаметра ГИ и дополнительные, так называемые геометрические, погрешности (их величина и структура – обратная сторона упрощенного выполнения карданова подвеса). При этом анализ показывает, что их неучет приводит к ошибкам компасирования, резко возрастающим на вертикальных и горизонтальных участках стволов.

Наряду с выявлением этих факторов – увеличением габаритов и кардановыми погрешностями – первые результаты разработки и внедрения ГИ с поперечной схемой привели к корректировке подхода к возможностям управления в интересах инклинометрии, положением трубной колонны по углу поворота отклонителя (ψ). Такое управление, естественно, возможно только при нахождении ГИ в составе компоновки низа буровой колонны (КНБК) в процессе сопровождения бурения, либо каротажа, осуществляемого с помощью спуска на трубах, причем последний вариант в последнее время становится все более распространенным. Если кинематическая схема ГИ предусматривает возможность поворота корпуса гироскопа вокруг оси скважинного прибора (двухколесная схема или вариант 2 в таблице), то, очевидно, что управление трубной колонной по углу ψ бессмысленно. В 3 схеме (см. табл.) такое управление придаст ГИ адаптивность к траектории, но в силу невозможности фиксировать этот угол с точностью, необходимой для инвариантного компасирования, ТХ такой одноколесной схемы все равно будут уступать варианту 2 (см. табл.).

Таким образом, мы можем констатировать бесспорную предпочтительность среди одноколесных поперечной схемы с вращением корпуса гироскопа вокруг оси скважинного прибора, несмотря на очевидное ограничение [1] траекторий, на которых возможно инвариантное компасирование ($|1 - \cos A| \gg 0$, где A – азимут скважины), и в этой связи – неэффективность использования управления трубной колонной по углу поворота отклонителя и привычную невозможность комбинированного режима – совмещения в ГИ режимов компасирования и ГИ.

В четвертой строке таблицы предложена кинематическая схема ГИ, свободная от перечисленных недостатков. Проведем анализ ее основных свойств.

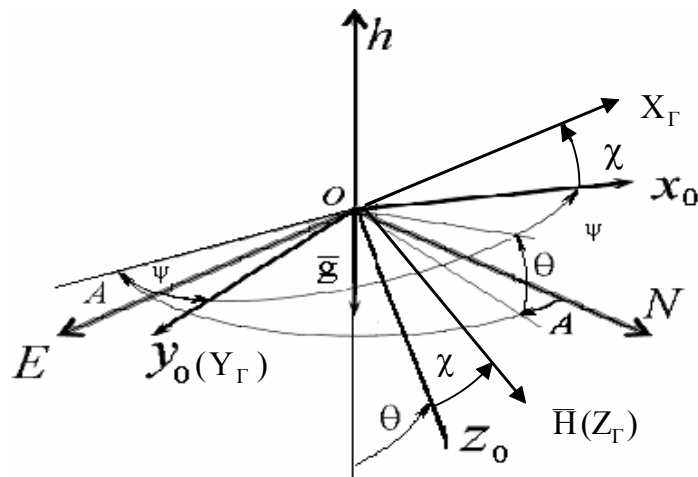


Рис. 1. Системы координат

Введем следующие системы координат (рис. 1): ENh – горизонтная с географической ориентацией осей (соответственно на восток, север и по вертикали места); $ox_0y_0z_0$ – связанная с корпусом гириноклинометра, ось oz_0 которой ориентирована вдоль продольной оси, а оси ox_0, oy_0 лежат в плоскости поперечного сечения СП; $oX_GY_GZ_G$ – связанная с корпусом ДУС, ось oZ_G которой совпадает с вектором кинематической оси гироскопа.

тического момента, а оси oX_G, oY_G – с измерительными осями ДУС. Положение системы координат $o x_0 y_0 z_0$ относительно географического трехгранника ENh определяется тремя ранее введенными углами – A, θ и ψ .

Положение системы координат $oX_G Y_G Z_G$ относительно $o x_0 y_0 z_0$ определяется углом χ (см. рис. 1) разворота вектора кинетического момента гироскопа относительно оси скважинного прибора. Очевидно, этот угол является тем универсальным параметром, который позволяет перейти от рассмотрения обобщенной схемы ГИ к конкретным вариантам ее реализации. Действительно, продольной и поперечной схемам соответствуют значения $\chi = 0^\circ$ и 90° соответственно (строки 1–3 таблицы).

Принципиальным отличительным признаком схемы, описываемой в четвертой строке таблицы, является возможность управления углом χ , реализуемая с помощью поворота вокруг оси, перпендикулярной вектору кинетического момента и лежащей в диаметральной плоскости скважины (определяемой аналогично диаметральной плоскости корабля). Такую схему ГИ, впервые описываемую в данной работе, назовем «диаметральной» и проведем здесь ее рассмотрение в соответствии с критериями таблицы.

Укажем еще раз, что при необходимости перехода в режим ГИ корпус гироскопа просто устанавливается в положение, $\chi = 0^\circ (180^\circ)$, и схема становится продольной.

Принципы, положенные в основу диаметральной схемы

Для синтеза закона управления углом χ в режиме компасирования необходимо рассмотреть выражения для сигналов измерительных каналов ДУС. При этом воспользуемся наиболее общей моделью уходов для гироскопа с механическим носителем вектора кинетического момента [2] и примем во внимание выражения для проекций кажущегося ускорения на измерительные оси ДУС:

$$\begin{aligned} n_{X_G} &= n_{x_0} \cos \chi - n_{z_0} \sin \chi, \\ n_{Y_G} &= n_{y_0} \end{aligned} \quad (1)$$

при том, что

$$n_{x_0} = -g \sin \theta \cos \psi, \quad n_{y_0} = -g \sin \theta \sin \psi, \quad n_{z_0} = g \cos \theta.$$

В соответствии с вышеизложенным имеем:

$$\begin{aligned} \omega_{X_G} &= \Omega \sin \varphi (\sin \theta \cos \psi \cos \chi + \cos \theta \sin \chi) + \\ &+ \Omega \cos \varphi \cdot [\cos A (\cos \theta \cos \psi \cos \chi - \sin \theta \sin \chi) + \sin A \sin \psi \cos \chi] + \tau_{X_G} - \\ &- \rho_{X_G} (\sin \theta \cos \psi \cos \chi + \cos \theta \sin \chi) + \varepsilon_{X_G} \sin \theta \sin \psi, \end{aligned} \quad (2)$$

$$\begin{aligned} \omega_{Y_G} &= \Omega \sin \varphi \sin \theta \sin \psi - \Omega \cos \varphi \cdot (\sin A \cos \psi - \cos A \cos \theta \sin \psi) + \tau_{Y_G} - \\ &- \rho_{Y_G} \sin \theta \sin \psi - \varepsilon_{Y_G} (\sin \theta \cos \psi \cos \chi + \cos \theta \sin \chi) \end{aligned} \quad (3)$$

В выражениях (1)–(3): Ω – угловая скорость суточного вращения Земли; φ – широта места объекта; g – ускорение свободного падения; $\omega_{X_G}, \omega_{Y_G}$ – выходные сигналы измерительных каналов ДУС; $n_{x_0}, n_{y_0}, n_{z_0}$ – составляющие кажущегося ускорения, измеряемые блоком акселерометров; τ_{X_G}, τ_{Y_G} – составляющие скорости ухода ДУС, не зависящие от ускорения (корпусные); ρ_{X_G}, ρ_{Y_G} – составляющие скорости ухода ДУС, пропорциональные проекциям ускорения на оси чувствительности измерительных каналов (дебаланс); $\varepsilon_{X_G}, \varepsilon_{Y_G}$ – составляющие скорости ухода ДУС, пропорциональные проекциям ускорения на оси, ортогональные к осям чувствительности измерительных каналов (квадратурные составляющие).

Чтобы сформулировать требования к углу χ по критерию адаптивности, т.е. обеспечения компасирования при любых параметрах траектории, рассмотрим (2) и (3) как систему двух линейных уравнений относительно неизвестных $\sin A$, $\cos A$ и найдем ее главный определитель

$$\Delta = -\cos\theta\cos\chi + \sin\theta\sin\chi\cos\psi. \quad (4)$$

Значение Δ , согласно (4), определяется зенитным углом ствола на данном участке траектории (θ), положением скважинного прибора относительно его продольной оси (ψ), а также углом разворота (χ) корпуса гироскопа в диаметральной плоскости скважины. Вполне очевидно, что мы не в состоянии изменять первый параметр. Не столь однозначна ситуация с углом ψ – в ряде случаев на его значение можно эффективно влиять, однако в данном случае – наиболее общего рассмотрения задачи – будем считать этот параметр неуправляемым. Рассмотрим условия обнуления Δ , не зависящие от значения управляемого параметра χ . Это возможно, если ГИ в горизонтальной скважине занимает такое положение, что ось разворота по углу χ оказывается вертикальной ($\cos\theta = \cos\psi = 0^\circ$). В этом случае соответствующий сигнал ($\omega_{\chi\Gamma}$) уже не несет информацию об азимуте, в чем легко убедиться, проанализировав уравнение (3), а из (2) получим (положив для определенности $\theta = \psi = 90^\circ$)

$$\sin(A - \chi) = \frac{\omega_{\chi\Gamma} - \omega_{\chi\text{др.}}}{\Omega \cos\varphi}, \quad (5)$$

(где $\omega_{\chi\text{др.}}$ – суммарный дрейф), откуда следует, что, несмотря на обнуление Δ (определяемое ориентацией скважинного прибора), задача выработки A решается путем управления углом χ .

Во всех других случаях обнуления главного определителя системы, как следует из (4), можно избежать, если, управляя углом χ , соблюдать условие

$$\text{tg}\chi_0 \neq \frac{\text{ctg}\theta}{\cos\psi}. \quad (6)$$

Отметим, строгости ради, что и в случае знака равенства в выражении (6) компасирование остается возможным, но его эффективность уже обретает параметрическую зависимость от траектории и положения ГИ в скважине и, как можно убедиться из анализа, например (3), теряется при условии

$$\text{tg}A_0 = -\frac{1}{\text{tg}\psi \cos\theta}. \quad (7)$$

В общем же случае управления углом χ мы показали, что «диаметральная» схема обеспечивает компасирование при любых траекториях скважин и разворотах скважинного прибора, и теперь проведем анализ возможностей и условий реализации режима, инвариантного к пусковым погрешностям ДУС.

Итак, если появлению знака равенства в соотношении (6) соответствует в общем случае потеря эффективности компасирования, то естественно было бы предположить, что при

$$\text{tg}\chi_0 = -\text{tg}\theta \cos\psi \quad (8)$$

эта эффективность окажется максимальной. При этом значение угла χ_0 будет на 90° отличаться от его же значения в случае выполнения равенства (6). Подставив (8) в (2) и представляя последнее в форме, предполагающей разрешение относительно тригонометрических функций азимута, получим:

$$\begin{aligned} \Omega \cos\varphi \cdot [\cos A(\cos\theta \cos\psi \cos\chi_0 - \sin A \sin\chi_0) + \sin A \sin\psi \cos\chi_0] = \\ = \omega_{\chi\Gamma} - \tau_{\chi\Gamma} + \varepsilon_{\chi\Gamma} \sin\theta \sin\psi \end{aligned} \quad (9)$$

или

$$\sin(A - A_0) = \frac{\omega_{x\Gamma} - \tau_{x\Gamma} + \varepsilon_{x\Gamma} \sin \theta \sin \psi}{\Omega \cos \varphi}, \quad (10)$$

где значение A_0 при условии (8) определяется выражением (7).

Приведенные выше соотношения (7)–(10) позволяют сделать следующие выводы:

- неопределенности дрейфов типа осевого дебаланса ротора гироскопа, являющиеся для механического носителя вектора кинетического момента основными возмущающими факторами, не оказывают влияние на точность компасирования;
- устанавливая ДУС в диаметральной плоскости скважинного прибора последовательно в положения χ_0 и $\chi_0 + 180^\circ$, мы можем исключить влияние прочих составляющих дрейфов гироскопа, вычисляя

$$\sin(A - A_0) = \frac{\omega_{x\Gamma}(\chi_0) - \omega_{x\Gamma}(\chi_0 + 180^\circ)}{2\Omega \cos \varphi}. \quad (11)$$

Эти выводы, как мы и предполагали, позволяют судить о принципиальной идентичности диаметральной и поперечной схем с точки зрения реализации компасирования, инвариантного к пусковым погрешностям ДУС. Однако здесь же необходимо подчеркнуть и отличие, имеющее не менее принципиальный характер. В поперечной схеме эффективность инвариантного компасирования максимальна в точке $A = 0^\circ (180^\circ)$, снижается до полной потери по мере достижения значения $A = 90^\circ (270^\circ)$, и повлиять на эту ситуацию никоим образом невозможно. В диаметральной схеме при одном и том же значении A_0 , определяемом формулой (7), эффективность определяется сочетанием параметров траектории (θ, A) и углов χ и ψ . Первый из них является полностью управляемым, и выполнение (8) является условием реализации алгоритма (11). Углом поворота отклонителя ψ , как уже было сказано, при нахождении инклинометра в составе КНБК управлять можно, но с невысокой точностью. Однако в «диаметральной» схеме достаточно практически произвольного (с точностью до «четверти») поворота по ψ для того, чтобы нарушить неблагоприятное сочетание параметров A и A_0 , определяемого (7), а затем вновь провести установку χ_0 в соответствии с (8) и т. д.

Конструктивные особенности реализации схемы

Проведенный сравнительный анализ поперечной и диаметральной схем был бы неполон без рассмотрения конструктивно-технических особенностей их реализации. В качестве определяющих габаритных характеристик выбраны, естественно, диаметры. При этом, чтобы уйти от жесткой привязки к конкретным изделиям, сравнению подвергаются их относительные размеры.

Разрабатываемая в данной работе конструкция, представляющая реализацию диаметральной схемы (рис. 2), сравнивается с уже существующей конструкцией, разработанной ранее и реализующей поперечную схему (рис. 3).

Коэффициенты, связывающие между собой относительные размеры в данном сравнении, получены из геометрии реальных конструкций. Соответствующими буквами на чертежах обозначены геометрические размеры сопоставимые между двумя рассматриваемыми конструкциями. Так как диаметр прибора является одним из основополагающих показателей при данной разработке, то ось, по которой расположены интересующие нас для сравнительного анализа конструкций размеры, выбрана перпендикулярной диаметральной плоскости скважинного прибора.

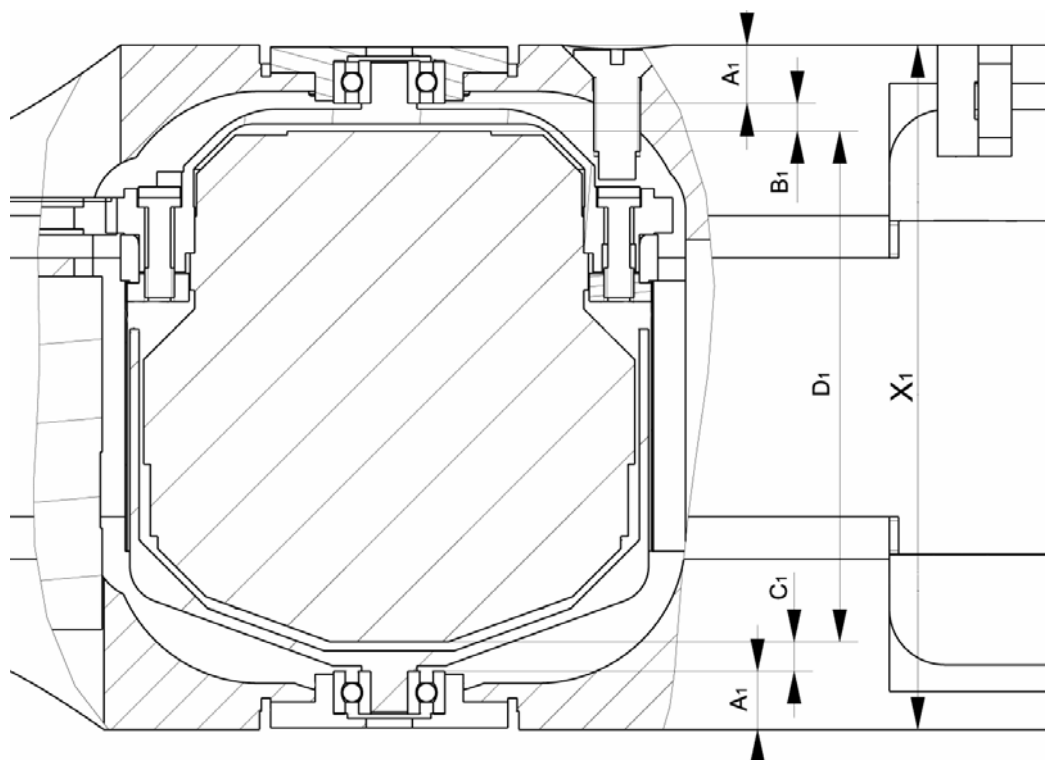


Рис. 2. Конструктивная реализация поперечной схемы

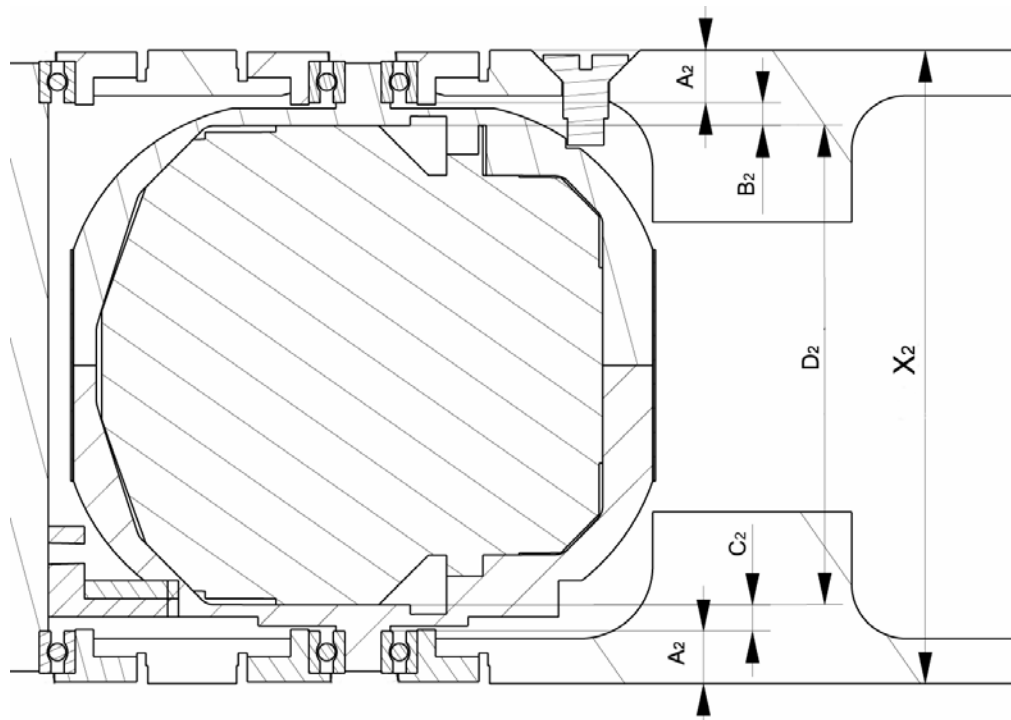


Рис. 3. Конструктивная реализация диаметральной схемы

Ниже приведены обозначения размеров и формулы, характеризующие их соотношения: X – диаметр прибора; A – размер подшипникового узла; B – толщина стенки рамки гироскопа; C – толщина стенки рамки гироскопа (другая сторона); D – размер гироскопа по измеряемой оси.

Учитывая изменение положения гироскопа в диаметральной схеме относительно поперечной, размер D можно считать основным фактором, влияющим на интересующий нас внешний диаметр прибора X . Но в связи с различными рамками, в которых на-

ходится гироскоп, а также разной формой установочных колец для подшипников, размеры A, B, C , также будут влиять на результирующий размер X .

$$\frac{A_2}{A_1} = 0.92; \quad \frac{C_2}{C_1} = 0.91;$$

$$\frac{B_2}{B_1} = 0.83 \quad \frac{D_2}{D_1} = 0.96.$$

Вычислим, насколько диаметральная схема позволяет снизить внешний диаметр прибора X_2 относительно внешнего диаметра прибора, реализующего поперечную схему X_1 .

$$\frac{X_2}{X_1} = \frac{2A_2 + B_2 + C_2 + D_2}{2A_1 + B_1 + C_1 + D_1} = \frac{2 \cdot 0.91A_1 + 0.83B_1 + 0.91C_1 + 0.96D_1}{2A_1 + B_1 + C_1 + D_1}.$$

Подставив значения размеров поперечной схемы, вычислим искомое соотношение:

$$\left(1 - \frac{X_2}{X_1}\right) \cdot 100\% = \left(1 - \frac{2 \cdot 0.91 \cdot 3.15 + 0.83 \cdot 1.5 + 0.91 \cdot 1.6 + 0.96 \cdot 27.6}{2 \cdot 3.15 + 1.5 + 1.6 + 27.6}\right) \cdot 100\% = 5.4\%.$$

Можно сделать вывод о том, что с учетом реальных возможностей конструктивного исполнения диаметальной схемы она дает снижение значения внешнего диаметра корпуса прибора относительно поперечной схемы на 5.4%.

Заключение

Диаметральная схема является результатом эволюционного развития и видоизменения схем, реализованных ранее, последней из которых является поперечная схема.

В результате анализа различных режимов работы схемы показано, что схема:

- реализует возможность совмещения непрерывного режима аналитического гироскопа направления с точечным компасированием;
- применима при любых траекториях скважин;
- обладает свойством гирокомпасирования, инвариантного к пусковым погрешностям ДУС;
- при наличии дополнительной степени свободы относительно продольной оси скважины, например при работе инклинометра в составе компоновки низа буровой колонны, осуществляет гирокомпасирование, инвариантное к пусковым погрешностям ДУС при любых траекториях скважин.

Рассматривая все преимущества диаметальной схемы, освещенные в статье, и тот факт, что данная схема дает еще и выигрыш в габаритах, можно говорить о больших перспективах внедрения рассмотренного типа гириноклинометров.

Литература

1. Биндер Я.И. Аналитическое компасирование в инклинометрии скважин малого диаметра. // Гироскопия и навигация. 2003. № 2.
2. Биндер Я.И. Универсальный гириноклинометр с ориентацией главной оси двухосного датчика угловой скорости в диаметальной плоскости скважины. // Гироскопия и навигация. 2005. № 12.
3. Биндер Я.И., Падерина Т.В. Бесплатформенный гириноклинометр с ориентацией главной оси двумерного датчика угловой скорости в плоскости поперечного сечения скважины. // Гироскопия и навигация. 2004. № 2.

ИССЛЕДОВАНИЕ АЛГОРИТМОВ ОПТИМИЗАЦИИ КОНФИГУРИРОВАНИЯ И РАСПРЕДЕЛЕНИЯ ЗАКАЗОВ ПРИ РЕШЕНИИ ЗАДАЧ ТПП В СРЕДЕ ВИРТУАЛЬНОГО ПРЕДПРИЯТИЯ

Ю.Н. Фомина

Научный руководитель – к.т.н., доцент Е.И. Яблочников

Важнейшей характеристикой современной организации ТПП является широкое использование кооперации. Предприятия, выполняющие различные работы по ТПП, рассматривают использование современных форм кооперирования с другими компаниями как одно из основных направлений укрепления своего бизнеса. Предлагаемая методология дает возможность планировать реализацию распределенного ТП.

Введение

Важнейшей характеристикой современной организации ТПП является широкое использование кооперации, когда вокруг основных предприятий существует круг партнеров-субподрядчиков. Виртуальное предприятие – это временная межпроизводственная кооперация ряда юридически независимых предприятий, которая:

- создается в короткий срок и поддерживается с помощью современных информационных технологий;
- путем интеграции возможностей участников разрабатывает и производит актуальную продукцию и услуги;
- обходится без новых юридических образований, которые заменяются гибкой координацией.

Обычная форма организации предприятия подразумевает, что производители стараются наиболее выгодно для себя подобрать заказы под имеющиеся у них ресурсы. Участники виртуального предприятия, напротив, стремятся оптимальным образом подобрать ресурсы под имеющиеся заказы. Второй вариант ведения бизнеса позволяет получить наилучшее распределение ресурсов предприятия (например, нагрузки оборудования) [2]. Использование такой организации производства, как виртуальное предприятие, позволяет получать следующие преимущества:

- позволяет находить наиболее выгодных и надежных партнеров из-за более широкой и динамичной среды виртуального предприятия;
- предоставляет большие возможности для экономии времени и стоимости проектирования и изготовления;
- способствует увеличению эффективности работы предприятий-участников за счет наиболее оптимального распределения ресурсов;
- стимулирует распространение новых технологий, которые позволяют создавать продукцию с улучшенной конструкцией и лучшего качества;
- позволяет выпускать продукцию, оптимально соответствующую требованиям покупателей, за счет более динамичной структуры виртуального предприятия.

Благодаря более низкой себестоимости производства при одновременном сокращении сроков выпуска и улучшении качества продукции виртуальные предприятия на сегодняшний день уже составляют серьезную конкуренцию крупным традиционным компаниям.

Одной из основных задач конфигурирования ВП является задача размещения заказа, подразумевающая наличие определенного объема работ (заказ), которые необходимо выполнить, использовав для этого некоторые технологические мощности (ресурсы участников виртуального предприятия – потенциальных исполнителей заказа). Для ее реализации необходимо определить механизм поиска конкретного производителя (исполнителя заказа), а также критерии выбора этой компании среди участников виртуального предприятия.

Целью работы в данном случае является выполнение оценки и выбора среди участников виртуального предприятия исполнителей заказов на ТПП изделий. Для этого необходимо предложить:

- структуру для представления данных о ТПП для их последующего анализа;
- подход к оптимальному конфигурированию и распределению заказов на решение задач ТПП в среде виртуального предприятия.

Основная часть

Особенностью проектирования ТП в распределенной среде является его многовариантность. Во-первых, заданное состояние изделия может достигаться за счет реализации различных вариантов технологического процесса. Например, комплексная операция, выполняемая на токарно-фрезерном обрабатывающем центре с ЧПУ, может заменить последовательность из нескольких токарных и фрезерных операций. Выбор того или иного варианта маршрута осуществляет обычно инженер-технолог, исходя из собственного опыта. Однако перевод ТПП изделия в среду виртуального предприятия накладывает более жесткие требования на выбор варианта ТП и его обоснование. Во-вторых, каждой задаче ТПП может соответствовать несколько потенциальных исполнителей, предлагающие различные варианты ее решения.

Таким образом, в распределенной производственной среде проектировщик создает универсальный технологический процесс, поскольку не обладает информацией о конкретных доступных возможностях, ресурсах и мощностях производителя. Этот начальный технологический процесс называется ресурсо-независимым технологическим процессом (РНТП) и лишь указывает подразумеваемые процедуры, необходимые для создания изделия. При разработке РНТП проектировщик определяет, какие именно задачи ТПП будут выполнены сторонним производителем.

Поверхности, получаемые удалением материала, могут иметь требования к точности обработки (взаимного расположения элементов детали). Операции удаления материала, имеющие высокие требования к точности, должны быть выполнены без перезакрепления, другими словами, каждая из них не может быть выполнена различными производителями. Если процессы изготовления двух отверстий имеют требования к точности взаимного расположения, отверстия развертываются за одну операцию. Подобные требования влекут за собой необходимость выполнения этих процессов в пределах одного станка, т.е. операция состоит из набора переходов, которые должны быть произведены без перезакрепления одним и тем же производителем. Поэтому увязка с конкретными ресурсами (т.е. распределение ресурсов для РНТП) должна выполняться на уровне операций.

Обычно РНТП состоит из нескольких операций. РНТП можно представить в виде разветвленного графа, вершинами которого являются операции, а ребра определяют последовательность выполнения операций. Такое представление ТП позволяет отобразить его многовариантную структуру. При этом параллельные ветви графа (множество эквивалентных вершин) соответствуют различным вариантам реализации ТП (подграфам, содержащим множество эквивалентных вершин) [3]. Для каждой операции ТП определено множество заказов Z на выполнение работ средствами стороннего предприятия.

$$Z_i^k \in Z, i = 0, \dots, n,$$

где Z_i^k – i -й заказ на выполнение ТПП; n – число заказов на выполнение ТПП k -й операции. Таким образом, РНТП представляется двухуровневым графом. Верхний уровень называется операционным уровнем графа (ОУГ), а нижний – уровнем заказов (ЗУГ). На рис. 1 а, б представлен пример изображения РНТП в виде двухуровневого графа.

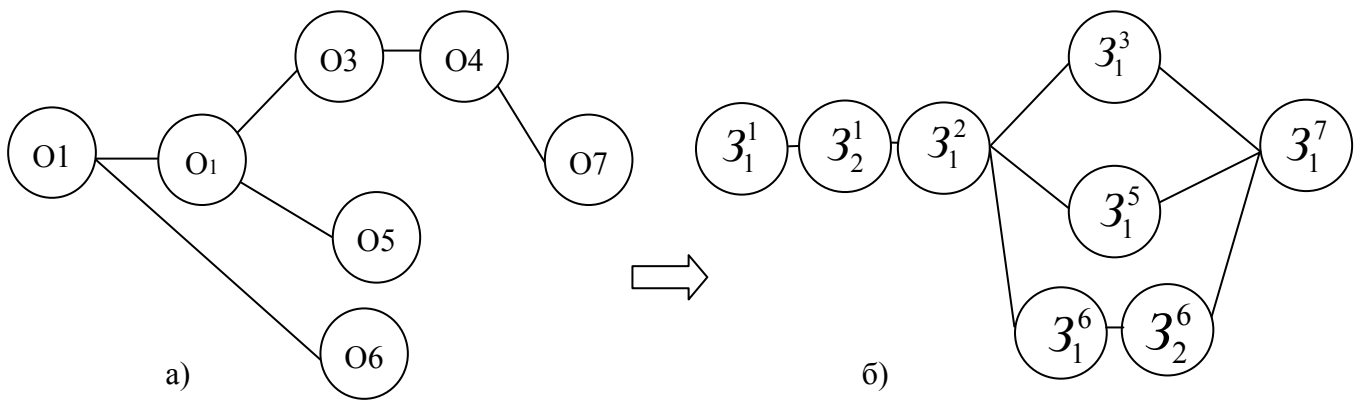


Рис. 1. Граф РНТП (а – операционный уровень, б – уровень заказов)

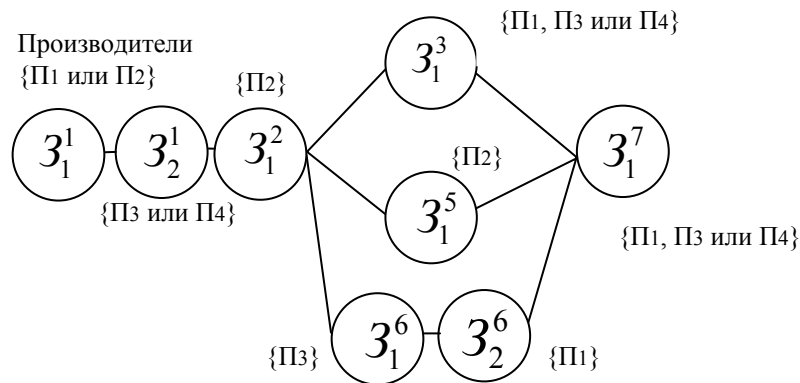


Рис. 2. Ресурсо-зависимый технологический процесс

Из рис. 1 видно, что множество заказов, соответствующих операциям, может быть пустым или включать один или несколько заказов.

Каждый узел ЗУГ определяет требования к его характеристикам, таким как тип, точность, связанные с ним геометрические элементы и другие. В соответствии с этими требованиями происходит выбор исполнителя заказа среди участников виртуального предприятия. Проектировщик составляет РНТП в соответствии с конструкцией изделия. Затем из опубликованного списка производителей выбираются конкретные производители, чьи производственные возможности и мощности соответствуют требованиям каждой операции. Следует отметить, что если поиск производителя не принес результатов, проектировщик может изменить конструкцию и/или составить новый РНТП. После выбора предприятие-заказчик передает конкретным производителям РНТП, который впоследствии становится ресурсо-зависимым технологическим процессом (РЗТП). В соответствии с этим определено множество потенциальных участников виртуального предприятия (потенциальных производителей).

$$P_i \in P, i = 0, \dots, l;$$

где P_i^k – потенциальный исполнитель i -го заказа на выполнение ТПП; l – общее число потенциальных исполнителей проекта ТПП.

Преобразование РНТП из рис. 1 в РЗТП показано на рис. 2.

На рис. 2 каждому заказу поставлены в соответствие возможные его исполнители. Следует отметить, что исполнитель может быть назначен на выполнение более чем одной операции.

Следующий шаг – выбор в ходе переговоров конкретного единственного производителя из ранее подобранных для каждого заказа. Подбор вариантов и определение характеристик выполнения заказа – предложений от потенциальных исполнителей в виде, например, приблизительных стоимости и времени выполнения заказа – осуществ-

ляется под управлением современных информационных технологий. Необходимо оценить каждого производителя, опираясь на полученные в ответ предложения на участие. На основании полученных данных на последнем этапе создается распределенный технологический процесс (РТП). РТП – это оптимальный набор производителей, каждому из которых поручается группа операций, что обеспечит проектировщику оптимальные или близкие к ним характеристики производственных параметров, таких как стоимость, качество и время изготовления. В данной работе предлагается методология оценки и оптимизации распределения задач ТПП среди потенциальных исполнителей, рассмотрены подходы к составлению окончательного РТП.

Прежде чем приступить к оценке и выбору производителей для РНТП, необходимо провести подготовку данных для сотрудничества. Процедура подготовки производится в следующем порядке:

- упрощение: исключение из РЗТП таких операционных узлов и/или изначально необоснованно выбранных производителей, которые на данный момент недоступны или считаются неэкономичными;
- расширение РЗТП за счет увеличения альтернатив для выбора производителя так, что каждый узел графа характеризовался бы единственным производителем.
- объединение: группирование операционных узлов, которые могут быть выполнены одним производителем.

Цель упрощения – исключить нереализуемые или неэкономичные операции или варианты производителей из РЗТП. Следующий этап – расширение – для расширения сокращенного РНТП путем дублирования узлов графа. Это означает, что операционный узел, включающий в себя несколько возможных производителей, дублируется столько раз, сколько существует вариантов производителей. Например, если заказ i на рис. 3а обладает m возможными вариантами производителей, то он дублируется m раз в набор $(z_{i1}^k, z_{i2}^k, \dots, z_{im}^k)$ как показано на рис. 3б.

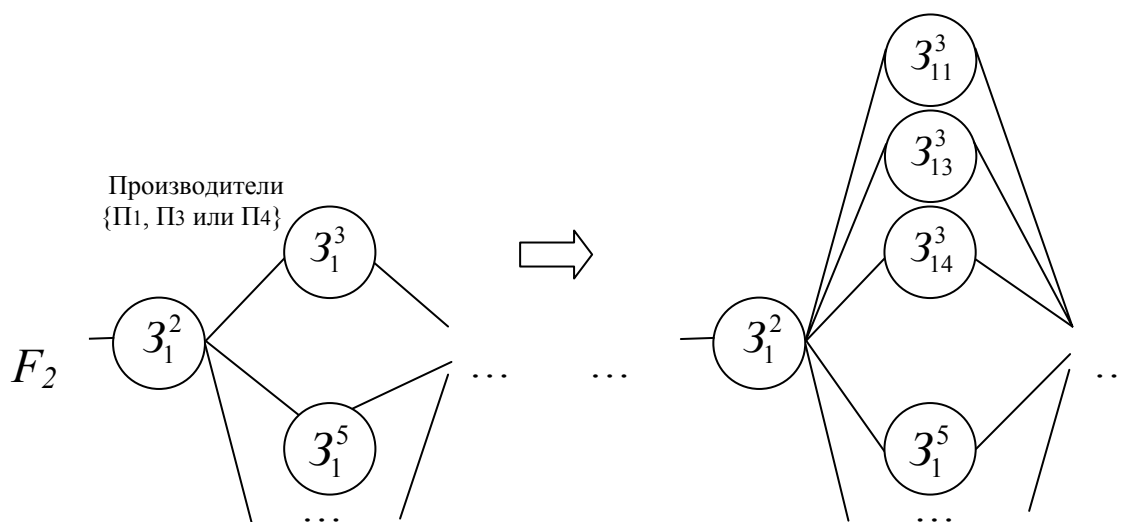


Рис. 3. Расширение РЗТП (уровень заказов графа РЗТП)

В этом случае каждая дублированная операция несет одну и ту же информацию, но предназначается разным производителям $(P_1, \dots, P_j, \dots, P_m)$. Результатом этапа расширения является увеличенный РЗТП, в котором каждый операционный узел связан только с одним производителем.

На последнем этапе производится объединение узлов ЗУГ, которые могут быть выполнены одним производителем, группируются они с целью снижения транспортных затрат и расходов на управление. После окончания подготовки данных для сотрудничества можно приступить к выбору и оценке исполнителей, т.е. к формированию РТП. Для

проведения такого анализа необходимо выделить критерии выбора оптимального решения. В самом простом случае в качестве критериев выбора оптимального варианта РЗТП могут быть использованы такие показатели, как время и стоимость выполнения заказа. При этом необходимо обследовать все множество возможных решений и выбрать то из них, которое позволяет минимизировать эти показатели.

Данная задача аналогична «задаче коммивояжера» (ЗКВ) [4, 5]. В исходном варианте ЗКВ состоит в том, чтобы по данному списку городов определить в каком порядке коммивояжер должен посетить каждый из них по одному разу, чтобы получившийся маршрут был кратчайшим из возможных или хотя бы близким к таковому. ЗКВ может быть решена несколькими методами. Вот некоторые из них:

- «алгоритм жадности» или метод поиска «ближайшего соседа»;
- метод полного перебора;
- метод «ветвей и границ»;
- метод «восхождения»;
- генетический алгоритм (ГА).

Из всех перечисленных методов наиболее длительным в плане расходования машинного времени является метод полного перебора. Наиболее выгодными являются метод «ветвей и границ» и генетические алгоритмы. Наибольший экономический эффект получается при решении ЗКВ с использованием генетических алгоритмов. Использование того или иного метода поиска оптимального решения зависит от размера пространства поиска. Если проектировщик имеет дело с небольшим ТП, содержащим не так много вариантов исполнения, то в данном случае имеет смысл воспользоваться, например, «алгоритмом жадности». Это позволит определить оптимум. Однако, если необходимо проанализировать более сложный ТП, который может включать несколько сотен операций и множество вариантов реализации ТП, а также множество потенциальных исполнителей для каждой задачи ТПП, поиск оптимума может потребовать значительных машинных ресурсов и занять продолжительное время. Стоит отметить, что в общем случае ГА не находят оптимального решения очень трудных задач, однако использование данного метода дает достаточно хороший результат.

Изложенные условия позволяют считать ГА наиболее эффективным методом для определения оптимальной последовательности производителей при решении задач ТПП в среде виртуального предприятия.

ГА, один из производно-свободных методов оптимизации, основан на понятии естественного отбора и эволюционных процессов. В отличие от методов наследования, таких как искусственные нейронные системы, в ГА используется большой объем вычислений для получения оптимального решения, что обусловлено характерными для ГА повторяющейся структурой и случайностью сочетаний. Тем не менее, непрерывная решается задача или дискретная, согласно ГА будут найдены решения с помощью характерной предметной функции, и эти решения будут наиболее оптимальными вариантами во всем пространстве решений. Благодаря этим особенностям область применения ГА широка и разнообразна, включает в себя задачи проектирования промышленного оборудования, проектирования зданий, определения параметров предпочтительной модели и оптимизации параметров процесса. После достижения критерия остановки в качестве оптимального решения отбирается вариант (квазиоптимум), обладающий лучшим сочетанием интересующих нас качеств [3].

Генетические алгоритмы обладают следующими достоинствами:

- позволяют решать задачу оптимизации с любым количеством вершин в графе ЗУГ;
- разрешают распараллелить задачу;
- допускают ограничение решения задачи как по времени, так и по заданному критерию останова;

- обладают высоким быстродействием и могут использовать решения, полученные другими методами, как начальные приближения с их последующей оптимизацией;
- имеют большое количество настроек, позволяющих изменять условия и время сходимости.

Относительно проблематики ТПП хромосома ГА соответствует варианту РЗТП. Хромосома состоит из генов, каждый ген соответствует определенному заказу РЗТП (имеется в виду уровень заказов графа РЗТП) и должен быть закодирован таким образом, чтобы однозначно описывать характеристики заказа (время и стоимость его выполнения). При этом такие данные, как конкретные заказ, исполнитель, время и стоимость выполнения заказа, являются фенотипом описываемой системы. В свою очередь, последовательность генов, содержащая эту же, но уже закодированную информацию, является генотипом данной системы. Таким образом, для решения задачи нам необходимо представить характеристики каждого заказа в форме, подходящей для использования в ГА. Все дальнейшее функционирование механизмов генетического алгоритма производится на уровне генотипа, позволяя обойтись без семантики описываемой системы. Это обстоятельство как раз и определяет широкое применение ГА в различных областях.

Генетические алгоритмы решают задачи, работая с популяцией из некоторого числа наугад взятых решений. Относительно рассматриваемой проблематики ТПП под решением понимается вариант РЗТП. Далее с помощью генетических операторов происходит изменение хромосом выбранной популяции, что позволяет выйти за ее рамки и определить оптимальное решение на всем множестве решений с хорошей степенью вероятности. К генетическим операторам относятся скрещивание (его также называют кроссовер или кроссинговер), мутация и инверсия.

Итак, если ЗКВ включает несколько сотен объектов, то ее решение является трудной задачей для современных ПК. При этом объем пространства поиска очень велик, даже лучший из выбранных вариантов может оказаться плохим. Для решения таких задач следует использовать ГА. Для реализации ГА необходимо симитировать «выживание сильнейших» хромосом, выделяя их среди остальных. Для этого требуется:

- в начальный момент времени $t = 0$ необходимо случайным образом сформировать начальную популяцию (первое поколение), состоящую из k особей: $Popul_0 = \{A_1, A_2, \dots, A_k\}$, где A_i – вариант РТП, $i \in (0, k)$;
- вычислить приспособленность каждой особи из этой популяции и популяции в целом с помощью, так называемой, фитнес-функции: $F_{A_i} = fit(A_i)$, где $i \in (0, k)$ (значение этой функции определяет, насколько хорошо подходит особь, описанная данной хромосомой, для решения задачи), $F_t = fit(B_t)$, где t – поколение популяции (первоначально $Popul_0$);
- выбрать особь A_c из популяции: $A_c = get(Popul_t)$;
- с определенной вероятностью выбрать вторую особь из популяции $A_{c1} = get(Popul_t)$ и провести операцию кроссовера: $A'_c = cros\ sin\ g(A_c, A_{c1})$;
- с определенной вероятностью выполнить операцию мутации: $A''_c = mutati\ on\ g(A_c)$;
- с определенной вероятностью выполнить операцию инверсии: $A'''_c = inversion(A_c)$;
- поместить полученную хромосому в новую популяцию: $insert\{B_{t+1}, A'''_c\}$;
- выполнить операции, начиная с пункта 3, k раз;
- увеличить номер текущей популяции (эпохи) $t = t + 1$;
- если выполнилось условие останова, то завершить работу, иначе выполнить операции, начиная с пункта 2 [6].

Теперь необходимо более подробно рассмотреть отдельные этапы алгоритма. Наибольшую роль в успешном функционировании алгоритма играет этап отбора родительских хромосом (пункты 3 и 4). При этом возможны различные варианты. Наиболее часто используется метод отбора, называемый рулеткой. «Колесо рулетки» используется для выбора из текущего поколения хромосом, которые сохранятся в следующем поколении. Другой часто используемый метод – турнирный отбор. Он заключается в том, что случайно выбирается несколько особей из популяции (обычно два) и победителем становится особь с наибольшей приспособленностью. Кроме того, в некоторых реализациях алгоритма применяется так называемая стратегия элитизма, которая заключается в том, что особи с наибольшей приспособленностью гарантированно переходят в новую популяцию. Использование элитизма обычно позволяет ускорить сходимость генетического алгоритма. Недостаток использования стратегии элитизма – в том, что повышается вероятность попадания алгоритма в локальный минимум (ЗКВ относится к многоэкстремальным задачам) [3]. Другой важный момент – определение критериев останова. Обычно в качестве них применяются или ограничение на максимальное число эпох функционирования алгоритма, или определение его сходимости, обычно путем сравнения приспособленности популяции на нескольких эпохах и остановки при стабилизации этого параметра.

Вычисление приспособленности хромосомы – это единственная часть программы, в которой необходимо иметь представление о том, что же на самом деле кодирует хромосома. Поэтому для каждой прикладной задачи необходимо писать собственную фитнес-функцию. В задаче по оптимизации распределения задач РЗТП в среде виртуального предприятия фитнес-функция будет зависеть от двух переменных (первоначально были определены два критерия выбора варианта заказа: время и стоимость его выполнения) [7]. При этом задача назначения приоритетов на критерии выбора предоставляется для выполнения проектировщику. Это обстоятельство связано с тем, что методы ведения бизнеса различных предприятий могут значительно отличаться. Например, одна компания при размещении своего заказа может руководствоваться такими приоритетами, как качество и сроки исполнения, а другая – стоимостью его реализации.

Приоритеты критериев в фитнес-функции будут отражаться в виде весовых коэффициентов. Таким образом, фитнес-функция (для двух критериев) распределения заказов среди участников виртуального предприятия будет выглядеть следующим образом:

$$F_{Ai} = \text{fit}(A_i)$$

$$F_{Ai} = \omega_1 C_{Ai} + \omega_2 T_{Ai},$$

где C_{Ai} и T_{Ai} – соответственно стоимость и время реализации совокупности заказов РЗТП среди выбранных производителей, а ω_1 и ω_2 – их весовые коэффициенты.

$$\omega_1 + \omega_2 = 1.$$

При этом стоимость и время РЗТП равняются $C_{Ai} = \sum_{n=0}^{n=\max(n)} C(Z_{in})$, $T_{Ai} = \sum_{n=0}^{n=\max(n)} T(Z_{in})$,

где $C(Z_{in})$ и $T(Z_{in})$ – соответственно, стоимость и время выполнения n -го заказа i -го варианта РЗТП, а $\max(n)$ – количество заказов в данном варианте РЗТП. Следует отметить, что n – это порядковый номер заказа в последовательности заказов, относящихся к выбранному РЗТП, а не обозначение соответствия заказа конкретной операции, как показано на рис. 1. Таким образом, искомая фитнес-функция определяется по формуле

$$F_{Ai} = \omega_1 \sum_{n=0}^{n=\max(n)} C(Z_{in}) + \omega_2 \sum_{n=0}^{n=\max(n)} T(Z_{in}).$$

Как было сказано ранее, хромосомы содержат закодированную информацию. Требуется определить, какой код необходимо использовать для описания РЗТП, т.е. найти такой вид кодирования РЗТП, который позволит по найденной хромосоме (результату решения

ЗКВ с помощью ГА) однозначно определять, какой последовательности заказов, а также их исполнителей он соответствует. Кроме того, при определении приспособленности особей популяции необходимо по коду хромосомы определить конкретные значения $C(Z_{in})$ и $T(Z_{in})$ заказа. Для этого необходимо построить матрицу приспособленности, столбцы которой будут соответствовать потенциальным исполнителям заказов РЗТП, а строки – заказам, входящим в РНТП. Матрица будет иметь размерность $X \times Y$, где X – число всех заказов оцениваемого РНТП, а Y – численность участников виртуального предприятия, задействованных в выполнении ТПП данного изделия. Ячейки матрицы будут содержать значения фитнес-функции (приспособленности) для конкретных заказов и исполнителей ($F_z = \omega_1 C(Z_n) + \omega_2 T(Z_n)$). Полученная матрица представлена на рис. 4.

	Π_1	Π_2	...	Π_y
Z_1	F_{11}	F_{12}	...	F_{1y}
Z_2	F_{21}	F_{22}	...	F_{2y}
...
Z_x	F_{1x}	z_i^s	...	F_{xy}

Рис. 4. Матрица приспособленности для анализируемой системы

Итак, хромосома РЗТП состоит из генов, каждый из которых соответствует заказу РЗТП. Гены в хромосоме должны быть расположены в той же последовательности, что и заказы в графике РЗТП (ЗУР). Для выполнения исходных требований по кодированию РЗТП ген должен состоять из номера заказа и номера исполнителя. При этом для каждого гена становится возможным определить по матрице приспособленности его значение фитнес-функции. Следует отметить, что часть ячеек матрицы будет иметь нулевые значения. Это обстоятельство связано с тем, что конкретному заказу соответствует ограниченное множество потенциальных исполнителей.

Если допустить, что число потенциальных исполнителей проекта имеет трехзначное значение, а число заказов оцениваемого проекта – четырехзначное, то для кодирования каждого гена потребуется использовать десятичный код, состоящий из семи цифр. Например, для последовательности заказов, представленных на рис. 2, хромосома может выглядеть так, как указано на рис. 5.

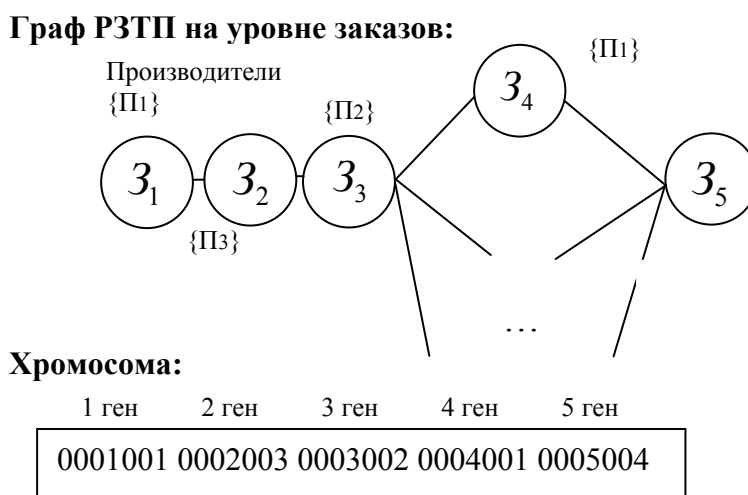


Рис. 5. Пример построения хромосомы для РЗТП

Поиск оптимального решения задачи ЗКВ с помощью ГА также включает в себя использование генетических операторов кроссинговера, мутации и инверсии. Исполь-

зование этих операций позволяет расширить область поиска и исключить возможность принятия в качестве решения задачи локального оптимума выбранной популяции. При кроссинговере хромосомы группируются в пары, и случайным образом выбирается точка кроссинговера. После этого «генетический материал» родителей смешивается. Однако если в выбранной популяции представлен не весь «генетический материал», то использование скрещивания не позволит получить недостающие гены, что исключит из анализа часть множества решений. Мутация предназначена для того, чтобы избежать подобных ситуаций и увеличить вариабельность популяции. Выполнение этой операции приводит к тому, что одна произвольно выбранная ячейка хромосомы случайным образом меняет свое значение. Кроме того, используется также оператор инверсии, который разбивает хромосому на две части и затем меняет их местами [7].

В соответствии с ГА лучшая хромосома (представляющая вариант РЗТП) получается в последнем поколении. Выбранная хромосома должна быть декодирована в РТП, для этого используется матрица приспособленности. По сути, данный процесс является обратным процессу кодирования РЗТП. Этот этап является последним, его результат – получение решения задачи КМВ для оптимизации конфигурирования и распределения заказов на решении задач ТПП в среде виртуального предприятия.

Заключение

Целью данной работы является оценка и выбор среди участников виртуального предприятия исполнителей заказов на ТПП изделий. Одной из основных задач конфигурирования ВП является задача размещения заказа. Для ее реализации необходимо определить механизм (алгоритм) поиска конкретного производителя (исполнителя заказа), а также критерии выбора этой компании среди участников виртуального предприятия.

При выполнении исследования был предложен ряд методик по оптимизации конфигурирования и распределения заказов при решении задач ТПП в среде виртуального предприятия:

- методология упрощения составления и оценки технологического процесса в структуре распределенного производства;
- метод представления РЗТП в виде двухуровневого графа, что позволяет описывать и оценивать множество заказов рассматриваемого проекта;
- методика (генетический алгоритм) оценки и выбора оптимальной последовательности производителей для реализации РТП;
- методика вычисления фитнес-функции для нахождения оптимального РТП;
- метод кодирования хромосомы РЗТП, который охватывает все множество решений при поиске оптимальной конфигурации РТП.

Предложенная методология дает возможность объединения планирования и производства в распределенной производственной среде. В конечном счете, ее можно использовать как быструю и эффективную форму виртуального производственного предприятия.

Литература

1. Schuh G., Millarg K., Goransson A. Virtuelle Fabrik: neue Marktchancen durch dynamische Netzwerke // Munchen, Wien: Carl Hanser Verlag, 1998.
2. Яблочников Е.И. Автоматизация технологической подготовки производства в приборостроении // Учебное пособие. СПб: СПбГИТМО (ТУ), 2002. 92 с.
3. Романовский И.В. Алгоритмы решения экстремальных задач./ М.: Наука, 1977. 352 с.

4. Cho H. and Wysk R.A. Intelligent workstation controller for computer integrated manufacturing: problem and models. // Journal of manufacturing Systems. 1995/14(4). P. 252–263.
5. Joo J., Park S. and Cho H. Adaptive and dynamic process planning using neural networks. // International Journal of Production Research. 2001. 39(13). P. 2923–2946.
6. Стариков А.В. Генетические алгоритмы - математический аппарат. BaseGroup Labs, 2001. 5 с.
7. Clement R.P., Wren A. Genetic Algorithms and Bus-Driver Scheduling // Presented at the 6th International Conference for Computer-Aided Transport Scheduling, Lisbon, Portugal, 1993. 9 с.

ИСПОЛЬЗОВАНИЕ CMS-СИСТЕМЫ ДЛЯ РАЗРАБОТКИ МОДУЛЯ РАСЧЕТА РЕЖИМОВ РЕЗАНИЯ

В.В. Богданов, Р.В. Бондаренко

Научный руководитель – д.т.н., профессор Д.Д. Куликов

В статье рассмотрена возможность разработки модуля расчета режимов резания в виде CMS системы

Введение

Обеспечение эффективного взаимодействия между пользователем и САПР ТП возможно осуществить на основе развитого и интуитивно понятного пользовательского интерфейса, позволяющего сделать процесс работы с ЭВМ как можно более простым. Главное преимущество использования web-приложений заключается в возможности реализации гибкого и функционального пользовательского интерфейса, основанного на получивших широкое распространение (и поэтому достаточно привычных для пользователя) программных средствах, таких как браузер или офисные приложения.

Расчет режимов резания является важной задачей на уровне проектирования перехода. Расчет заключается в определении значений скорости, подачи и глубины обработки на каждом рабочем ходу для заданного перехода, обеспечивающих требуемую точность, чистоту поверхности и устойчивую работу технологической системы при наименьшей себестоимости обработки.

В настоящее время существует немало программных продуктов для расчета режимов резания, таких как «Расчет режимов резания» фирмы Аскон (позволяет автоматически получить скорректированные по паспорту станка обороты шпинделя и подачу, скорость и мощность резания, машинное время, вспомогательное время на основной переход, энергозатраты), система расчета режимов резания КОМПАС-АВТОПРОЕКТ (функционирует в среде САПР технологических процессов КОМПАС-АВТОПРОЕКТ и предназначена для расчета режимов обработки, основного (машинного) и вспомогательного времени на основной переход), модуль расчета режимов резания системы TechnologiCS, подсистема ТехноПро/Резание (предназначена для расчета режимов резания и норм времени при проектировании операционных, маршрутно-операционных и маршрутных технологических процессов, а также для автоматизации выполнения любых расчетов, сопутствующих проектированию технологии) и другие. Все эти продукты имеют, соответственно, свои достоинства и недостатки, но ни один продукт не реализован в виде WEB-приложения.

Разработкой модуля расчета режимов резания уже давно занимается кафедра технологий приборостроения СПбГУ ИТМО. В результате многочисленных попыток удалось создать модуль расчета режимов резания для механообработки, который спроектирован на основе WEB-технологий, а в качестве модели исходных данных модуль использует и отображает XML-документ.

Реализация данной системы в виде WEB-приложения предоставляет возможность любому клиенту, имеющему браузер, полноценно работать и создавать технологические процессы посредством WEB-интерфейса. Модуль расчета режимов резания позволяет на основе общего WEB-интерфейса взаимодействовать с системой проектирования посредством использования XML-структур данных.

Решением проблемы является преобразование имеющейся системы к виду CMS-системы, которая позволяла бы добавлять и преобразовывать уже имеющиеся модули режимов резания без участия программиста, т.е. непосредственно персоналом или бизнес-аналитиком в структуре предприятия. Актуальность задачи объясняется тем, что предприятие не в состоянии организовать доработку имеющегося программного продукта в сжатые сроки в соответствии с непрерывно меняющимися техническими усло-

виями. Создание же самостоятельной CMS-системы значительно снижает сроки доработки и не требует дополнительных финансовых вложений, что крайне важно в нынешнее время.

CMS-система – модуль расчета режимов резания

Использование WEB-приложений предоставляет мощные возможности для централизованного управления распределенными системами и интеграции распределенных данных, что позволяет интегрировать разрабатываемую систему практически с любыми программными и аппаратными средствами, используемыми на предприятии. Одним из центральных вопросов, возникающих при выборе методологии построения WEB-приложений для решения задач ТПП, является выбор архитектурного подхода к созданию и интеграции приложений. От выбора архитектуры разрабатываемого приложения будут напрямую зависеть возможности масштабируемости, интеграции, функциональности и, в конечном итоге, удобство использования готового программного продукта пользователем.

Кроме того, от выбора архитектуры также напрямую зависят количество затрат на разработку и внедрение программного продукта, а также на его поддержку в процессе использования. При выборе методов реализации приложения необходимо учитывать не только возможности, которые дает та или иная архитектура проектируемого приложения, но и возможности конечного пользователя. Также следует учитывать, что программный продукт будет внедряться в уже существующую схему производственного процесса и поэтому должен иметь как можно более гибкие возможности для интеграции с другими подсистемами ТПП и системой электронного документооборота. Продуманная интеграция разрабатываемого программного модуля в существующий производственный процесс позволит значительно улучшить удобство использования и функциональность программного модуля, а также осуществлять более эффективную работу с данными, по каким-либо причинам хранимыми вне разрабатываемых приложения и базы данных. Это обеспечит переход к использованию нового программного средства в несколько этапов, не нарушая отлаженный производственный процесс, что в некоторых случаях может стать дополнительным преимуществом.

Одним из наиболее перспективных направлений на сегодняшний день является разработка CMS-систем, позволяющих управлять содержимым WEB-порталов.

CMS-система может быть применена для управления иерархической системой WEB-страниц с исходными данными, в дальнейшем CMS-система, производя расчеты, выводит результирующую WEB-страницу. CMS-система может задавать сеть WEB-страниц и последовательность их вывода.

Разработка модуля расчета режимов резания в виде CMS-системы позволит предприятию, установившему у себя такую систему, значительно снизить затраты и сроки на доработку, которые впоследствии часто бывают необходимы. Имея универсальный алгоритм работы для расчета режимов резания, в систему не нужно вносить специальных программных доработок; если возникнет необходимость внести изменения (добавить или убрать метод обработки), достаточно лишь изменить шаблон, из которого берутся входные параметры либо просто удалить шаблон для конкретной обработки.

В настоящее время существует два метода расчета режимов резания: расчетно-аналитический (по эмпирическим формулам, составление математической модели) и опытно-статистический (по справочным таблицам).

В случае использования расчетно-аналитического метода режимы резания определяются по формулам, содержащим несколько переменных, в том числе и степенных, а также поправочные коэффициенты. Несмотря на перспективность применения расчетно-аналитического метода, его использование встречает ряд трудностей:

- данные, приведенные в справочниках, являются неполными и не очень точными, так как наблюдается разброс в рекомендуемых значениях по разным источникам;
- нужно определять достаточно много переменных и показателей степени, следовательно, организовывать базу данных и поиск по ней;
- всегда найдется ситуация, для которой не будет констант и показателей степени;
- точность задания констант и показателей степени невелика, следовательно, точность результата будет незначительной.

Поэтому при автоматизированном проектировании ТП большее применение получил опытно-статистический (табличный) метод. Этот метод более прост и не содержит трудно определяемых факторов. Для этого метода имеется ряд государственных стандартов.

Для хранения и использования технологической информации посредством базы знаний существует много различных способов, основанных на использовании различных программных средств, имеющих свои достоинства и недостатки. Выбор способа организации базы знаний зависит от предъявляемых к ней требований, условий работы с ней, ее назначения, необходимости внесения в нее изменений и многих других факторов. Информацию в базе знаний можно изменять в соответствии с изменением условий производства настраиваясь, таким образом, на конкретную производственную среду.

Преимущества опытно-статистического метода:

- простота;
- малое время поиска результата;
- используется меньшее количество таблиц (по сравнению с расчетно-аналитическим методом);
- легкая адаптация при изменении производственной среды [1].

При разработке универсального алгоритма для CMS-системы использовался опытно-статистический метод, в котором осуществляется выбор режимов резания из таблиц, составленных на основе статистических данных, и корректировка результатов с помощью поправочных коэффициентов. Формирование базы «Режимы резания» основано на анализе нормативно-справочной информации по режущему инструменту станков и соответствующих этим станкам паспортов, которые имеются на конкретном предприятии. В первую очередь для занесения в базу знаний отбираются те документы, которые будут использованы для задач автоматизации.

Последовательность вычисления режимов резания – следующая.

- 1) В зависимости от заданных технологических факторов осуществляется определение табличного значения подачи на оборот S_0 (в мм).
- 2) Осуществляется определение поправочных коэффициентов на подачу, в зависимости от заданных технологических параметров.
- 3) После определения поправочных коэффициентов производится корректировка табличного значения подачи на эти коэффициенты:
 $S_p = S_0 \cdot (K_1 \cdot K_2 \cdot \dots \cdot K_i)$, где K_i – соответствующие поправочные коэффициенты.
- 4) Осуществляется определение табличного значения скорости резания V_0 (в м/мин).
- 5) Осуществляется определение поправочных коэффициентов к скорости резания в зависимости от заданных технологических параметров.
- 6) Осуществляется корректировка табличного значения скорости резания:
 $V_p = V_0 \cdot (K_1 \cdot K_2 \cdot \dots \cdot K_i)$, где K_i – соответствующие поправочные коэффициенты.
- 7) Производится расчет частоты вращения.
- 8) После определения «расчетных» значений режимов резания осуществляется их приведение к ряду подач и частот вращения выбранной модели оборудования.
- 9) После определения приведенных значений подачи и частоты вращения осуществляется расчет фактической скорости резания.
- 10) Далее осуществляется расчет основного технологического времени T_0 (мин).

Исходные данные (информация о параметрах обработки, материале заготовки, выбранном инструменте и т.д.) находятся в шаблоне (XML-документ) с входными параметрами (рис. 1). Откорректированные исходные данные обрабатываются, и при помощи простой формулы, полученной на основе универсального алгоритма, производится расчет режимов резания. Выходные данные могут быть переданы обратно в шаблон (рис. 2).

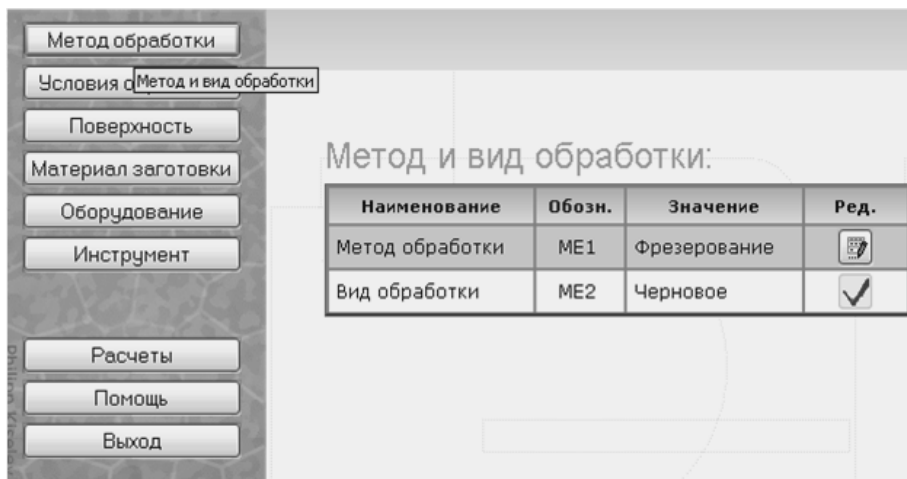


Рис. 1. XML-документ с входными параметрами

В качестве входной информации используется модель процесса в XML-документе, а результаты расчета заносятся в эту же модель. При таком подходе достаточно просто создать WEB-приложение, реализующее задачу формирования модуля расчета режимов резания.

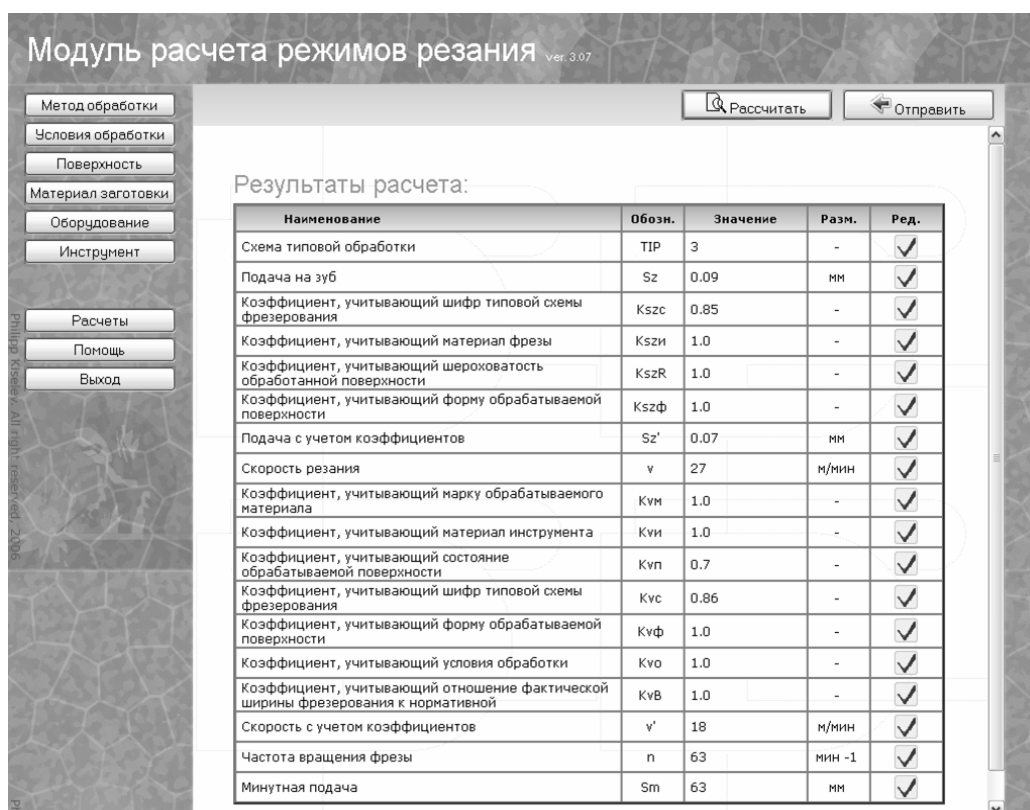


Рис. 2. XML-документ с выходными параметрами

Практическая реализация

В качестве практической реализации на данном этапе была разработана методика и комплекс алгоритмов для поставленной задачи, который впоследствии станет каркасом для доработки нынешней системы и разработки последующих. С целью максимальной адаптации и интеграции с существующими приложениями при разработке данной CMS-системы планируется использовать Microsoft Content Management Server (MCMS) и Microsoft BizTalk Server.

Content Management Server позволяет эффективно управлять информационным наполнением WEB-узла на протяжении всего его жизненного цикла, начиная с создания и коллективного редактирования документа и заканчивая его внедрением. Продукт поддерживает автоматизацию процесса публикации, персонализацию предоставления данных пользователям в зависимости от их профиля, а также отделение содержания от представления. Создание информационного наполнения и управление им в Content Management Server осуществляется с помощью WEB-интерфейса и, как и в случае многих других продуктов подобного назначения, не требует навыков WEB-дизайна или программирования. Сведения о задачах, возникающих в процессе коллективной работы над документом (например, о необходимости визирования документа), автоматически направляются пользователям по электронной почте, при этом возможен контроль версий и отслеживание путей прохождения документов. Продукт интегрируется с Microsoft SharePoint Portal Server, Microsoft Commerce Server, Microsoft SQL Server, Microsoft BizTalk Server, что позволяет использовать его в комплексных решениях масштаба предприятия [2].

Сервер BizTalk Server состоит из принимающих и передающих адаптеров, принимающих и передающих конвейеров, оркестровок, почтового ящика BizTalk Server и машины бизнес-правил [2].

Заключение

Удаленная работа с технологическими процессами посредством почти любого браузера дает возможность создавать виртуальные рабочие места, а также предоставляет средства для масштабирования подобного рода технологических систем. Одним из важных аспектов подобного рода приложений является использование стандартных коммуникационных протоколов, таким образом, обеспечивается совместимость почти любых программно-аппаратных платформ, что является плюсом при современных темпах развития информационных и промышленных технологий.

Механизм адаптации необходимо строить на основе WEB-технологий, включая адаптацию по входу и по выходу на основе аппарата шаблона и образцов. Решение этой задачи позволит ускорить адаптацию подсистем к меняющейся проблемной среде и повышающемуся уровню автоматизации подсистем. Поэтому разработка данной CMS-системы в настоящее время обеспечивает наибольшую адаптацию по входным и выходным параметрам, а также, как следствие, адаптацию к проблемной среде.

Литература

1. Киселев Ф.В. Выпускная квалификационная работа. СПбГУ ИТМО, 2007.
2. <http://www.microsoft.com>

ИНТЕГРАЦИЯ СИСТЕМ АВТОМАТИЗИРОВАННОЙ ПОДГОТОВКИ ПРОИЗВОДСТВА В ЕДИНОМ ИНФОРМАЦИОННОМ ПРОСТРАНСТВЕ

В.В. Богданов

Научный руководитель – д.т.н., профессор Д.Д. Куликов

В статье рассматриваются возможности интеграции систем автоматизированной технологической подготовки производства в единое информационное пространство предприятия при помощи современных средств обеспечения межпрограммного взаимодействия.

Введение

Современная методология автоматизации технологической и конструкторской подготовки производства подразумевает внедрение на предприятии целого комплекса программных средств, начиная от различного рода САПР и заканчивая PDM- и ERP-системами. В условиях использования значительного количества разнообразного программного обеспечения возникает необходимость в организации единого информационного пространства предприятия и интеграции в него всех используемых на предприятии средств автоматизации обработки и хранения информации. На сегодняшний день решение проблемы обеспечения интеграции программного обеспечения в единую информационную структуру предприятия является одной из важнейших задач ИТ-персонала множества предприятий. Решения, внедряемые на предприятиях, во многом основываются на предыдущем опыте сотрудников ИТ-персонала и, зачастую, не удовлетворяет требованиям, предъявляемым к современным системам автоматизированной подготовки производства. Для определения возможных путей решения проблемы интеграции автоматизированных систем подготовки производства был произведен общий анализ всех существующих на сегодняшний день методов интеграции подсистем автоматизированной подготовки производства в единое информационное пространство предприятия. Результаты анализа и предложения по применению некоторых технологий приведены в данной статье.

Постановка задачи

В условиях жестких требований рынка к сокращению сроков проектирования и подготовки производства, к повышению качества изделий необходим выход на качественно новый уровень компьютеризации конструкторской и технологической подготовки производства. Этот уровень обеспечивает применение CALS-методологии, суть которой состоит в непрерывной информационной поддержке разработчиков на всех этапах жизненного цикла изделия (ЖЦИ). Стратегическое решение задачи перехода к CALS-технологиям предполагает применение интегрированных решений по следующим направлениям [1]:

- сквозная компьютеризация всего спектра инженерных задач в проектировании и подготовке производства, с выбором базовых CAD/CAM/CAE-систем и поддержкой необходимых форматов данных для обмена конструкторской и технологической информацией;
- организация единой базы данных проекта для поддержки всех этапов ЖЦИ, компьютеризация управления проектированием и подготовкой производства на основе применения PDM-систем;
- учет фактора кооперации предприятий при работе над проектом. Применение специальных средств, поддерживающих оперативный обмен конструкторско-

технологической информацией между заказчиком и исполнителем, процессы коллективного принятия решений.

Одной из основных задач в рамках применения CALS-методологии является интеграция всех данных, получаемых с помощью различных CAD-систем. На предприятиях, где доля конструкторского проектирования с использованием CAD-систем уже весьма значительна, специалистами и специализированными фирмами в рамках крупных проектов выполняется значительный объем работ по преобразованию форматов данных и созданию системы информационного взаимодействия. Однако отработка информационного взаимодействия разных систем с использованием нейтральных форматов или прямых интерфейсов – только часть решения проблемы. Большую сложность представляет интеграция всей информации (результатов деятельности всех специалистов), с обеспечением возможности ее многократного использования. На практике по-прежнему всю информацию выводят на бумажные носители, и это объясняется, в том числе, и неготовностью участников процесса принять информацию в электронном виде, неспособностью служб технической документации управлять электронными архивами и т.д. Решение этой проблемы создатели CALS-технологий видят в применении PDM-систем [1]. Информация, создаваемая на этапе технической подготовки производства, составляет большую часть общей информации ЖЦИ. Сюда входит информация конструкторских проектов изделий основного производства и конструкторских проектов оснастки для изготовления этих изделий, информация технологических процессов изготовления изделий и технологических процессов изготовления оснастки, информация о стандартных изделиях и материалах и т.д. Вся эта информация формируется в различных системах. Большинство существующих сегодня систем разрабатывались как автономные, и вследствие этого при интеграции их с другими системами необходимы серьезные доработки с привлечением высококвалифицированных программистов. Для успешной реализации внедрения PDM-системы необходимо обеспечить возможность взаимодействия этих систем как с PDM-системой, так и между собой. Дополнительно к этому, в рамках применения CALS-методологии для автоматизации производства необходимо предусмотреть возможность кооперации предприятий при работе над проектом. Это подразумевает под собой организацию системы обмена информацией не только в рамках предприятия, но и между различными предприятиями, участвующими в проекте. Таким образом, задача может быть разделена на 2 этапа: объединение подсистем автоматизации подготовки производства внутри предприятия, и организация обмена информацией между предприятиями.

Анализ существующих решений

Процесс обеспечения интеграции информационных подсистем предприятия в единую информационную систему сильно зависит от применяемого программного и аппаратного обеспечения. Каждое предприятие уникально. Одни программные средства могут обладать большими возможностями для интеграции, чем другие. Потенциальная выгода от интеграции информационных систем предприятия обычно выше для больших предприятий. На сегодняшний день существует и успешно применяется несколько подходов к обеспечению интеграции приложений. Различия между этими типами интеграции приложений предприятия основаны на том, где применяются эти подходы. Основные стратегии и точки интеграции включают в себя следующее [2]:

- интеграция источников данных;
- интеграция прикладных интерфейсов программирования;
- интеграция бизнес-логики.

В процессе интеграции информационных систем предприятия могут одновременно использоваться несколько подходов к интеграции, если это необходимо. В некото-

рых случаях представляется невозможным использовать те или иные средства для интеграции ПО. Это зависит от программно-аппаратной среды предприятия.

Метод интеграции, основанный на использовании общих источников данных, предполагает создание единой системы обработки и хранения данных предприятия. Приложения должны хранить данные в одном из существующих хранилищ данных. Основная масса существующих на сегодняшний день приложений использует в качестве хранилища данных реляционные системы управления базами данных. Информация извлекается из источника данных автоматически, трансформируется в соответствии с потребностями бизнес-процессов и заносится в другую БД [2]. Для обеспечения успешной интеграции источников данных необходимо четко представлять себе структуру метаданных, хранящихся в БД информационных систем. Должны быть известны схемы баз данных и значение различных полей данных, которые подвергаются изменениям. Простая операция обновления данных может потребовать внесения изменений в несколько баз данных и множество таблиц. В некоторых случаях необходимыми являются некоторые представления о процессах, происходящих внутри приложения. Это затрудняет процесс интеграции. Метод интеграции, основанный на использовании общих источников данных, часто оказывается самым дешевым. Средства и методы обмена информацией хорошо развиты в современном мире информационных технологий и сравнительно дешевы по отношению к другим методам интеграции приложений предприятия.

Метод интеграции, основанный на использовании общих прикладных интерфейсов программирования (API), использует программные интерфейсы, предоставляемые различными приложениями в качестве точки интеграции. В зависимости от предоставляемых программным продуктом API бизнес-логика и процессы приложения могут быть такими же доступными, как и данные. Этот подход к интеграции наиболее подходит для больших приложений (таких, как PDM- или ERP-системы), содержащих сложную бизнес-логику и большие недокументированные или сложные для понимания базы данных [3]. Например, PDM-система SmarTeam может использовать в качестве хранилища данных различные СУБД (такие как Oracle, DB2 или Microsoft SQL Server), что дает широкие потенциальные возможности для интеграции с другими системами на основе общего источника данных. Но метаданные базы данных могут быть комплексны, что очень сильно затруднит процесс интеграции с использованием общих источников данных без потенциальных нарушений целостности БД и системы SmarTeam в целом. Поэтому система предоставляет целый набор API, которые могут использоваться в качестве точки интеграции. Значительным недостатком этого подхода является то, что процесс интеграции ограничен теми API, которые предоставляются разработчиками ПО. Одни приложения могут иметь огромный набор программных интерфейсов, предоставляющих доступ практически ко всем объектам приложения, другие же могут иметь очень ограниченный набор API. Для обеспечения своему приложению конкурентоспособности большинство разработчиков стремится соблюсти баланс в предоставлении API и доступа к логике приложения. Поэтому зачастую необходимые для интеграции API могут не входить в состав того или иного приложения. Кроме того, несмотря на то, что разработчики предоставляют API для интеграции своего приложения, эти API могут быть реализованы с помощью устаревших методов, а также быть достаточно сложными для понимания и корректного использования.

Метод интеграции, основанный на использовании бизнес-логики, обычно требует наиболее распространенных изменений в приложениях. Цель заключается в совместном использовании бизнес-логики, принятой в компании. Например, изменение какой-либо информации может затронуть несколько приложений. Используя распределенную бизнес-логику, можно создать единые методы управления информацией, которые будут распределены между всеми приложениями, которым необходимы эти функции [2].

Если происходят изменения в логической структуре приложений, то внесение изменений в программный код производится централизованно. Недостатком этого метода является то, что необходимо вносить изменения во все существующие приложения.

Предлагаемое решение

На сегодняшний день существуют различные информационные технологии, которые позволяют реализовать описанные выше методы интеграции программного обеспечения и различные их комбинации. В качестве таких технологий можно привести технологии CORBA, DCOM или SOA [3]. Исходя из того, что активы программного обеспечения и опыта его внедрения на разных предприятиях различны, требуется найти как можно более универсальные методы для обеспечения информационной интеграции программной среды предприятия. На сегодняшний день достигнуть достаточной универсальности позволяет применение технологии веб-сервисов. Веб-сервисы представляют собой технологию, основанную на применении свободно связанных компьютерных систем, которые предоставляют сервисы друг для друга. Это позволяет создать среду межпрограммного взаимодействия, которая очень слабо зависит от того, какие приложения используются в этой среде. Назовем основные особенности веб-сервисов.

- Веб-сервисы разработаны для обеспечения взаимодействия одних программных модулей (объектов) с другими. Соединенные приложения предоставляют сервисы для обмена информацией или выполнения бизнес-логики.
- Веб-сервисы используют стандарты XML (extensible markup language) для обмена информацией. Стандарт XML является универсальным средством для представления информации и обмена ею.
- Решения, основанные на технологии веб-сервисов, являются легко масштабируемыми.
- Веб-сервисы используют общепринятые стандарты internet. Это дает огромное преимущество перед другими технологиями, поскольку стандарты internet применяются во всем мире и не зависят от программной и аппаратной среды.

SOA (Service-oriented architecture, сервис-ориентированная архитектура) – распределенная вычислительная окружающая среда, разработанная, чтобы, в том числе, обеспечить возможность сравнительно быстрой адаптации и интеграции различных приложений. SOA обладает достаточной гибкостью, чтобы рассматривать элементы бизнес-процесса и основную инфраструктуру ИТ как безопасные, стандартизированные компоненты, которые могут многократно использоваться и объединяться, чтобы адекватно отображать изменяющиеся приоритеты [3]. Основной SOA является всесторонний набор программного обеспечения для управления информационными потоками и процессами обмена информацией, необходимый для координирования сервисов. В качестве примеров таких программных продуктов можно привести Microsoft BizTalk Server или IBM WebSphere Business Integration Server. Эти приложения базируются на современных программных платформах Java 2EE или Microsoft .NET и имеют встроенную поддержку для работы с веб-сервисами. Такие программные продукты имеют 2 сценария применения [3]:

- интеграция приложений внутри одного предприятия (EAI, Enterprise Application Integration);
- интеграция приложений, обслуживающих различные предприятия (B2B, Business to business integration).

На рис. 1 приведена схема, по которой реализуется процесс EAI. Каждое приложение, участвующее в процессе интеграции, связывается со службой обмена сообщениями, которая входит в состав сервера, являющегося ядром системы информационного взаимодействия предприятия (BizTalk Server, IBM BIS, и т.д.) Связь со службой со-

общений происходит посредством адаптера данных, который преобразовывает данные из формата, специфического для данного приложения, в общий формат данных сервера, построенный на основе XML. Если адаптер данных в составе используемого ПО отсутствует, он может быть разработан сотрудниками предприятия, занимающимися внедрением этого ПО, или взаимодействие со службой обмена сообщениями может быть налажено без использования адаптера данных, поскольку служба обмена сообщениями может быть представлена как веб-сервис. Разработка собственного адаптера данных не сложнее интеграции приложений через API и может быть облегчена наличием в составе сервера интеграции универсальных адаптеров данных, которые после некоторых доработок могут быть использованы в работе с конкретным приложением. Для управления очередью сообщений в состав сервера интеграции входит служба управления потоками данных. Эта служба отвечает не только за доставку данных между приложениями, но и может быть использована для реализации дополнительной бизнес-логики. Таким образом, в случае необходимости можно избежать разработки дополнительного ПО, заменив его стандартными компонентами сервера интеграции.

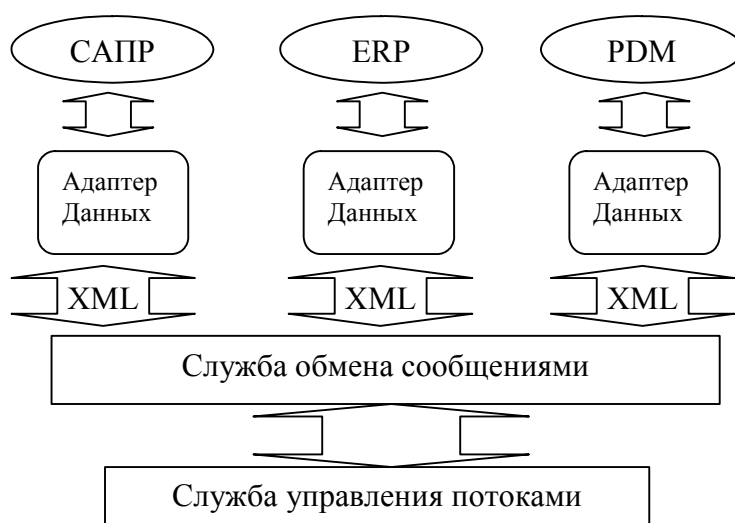


Рис. 1. Схема организации EAI

Для реализации обмена информацией между предприятиями используется аналогичная по сути схема. Сервер интеграции выступает в роли веб-сервиса, предоставляющего доступ к информационным ресурсам предприятия. Процесс обмена информацией между различными предприятиями также реализован при помощи XML. Поскольку информационные системы различных предприятий различны, то становится необходимым применение возможностей службы управления потоками данных по организации дополнительной бизнес-логики. Происходит выборка информации из информационных источников предприятия (в процессе могут использоваться хранилища данных нескольких приложений сразу), ее обработка, и передача во внешний мир информационного сообщения, отвечающего запросам стороннего предприятия. Аналогично, передаваемая сторонним предприятием информация может дробиться и заноситься в хранилища данных различных приложений.

Заключение

Предложенный метод интеграции систем подготовки производства позволяет интегрировать различные приложения, имеющие различные программные и аппаратные платформы. Метод является масштабируемым и достаточно простым в реализации за счет использования общепринятых и широко распространенных стандартов. Дополнительным преимуществом является использование в процессе интеграции разработанных

ных ведущими производителями программного обеспечения компонентов, что в значительной степени облегчает реализацию интеграции на практике. Минусом такого метода является необходимость внедрения на предприятии еще одного дорогостоящего программного продукта. При выборе решения необходимо учитывать возможности потребности конкретного предприятия.

Литература

1. Куликов Д.Д., Падун Б.С., Яблочников Е.И., Скуратов А.К., Тихонов А.Н. Методы автоматизации ТПП в приборостроении и машиностроении. СПб: СПбГУ ИТМО, 2006.
2. Roger Sessions, Janet Van Sickler. Software Fortresses: Modeling Enterprise Architectures. Addison-Wesley Press, 2003.
3. James Wilson, William Harding, Steven Baker, Jim Christensen, Scott Vidican. Microsoft® .NET Server Solutions for the Enterprise. Microsoft Press, 2004.

КОНСТРУКТОРСКО-ТЕХНОЛОГИЧЕСКОЕ МОДЕЛИРОВАНИЕ В УСЛОВИЯХ МЕЛКОСЕРИЙНОГО МАШИНОСТРОИТЕЛЬНОГО ПРОИЗВОДСТВА

**А.Р. Денисов (Костромской государственный технологический университет)
Научный руководитель – д.т.н., профессор М.Г. Левин
(Костромской государственный технологический университет)**

В статье приводятся результаты сравнительного анализа основных подходов к формированию конструкторско-технологических моделей, на основании которого определяются принципы конструкторско-технологического моделирования для условий мелкосерийного машиностроительного производства, а также структура интегрированной САПР для осуществления такого проектирования.

Введение

В современных условиях для промышленных предприятий остро стоит вопрос сокращения сроков конструкторско-технологической подготовки производства. Одним из наиболее популярных методов решения этой задачи является внедрение интегрированной САПР. Для условий мелкосерийного машиностроения такую САПР целесообразно создавать путем интеграции в единый программный комплекс систем низкого уровня («легких САПР»), например, T-Flex CAD 3D (конструкторское проектирование) и T-Flex Технология (технологическое проектирование), компании «Топ Системы». Подобные САПР предоставляют достаточный набор инструментов для решения задач проектирования в мелкосерийном машиностроительном производстве и при этом обладают наименьшей стоимостью одного рабочего места.

Основным недостатком такого подхода является возникновение временных потерь при передаче данных из одной системы в другую [1], которые можно минимизировать путем внедрения принципов конструкторско-технологического моделирования методом унификации конструкторско-технологических элементов (КТЭ) [2]. Но на сегодняшний момент не существует однозначного понимания того, как должно проходить такое моделирование. Это ведет к тому, что в последних разработках в этой области, например, в [3–6], в основном повторяется то, что было сделано ранее в 70–80 г.г. прошлого века [2, 7, 8]. Однако такие попытки внедрить конструкторско-технологическое моделирование в условия машиностроительного производства потерпели неудачу, что отчасти было связано с низким уровнем развития вычислительной техники и информационных технологий (например, отсутствие 3D-графики), а также с системными просчетами при определении принципов проектирования. Избежать повторения ситуации можно путем выявления достоинств и недостатков методов конструкторско-технологического моделирования, для чего был проведен их сравнительный анализ.

Результаты сравнительного анализа методов конструкторско-технологического моделирования

Проведенный анализ методов конструкторско-технологического моделирования показал, что каждый из них использует один из четырех подходов к решению этой задачи, которые условно можно обозначить следующим образом: классификатор ЕСКД [9], языки на основе ЕСТПП ГОСТ 14.417-81 [2, 4, 5, 7, 8], модульное проектирование [3, 6], прикладной протокол AP224 стандарта STEP ISO-10303 [10].

Базовым КТЭ в ЕСКД является само изделие, т.е. конструкторско-технологическая модель (КТМ) объекта производства представляет собой классификационный код ЕСКД. Такой код не дает полного представления ни о конструкторской модели, ни о технологическом процессе (ТП) изготовления объекта производства, од-

нако по нему определяется функциональное назначение детали, что в дальнейшем позволит выделить в ее структуре наиболее важные КТЭ и функциональные связи между ними.

Базовым стандартом для второго подхода является **ГОСТ 14.417-81 ЕСТПП** [8], в основе которого лежит язык ФТЯ-ВХ [7]. В этом языке в качестве КТЭ используются конструктивные элементы, которые могут объединяться в комплексы с использованием размерных связей и связей отношения. Таким образом, КТМ представляет собой сетевой граф, где КТЭ выступают в качестве конечных вершин, а начальные и промежуточные вершины являются комплексами элементов. На современном этапе такой подход к формированию КТМ был реализован в САПР ТП «Вертикаль» [4]. Основным недостатком подхода является невозможность определения функциональных связей между элементами.

При использовании **модульной технологии** КТМ представляют в виде совокупности модулей поверхностей (МП), каждому из которых ставятся в соответствие модули технологического процесса изготовления, представляющие собой описание технологических переходов одного установа технологической операции (ТО). Все МП распределены по трем классам: рабочие (функциональные), базирующие и специальные (вспомогательные). Это позволяет при определении последовательности обработки МП, учитывать функциональное назначение элементарных поверхностей детали, и тем самым минимизировать погрешность их изготовления. Несмотря на то, что данный подход не регламентируется ни одним стандартом, модульное проектирование большинством специалистов признается стандартом конструкторско-технологического моделирования «де-факто».

AP224 является международным стандартом и входит в группу томов стандарта STEP ISO-10303. Он регламентирует язык конструкторско-технологического моделирования на основе унификации конструктивных элементов. AP224 имеет двухуровневую классификацию КТЭ. На верхнем уровне определяются классы КТЭ (поверхности вращения, пазы, бобышки и т.п.), а на нижнем уровне – типы КТЭ. Если набор классов является фиксированным, то для определения типов КТЭ используются специальные элементы, в качестве которых выступают различные профили, направляющие и концевые элементы. Использование профилей позволяет предприятию вводить собственные КТЭ в пределах одного класса, что делает методику, предлагаемую AP224, более гибкой. Однако метод имеет три существенных недостатка:

- 1) невозможность определения функциональных связей между КТЭ;
- 2) несоответствие технологического описания AP224 и российских ГОСТов;
- 3) высокая сложность кодирования КТМ и ручного (без ЭВМ) восприятия сформированных моделей.

Проведенный сравнительный анализ показал, что наиболее «продуманным» является метод модульного проектирования, однако его недостатки ограничивают область применения среднесерийным производством. В условиях же мелкосерийного производства единственным возможным методом формирования КТМ остается проектирование на основе унификации конструктивных элементов. При этом от такого метода требуется:

- учет функциональных связей между КТЭ. Решить эту проблему можно через использование классификатора ЕСКД, где каждая группа деталей определяет набор таких связей;
- наличие удобного для проектировщика интерфейса при формировании КТМ;
- возможность быстрой адаптации существующих классификаторов (деталей и КТЭ) под условия предприятия с использованием удобного для проектировщика интерфейса;

- возможность автоматической генерации обменного файла AP 224 с описанием сформированных КТМ. Это условие необходимо при работе с иностранными партнерами.

Структура интегрированной САПР

При создании интегрированной САПР, использующей принципы конструкторско-технологического моделирования, в ее состав вводятся специальные модули, предназначенные для формирования КТМ, автоматизированного синтеза конструкторской и технологической моделей и передачи их в соответствующие «легкие» САПР посредством, например, WinAPI интерфейса, обменных файлов или запросов к серверу данных САПР. В виде компонентной диаграммы структура интегрированной САПР будет иметь вид (рис. 1).

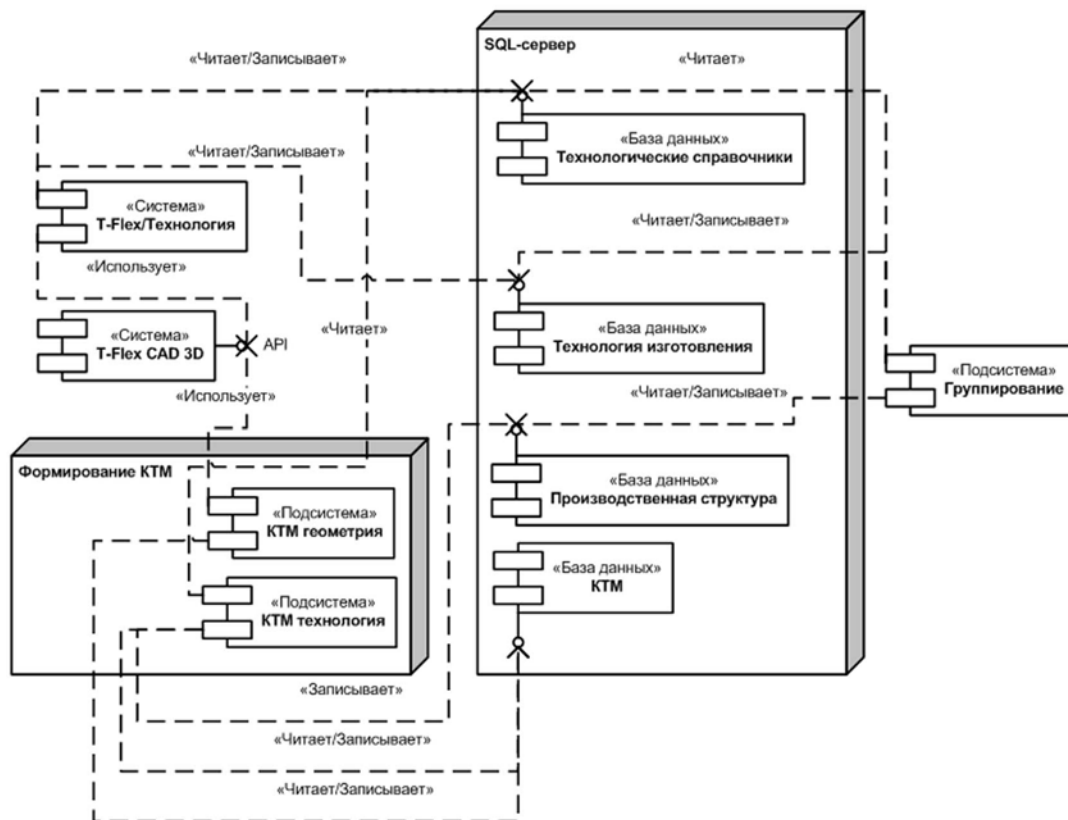


Рис. 1. Компонентная модель системы конструкторско-технологического проектирования

Синтез КТМ осуществляется в среде модуля «КТМ-геометрия» с использованием базы КТЭ, которая должна быть сформирована, исходя из потребностей предприятия. Впоследствии, при изменении номенклатуры выпускаемых изделий, база данных может быть дополнена.

При формировании КТМ в первую очередь вводится код детали по классификатору ЕСКД. Это позволяет определить функциональное назначение детали и, соответственно, рабочие поверхности в КТМ [6] и функциональные связи между ними. Таким образом, формируется первичный граф связей, который в дальнейшем необходимо дополнить базирующими и вспомогательными КТЭ. На основе полученной модели можно осуществлять технологическое проектирование (необходимое условие), однако для автоматизированного формирования конструкторской документации, а также для уменьшения числа вариантов ТП, необходимо определить связи между всеми КТЭ-

моделями (достаточное условие). При этом в первую очередь необходимо сформировать граф отношений КТЭ, а потом, если необходимо, дополнить его недостающими размерными связями, например, допусками на взаимное расположение поверхностей.

Важным этапом конструкторского проектирования является автоматизированное формирование 3D-модели детали. Автоматизированный синтез 3D-модели также осуществляется при помощи модуля «КТМ-геометрия» и затем передается в САД-систему. Качество получаемой модели напрямую зависит от полноты КТМ: в случае однозначного указания пространственного расположения КТЭ 3D-модель будет сформирована автоматически, в противном случае – потребуется ее доработка.

Дальнейшее оформление конструкторской документации осуществляется параллельно технологическому проектированию, что позволяет сократить сроки освоения новой продукции.

При выполнении технологического проектирования может быть использован любой из известных методов: на основе детали аналога, типовой, групповой и единичный методы. Если при реализации первых двух методов достаточно найти в базе данных соответствующий ТП и скорректировать его для конкретной детали, то при проектировании единичного и группового ТП предварительно необходимо синтезировать маршрут обработки. Синтез маршрута обработки осуществляется в среде модулей «КТМ-геометрия» и «Группирование».

Модуль «КТМ-технология» предназначен для автоматического синтеза графа ТП изготовления деталей по алгоритму, входными данными для которого являются КТМ детали и количество деталей в партии. В основе алгоритма лежит задача определения последовательности установов, необходимых для изготовления каждой КТЭ-детали. Для этого каждому типу КТЭ ставится в соответствие дерево установов, корнем которого является формируемый КТЭ, а листьями – КТЭ, присутствующие на заготовке (рис. 2).

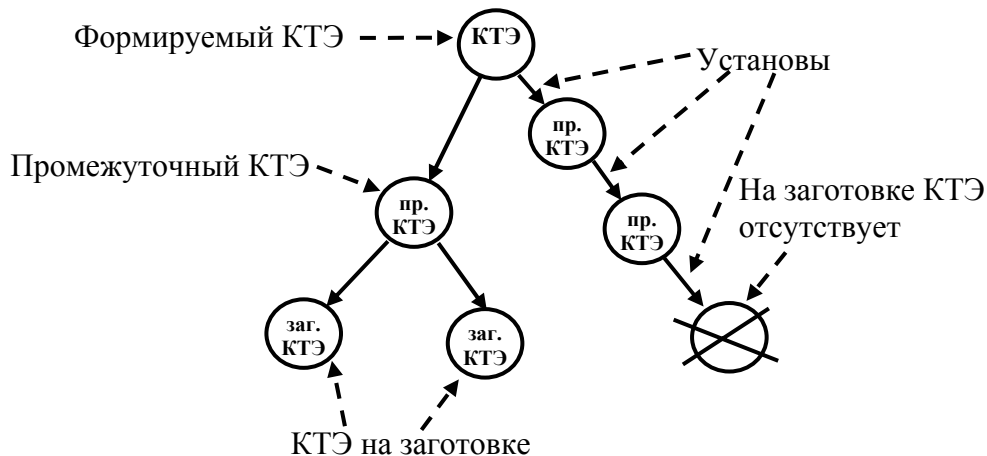


Рис. 2. Структура дерева обработки КТЭ

В качестве промежуточных вершин выступают КТЭ, полученные на промежуточных этапах, а в качестве ребер – выполняемые установы. Дополнительно каждому установу ставится в соответствие набор правил, определяющих возможность его применения для получения конкретного КТЭ, а также правил определения вариантов базирования при выполнении установов, например, [11].

Сформированный граф ТП передается в модуль «Группирование», предназначенный для решения следующих задач:

- автоматизированного синтеза операционного описания ТП изготовления изделия на производственных линиях, существующих на предприятии, и передачи его в САПР-систему;

- группирования изделий для их совместного изготовления на многономенклатурных производственных линиях;
- нормирования времени изготовления изделий, в том числе и для условий их изготовления на многономенклатурных производственных линиях;
- планирования загрузки производственных линий при изготовлении заданных объемов выпуска.

Заключение

В ходе исследования:

- проведен сравнительный анализ методов конструкторско-технологического моделирования и выявлено, что в условиях мелкосерийного машиностроительного производства целесообразно использовать формирование КТМ на основе унификации конструктивных элементов;
- определены требования, предъявляемые к системам конструкторско-технологического моделирования в условиях мелкосерийного машиностроительного производства;
- определена структура и назначение основных модулей интегрированной САПР.

Литература

1. Евгеньев Г.Б. Системология инженерных знаний: учеб. пособие для вузов // М.: МГТУ им. Н.Э. Баумана, 2001. 376 с.
2. Митрофанов С.П. Групповая технология машиностроительного производства. В 2-х т. Т.1. Организация группового производства. 3-е изд. перераб. и дополн. // Л.: Машиностроение, ЛО, 1983. 407с.
3. Боткин Ю.А., Голдовский П.С. Интегрированная САПР и модульное проектирование. // САПР и графика. 2005. №6. С. 45–50.
4. Андрейченко А. «Вертикаль» – новое поколение технологических САПР: объектный подход. // САПР и графика. 2005. №6. С. 30–35.
5. Беззуб А., Чилингаров К. «Старый новый» метод автоматизации проектирования техпроцессов. // САПР и графика. 2006. №6. С. 10–16.
6. Базров Б.М. Модульная технология в машиностроении. М.: Машиностроение, 2001. 368 с.
7. Цветков В.Д. Система автоматизации проектирования технологических процессов. М.: Машиностроение, 1972. 240 с.
8. ГОСТ 14.417-81. ЕСТПП. Проектирование автоматизированное. Входной язык для технологического проектирования. Язык описания детали.
9. ГОСТ 2.201-80 ЕСКД. Обозначение изделий и конструкторских документов
10. ISO 10303-224 Mechanical product definition for process plans using machining features.
11. Старостин В.Г., Лелюхин В.Е. Формализация процесса проектирования процессов обработки резанием. М.: Машиностроение, 1986. 136 с.

АКУСТИКО-ЭМИССИОННЫЙ МЕТОД КОНТРОЛЯ ПРИ ОЦЕНКЕ СТЕПЕНИ ДЕГРАДАЦИИ МЕХАНИЧЕСКИХ СВОЙСТВ И ОСТАТОЧНОГО РЕСУРСА РАБОТОСПОСОБНОСТИ ТРУБНЫХ СТАЛЕЙ

Н.П. Лузина

Научный руководитель – д.т.н., профессор В.Л. Ткалич

Рассмотрены возможности и перспективы акустико-эмиссионной диагностики для оценки структурного состояния металлических материалов на различных стадиях разрушения в условиях статического деформирования. Показано, что параметры акустической эмиссии (энергия, амплитудный и частотный спектры и др.) уже на стадии деформационного упрочнения позволяют оценить структурную деградацию материала (степень охрупчивания, зарождение микротрещин и т.п.), а также оценить остаточный ресурс работоспособности металлических материалов.

Введение

Зарождение и развитие трещин, пластическая деформация, коррозионное растрескивание, фазовые превращения и другие процессы динамической перестройки структуры материала сопровождаются излучением акустических (ультразвуковых) сигналов акустической эмиссии. Регистрация сигналов акустической эмиссии, определение параметров и координат источников сигналов акустической эмиссии позволяет на ранних стадиях структурных изменений идентифицировать дефекты структуры материала, контролировать скорость их развития, оценивать степень их опасности и прогнозировать остаточный ресурс работы материала и конструкции в целом.

Метод акустико-эмиссионной диагностики как метод неразрушающего контроля обладает не только более высокой производительностью (в сотни раз превосходящей производительность классических методов неразрушающего контроля, таких как рентген, ультразвук и др.), но и дает ответ на главный вопрос – имеются ли в материале конструкции потенциально опасные дефекты, которые в будущем могут стать причиной разрушения, и оценить степень их опасности, а также повысить безопасность эксплуатации, продлить ресурс работы оборудования, заменить гидравлические испытания сосудов давления пневматическими, существенно сократив сроки регламентных работ технологических установок.

Метод акустико-эмиссионного неразрушающего контроля основан на регистрации и последующей обработке параметров акустических сигналов ультразвукового диапазона, сопровождающих локальную перестройку структуры материала, зарождение и развитие микро- и макродефектов.

Метод акустической эмиссии (АЭ) реализуется в процессе активного нагружения контролируемого объекта. Для проведения АЭ диагностики к объекту контроля должны быть приложены статические и/или динамические нагрузки повышением давления при гидравлических или пневматических испытаниях, либо должны быть созданы поля напряжений механическим нагружением объекта.

Оценка степени повреждаемости конструкционных сталей при статическом и циклическом деформировании с использованием метода акустической эмиссии

В процессе пластической деформации пластичных металлических материалов наблюдается постепенное накопление и взаимодействие дефектов кристаллической решетки (вакансий, междоузельных атомов, дислокаций и дисклинаций, двойников, границ блоков и зерен и т.п.). Вследствие этого на определенной стадии деформации в материале зарождаются сначала субмикротрещины, затем микротрещины, которые,

достигнув критического размера, приводят к необратимой повреждаемости материала, развитие которой в конечном итоге приводит к катастрофическому разрушению.

Выявление степени критической повреждаемости в ответственных металлических конструкциях является в настоящее время одним из важнейших направлений в науке о прочности материалов. Особенно это важно для металлических конструкционных материалов, работающих в условиях циклических нагрузок, поскольку критический размер дефекта в этом случае чрезвычайно мал, а высокая частота нагружения может привести к быстрому его развитию и разрушению материала. В условиях циклического деформирования зарождение микротрещин чаще всего происходит при напряжениях ниже статического предела текучести за счет интенсивных процессов микропластической деформации на протяжении многих циклов нагружения в отдельных наиболее благоприятно ориентированных зернах поликристаллических материалов [6, 8, 9].

Обнаружение микротрещин (или других опасных дефектов) определенного размера связано с возможностью методов неразрушающего контроля (НК). Применение такого метода НК, как акустико-эмиссионная (АЭ) диагностика, открывает новые возможности в создании измерительных средств НК, позволяющих получить количественные характеристики структурного состояния и остаточного ресурса работы металлических конструкций [2–4, 6, 7, 11].

Принципиально новый уровень в исследованиях и контроле структурного состояния металлических материалов открывается при использовании акустико-эмиссионной диагностики неразрушающего контроля с регистрацией комплекса акустико-эмиссионных характеристик материала. В настоящей работе с использованием этого метода исследовали процессы пластической деформации и разрушения при статическом деформировании конструкционных сталей.

Исследование характеристик АЭ при статических испытаниях на растяжение

В качестве исследуемого материала были выбраны конструкционные стали 19Г (0,16–0,22 С, 0,20–0,40 Si, 0,7–1,00 Mn, $\leq 0,30$ Cr, $\leq 0,30$ Cu и $\leq 0,30$ Ni, вес. %) и 20Х13 (0,23С, 0,45Mn, 0,27Si, 0,013P, 0,008S, 12,8Cr, 0,18Ni, 0,1Cu, в %) и Cu $\leq 0,30$. Сталь 19Г используется для изготовления труб для магистральных трубопроводов, а сталь 20Х13 широко применяется в ответственных конструкциях.

На рис. 1 для примера представлена диаграмма изменения интенсивности $N_{АЭ}$ и суммарного числа сигналов АЭ N_{Σ} от времени (степени деформации), совмещенная с диаграммой напряжение–деформация при одноосном растяжении образца из стали 19Г в исходном структурном состоянии. Похожая диаграмма была получена и на образцах из стали 20Х13.

Изменение параметров АЭ коррелирует с основными стадиями деформации на диаграмме растяжения. В работе [9] предложено процесс пластической деформации до образования шейки в поликристаллических металлических металлах при статическом растяжении рассматривать состоящим из следующих стадий: микротекучести (макроупругая область, где соблюдается закон Гука), текучести (деформация Людерса-Чернова) и деформационного упрочнения. В случае отсутствия площадки текучести макроупругая область сразу переходит в стадию деформационного упрочнения.

Первый максимум интенсивности АЭ $N_{АЭ}$ регистрируется в области микротекучести (ниже предела текучести), а второй, с наивысшим пиком интенсивности, – в районе предела текучести, при этом в спектре сигналов АЭ преобладают высокочастотные компоненты (выше 400 кГц, рис. 2) [6].

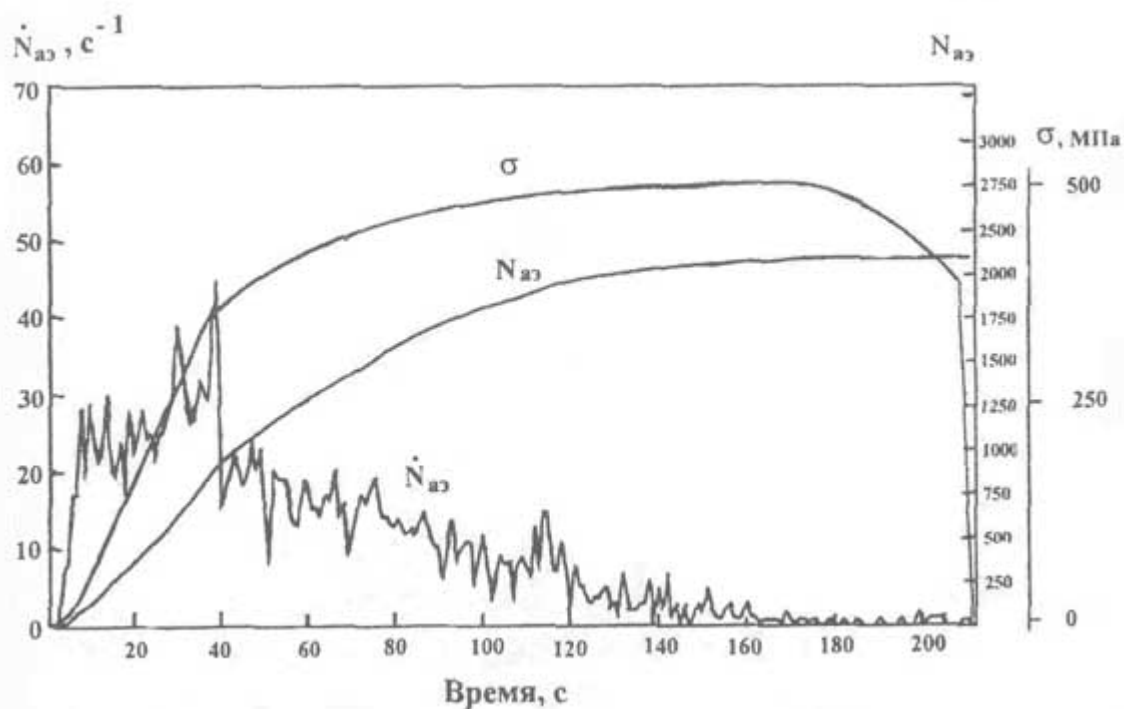


Рис. 1. Диаграммы изменения интенсивности сигналов N_{AE} и суммарного числа импульсов АЭ N_{AE} в зависимости от времени деформирования (степени деформирования) при статическом растяжении образца из стали 19Г в исходном структурном состоянии

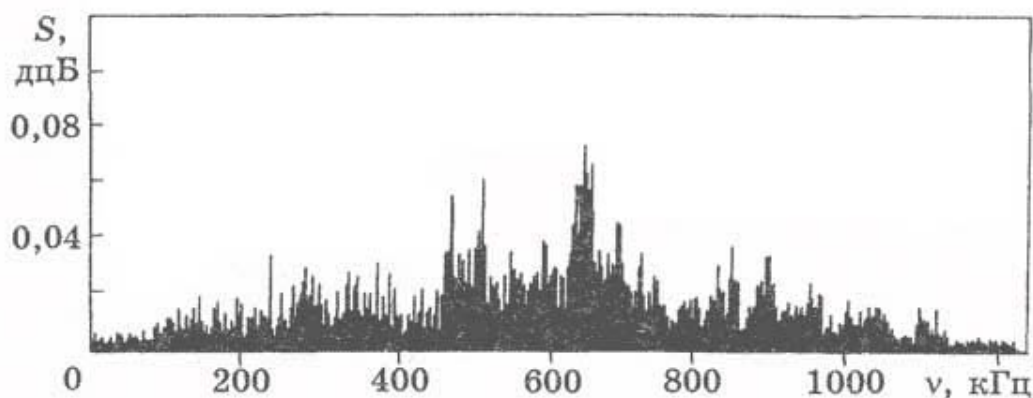


Рис. 2. Распределение спектральной плотности мощности S сигналов АЭ в зависимости от частоты сигналов при деформации в области предела текучести при статическом растяжении образца из стали 19Г в исходном структурном состоянии

Уже давно известно, что в железе, углеродистых и низколегированных сталях АЭ часто резко возрастает, начиная с напряжений меньших предела текучести. Для отожженного железа этот максимум наступает примерно при напряжении $0,7 \sigma_T$. Также известно [8, 9], что в макроупругой области у этих материалов наблюдается преимущественное пластическое течение поверхностных слоев металла.

Ряд исследований по изучению поведения деформации поверхностных слоев в области микротекучести был выполнен с использованием методов рентгеноструктурного анализ. Так, в работах [1, 5] было показано, что на зависимостях «деформация решетки – приложенное напряжение» при определенном напряжении, меньшем макроскопического предела текучести, у углеродистых сталей отмечается резкое изменение деформации кристаллической решетки, которая характеризуется параметром m^*

(рис. 3) с образованием «зуба» на этой кривой [1]. Если образец разгрузить, то возникает остаточная деформация решетки.



Рис. 3. Изменение параметра m^* с ростом напряжения в образце из углеродистой стали с 0,61 %C

В работах [8, 9] высказано предположение, что такое внезапное изменение параметров кристаллической решетки в области микротекучести в поверхностном слое углеродистых сталей может быть связано с мгновенным прохождением фронта Людерса-Чернова в поверхностных зернах поликристалла. Анализ также показал [8, 9], что движение и размножение дислокаций в этом случае преимущественно происходит в области границ зерен, а эксперименты по стравливанию поверхностного слоя показали, что этот эффект связан с более ранним пластическим деформированием поверхностных слоев металла. У сталей 20 и 09ГСФ максимум акустической активности приходится на область квазиупругих деформаций, и, по мнению авторов этих работ, он связан с поверхностными эффектами. Кроме того, в работах [8, 9] более детально рассмотрен вопрос об особенностях микропластической деформации в макроупругой области у углеродистых сталей с учетом поверхностных эффектов.

Таким образом, с учетом рассмотренных соображений, первый пик акустической эмиссии на стадии микротекучести можно связать с коллективным движением и размножением дислокаций в области границ зерен преимущественно поверхностного слоя, а также выходом их на поверхность. Зарождение новых дислокаций в основном происходит на границах зерен и микроконцентраторах напряжений непосредственно на поверхности металла. Единичный импульс акустической эмиссии в области микротекучести может произойти при одновременном прорыве примерно $4 \cdot 10^5$ дислокаций. Пики акустической эмиссии при прохождении деформации Людерса-Чернова на площадке текучести связаны с процессами прорыва дислокационной лавины во всем объеме металла на линии фронта текучести. При этом основными факторами также являются процессы выхода дислокаций на поверхность и преимущественное увеличение плотности дислокаций у границ зерен.

Для стадии деформационного упрочнения, наиболее интересной с точки зрения изменения структуры металла, характерна АЭ с преобладанием мощных дискретных импульсов. На этой стадии происходит значительное увеличение плотности дислокаций с выходом дислокационных скоплений на границы зерен и когерентные границы

фазовых включений, формированием дислокационной ячеистой структуры с критической плотностью дислокаций и образованием субмикротрещин. И.А. Одинг и Ю.П. Либеров [5] показали, что уже на стадии деформационного упрочнения существует критическое напряжение, при котором на поверхности металлов зарождаются субмикротрещины длиной порядка 1–5 мкм. Для железа это критическое напряжение при комнатной температуре соответствует пластической деформации 7–9 %. Выше этого уровня перечисленные процессы сопровождаются процессами АЭ значительной амплитуды и энергии. В распределении спектральной плотности мощности АЭ на стадии деформационного упрочнения выделяются два максимума [6]: первый в диапазоне 100–350 кГц, а второй – в диапазоне 350–800 кГц, при этом спектр простирается до 1 мГц (рис. 4).

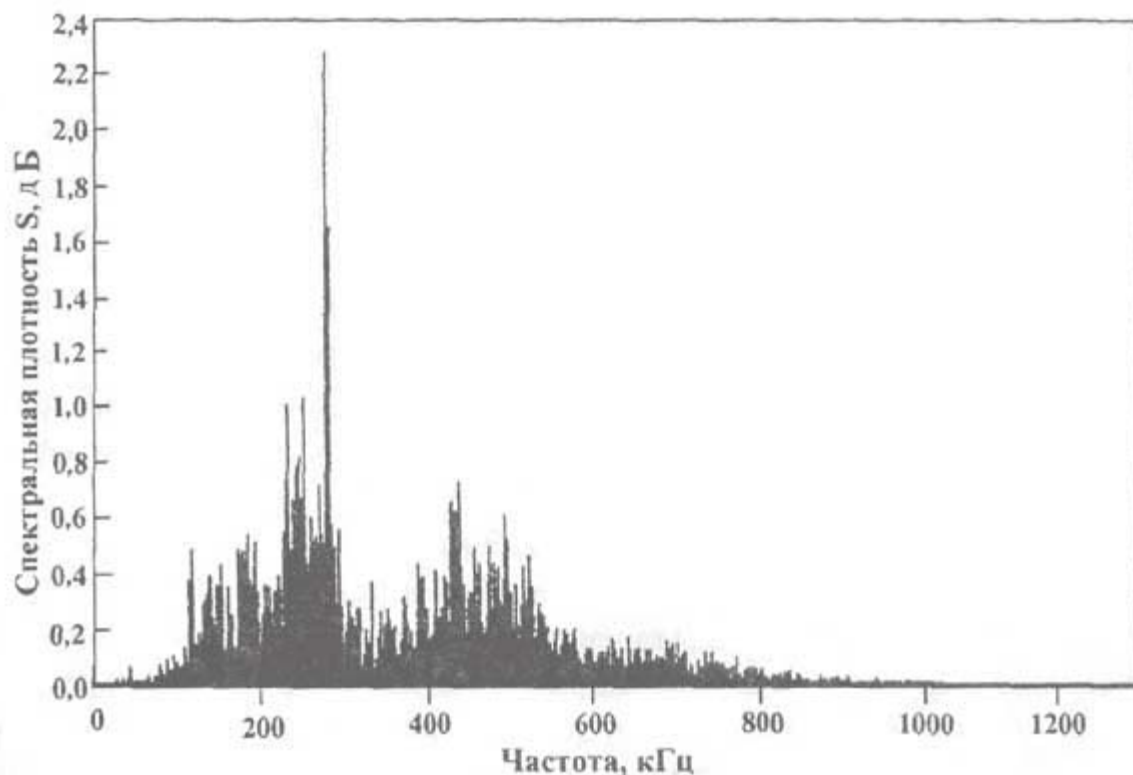


Рис. 4. Распределение спектральной плотности мощности S сигналов АЭ в зависимости от частоты на стадии деформационного упрочнения при статическом растяжении образца из стали 19Г в исходном структурном состоянии

Высокочастотный максимум спектральной плотности мощности АЭ (350–800 кГц) связан с релаксацией упругой энергии посредством увеличения плотности подвижных дислокаций, а низкочастотный (100–350 кГц) – с релаксацией упругой энергии при выходе дислокационных скоплений на межзеренные или межфазные границы, а также в момент образования микротрещин. Увеличение плотности дислокаций приводит к сокращению длины свободного пробега и плотности подвижных дислокаций (уменьшается амплитуда и энергия сигналов АЭ), а сокращение активно деформируемого объема обуславливает постепенное снижение уровня АЭ, который достигает своего минимума в момент локализации деформации (образования шейки).

Заключение

Рассмотрены возможности и перспективы акустико-эмиссионной диагностики для оценки структурного состояния металлических материалов на различных стадиях

пластической деформации и разрушения в условиях статического деформирования. Исследования проведены на образцах из конструкционных сталей 19Г и 20Х13. Показано, что параметры акустической эмиссии (энергия, амплитудный и частотные спектры и др.) на стадии деформационного упрочнения при статическом растяжении коррелируют со структурным состоянием металла и степенью его повреждаемости. Энергетические параметры, в частности, суммарная энергия акустической эмиссии, могут служить критериями структурного состояния и основой методики оценки остаточного ресурса металла.

Предлагаемая методика определения повреждаемости конструкционных сталей может быть использована для оценки остаточного ресурса ответственных металлических изделий (например, магистральных трубопроводов, сосудов давления, лопаток турбин и т.п.).

Литература

1. Faninger G. Einfluss des Gefüges auf das Verformungsverhalten unlegierter und chromlegierter Stähle. // *Harterei-Techn. Mitt.* В.22. 1967. 341–350.
2. Баранов В.М., Гриценко А.И., Карасевич А.М. Акустическая диагностика и контроль на предприятиях топливно-энергетического комплекса. М., Наука. 1998. 304 с.
3. Башков О.В., Семашко Н.А. Прогнозирование механических характеристик сплавов методом акустической эмиссии. // *Материаловедение.* 2004. С. 41–44.
4. Березин А.В., Козинкина Л.М., Рыбакова Л.М. Акустическая эмиссия и деструкция пластически деформированного металла. // *Дефектоскопия.* 2004. С. 9–14.
5. Одинг И.Л., Либеров Ю.П. Накопление дефектов и образование субмикроскопических трещин при статическом деформировании армо-железа. // *Изв. АН СССР. ОТН. Металлургия и топливо, горное дело.* 1964. С. 113–119.
6. Пенкин А.Г., Терентьев В.Ф. Оценка степени повреждаемости конструкционной стали 19Г при статическом и циклическом деформировании с использованием метода акустической эмиссии. // *Металлы.* 2004. С. 78–85.
7. Семашко Н.А., Шпорт В.И., Марьин Б.Л. Акустическая эмиссия в экспериментальном материаловедении. М., Машиностроение. 2002. 239 с.
8. Терентьев В.Ф. О пределе выносливости металлических материалов. // *Металловедение и термическая обработка металлов.* 2004. С. 15–19.
9. Терентьев В.Ф. Усталость металлических материалов. М., Наука. 2003. 254 с.
10. Терентьев В.Ф. Усталостная прочность металлов и сплавов. М., Интернет инжиниринг. 2002. 288 с.
11. Шоршоров М.Х., Гусев О.В., Пенкин А.Г. Акустическая эмиссия как метод исследования закономерностей деформации и разрушения при испытании композиционных материалов. Волокнистые и дисперсно-упрочненные композиционные материалы. М., Наука. 1976. С. 93–101.

**АЛГОРИТМЫ ИМИТАЦИОННОГО МОДЕЛИРОВАНИЯ
И СИТУАЦИОННОГО АНАЛИЗА ПРОЦЕССОВ
ПРОТИВОАВАРИЙНОГО УПРАВЛЕНИЯ СЛОЖНЫМИ
ЧЕЛОВЕКО-МАШИНЫМИ СИСТЕМАМИ
В ОБУЧАЮЩИХ СУДОВЫХ ТРЕНАЖЕРАХ****Н.А. Шарков (ФГУП «ЦНИИ им. акад. А.Н. Крылова»)****Научный руководитель – д.т.н., ст.н.с. Л.И. Ковтун
(ФГУП «ЦНИИ им. акад. А.Н. Крылова»)**

Дано описание алгоритма имитационного моделирования аварийных процессов на судах, связанных с характеристиками надежности и безопасности оборудования, ситуационный анализ которых позволяет обучаемому использовать процедуру обоснованного выбора противоаварийных управленческих решений. Алгоритм реализован в опытно-конструкторской работе, по материалам которой представлены также основные компоненты пользовательского интерфейса тренажера.

Введение

Изучение вопросов применения новых информационных технологий для профессионального обучения специалистов морской индустрии входит в область исследований судовых структурно-сложных многофункциональных человеко-машинных систем (СЧМ). Мировая практика показывает актуальность этого направления исследований, так как человеческий фактор доминирует среди всех причин техногенных аварий, экологических катастроф, потерь здоровья и жизни людей, а также дорогостоящего имущества и техники [1]. Традиционные технологии профессионального обучения позволяют достаточно эффективно моделировать физические процессы на аварийных судах. Однако они плохо учитывают особенности принятия управленческих решений человеком в аварийных условиях, которые характеризуются большими потоками переменной информации и дефицитом времени на принятие решений. Современные судовые микропроцессорные системы реального времени, контролирующие процессы, представляют только текущую информацию, но не обобщают всю прошлую и не прогнозируют будущее поведение сложных судовых комплексов, что необходимо для выработки оператором упреждающих противоаварийных воздействий [2]. Известные в инновационных технологиях экспертные системы не нашли применения в морской практике. Они требуют организации слишком сложного диалога. Чтобы осуществить поиск в базе точных данных, необходимо ответить на большое число вопросов, которые ЭВМ задает оператору. Это неприемлемо для специалиста средней квалификации в экстремальных условиях плавания при дефиците времени.

Наши предложения основаны на идеях ситуационного выбора [3]. С помощью эвристического фильтра результаты точного численного моделирования процессов преобразуются в обобщенные вербальные (словесные) и образные (когнитивно-графические) качественные оценки текущей ситуации на судне и потенциально опасные траектории ее развития. При этом учитываются ограниченные когнитивные (мыслительные) способности человека в условиях огромных потоков переменной информации и дефицита времени на принятие решений. Известно, что человек сохраняет возможность одновременного анализа не более 3–4 изменяющихся сущностей, а более 7–8 вызывает полное расстройство мыслительной деятельности. Предложенный алгоритм

ситуационного анализа обеспечивает обучаемому процедуру выбора противоаварийных управленческих решений на основе эвристического правила «реагировать прежде всего на самое потенциально опасное событие, а из всех самых опасных событий – на событие, лидирующее по времени». Указанная эвристическая процедура выбора реализуется итерационно на всех уровнях иерархически организованной системы управления судном, начиная с верхнего (базового) уровня, выбора режима использования судна. Для достижения этой цели оператору предлагается наглядная и обозримая картина опасных (критических) состояний судна, комплексов, систем, подсистем и основных агрегатов, изменяющихся в прогнозном интервале времени. Ситуационный анализ позволяет определить точки эффективного приложения управляющих воздействий для достижения целей борьбы за живучесть судна. Так определена область решения научных задач и методов для создания современных судовых тренажеров профессионального обучения операторов и лиц, принимающих решения (ЛПР).

Обоснование выбора на основе принципа сужения множества альтернативных решений – концептуальная основа ситуационного анализа структурно сложных многофункциональных человеко-машинных систем реального времени

Принятие точных и однозначных управляющих решений является функцией управления, которая, в отличие от функций регулирования, не может базироваться на предварительном описании огромного комбинаторного множества всех возмущающих воздействий и соответствующих управляющих воздействий, а также переборного алгоритма их поиска. Решение принимается для наиболее опасного варианта возможных возмущающих воздействий путем поиска наилучшего варианта противодействия возможным рискам для сложной многофункциональной системы. Основной характеристикой решения является его эффективность, т.е. степень и темп достижения целей, а также затраты ресурсов для принятия и реализации решения. Решение тем эффективнее, чем больше степень достижения целей и меньше затраты. Степень достижения целей и соответствующие затраты могут иметь как количественные, так и качественные оценки. Справедливо отмечено в [4], что «вся теория обоснования решений представляет собой неуволвимую игру качественного и количественного».

Общий подход к обоснованию управленческих решений представлен в [3–5]. Кратко изложим только те положения, которые поясняют ход последующих рассуждений и соответствующих предложений. В основе обоснования решения лежит *модель выбора*, включающая *предъявление* конечного множества основных факторов и альтернатив выбора, *механизм выбора* – установление неэквивалентности (различной ценности) их и, наконец, *правило выбора*, выражающее цель выбора в виде назначения приоритетов (шкалы оценок) выбора альтернатив. В общем виде задачу выбора можно представить в виде функции выбора C :

$$C: \{X''\} = C(X), \quad \{X''\} \subset \{X'\} \subset \{X\},$$

где $\{X\}$ – *универсальное* (полное) множество предъявления, включающее все принципиально возможные значащие факторы и альтернативы выбора; $\{X'\}$ – *допустимое* множество предъявления, выделяющее из универсального множества те факторы и альтернативы выбора, которые могут быть исполнены или соответствовать целям оперирующей стороны (например, конструктора, технолога, оператора или ЛПР) нормам или правилам, которым она придерживается; $\{X''\}$ – *предпочтительное* множество предъявления, выделяющее из всего допустимого множества только физически содержательные факторы, отвечающие физическим законам природы или условиям реальности.

Указанная функция выбора применяется в общем случае не однократно, а в многошаговых процедурах принятия решений в структурно-сложных иерархически организо-

ванных человеко-машинных системах. При этом альтернативы для выбора последовательно поступают в модель выбора на каждом новом шаге решения задачи по мере изменения возмущающих и управляющих воздействий. А максимальной эффективностью обладают те решения, которые устраняют *наибольшие потери* от возмущающих воздействий при *меньших затратах* на реализацию соответствующих управляющих воздействий.

Для разрешения проблемной ситуации, связанной со случайными возмущающими воздействиями, для которых не могут быть заданы заранее соответствующие управляющие воздействия, применяются критериальные, ситуационные или равновесные правила выбора управляющего решения. Выделим *ситуационный* подход к обоснованию управленческих решений, характерный для операторско-диспетчерской деятельности человека в сложных технических системах и комплексах машиностроения.

Выбор оператором решения из множества $\{X''\}$ еще сохраняет значительную неопределенность и сложность. Поэтому сразу осуществить выбор единственного решения невозможно без дальнейшего использования *принципа сужения множества* решений. Ограничение числа предъявлений оператору на этом шаге должно определяться когнитивными способностями человека разумно реагировать на 3–4 сущности, изменяющиеся во времени. Для этого определим некоторый набор *каталожных* (или типовых) ситуаций $\{S^c\}$. В наших задачах это пожар, затопление, воздействие морской стихии и т.п., а также структурные уровни описания и дискретные оценки опасности событий. Тогда в соответствии с [4] текущую *ситуацию* s^t из полного списка каталожных ситуаций можно представить выражением

$$s^t \equiv s_i^c \Leftrightarrow x(s_i^c) \succ x(s_j^c), \quad \forall j \neq i, \quad x \in \{X''\}, s_i^c \in \{S^c\}.$$

Это выражение читается: «как только устанавливается тождество текущей ситуации s^t любой каталожной s_i^c из полного множества ситуаций $\{S^c\}$, так сразу (и обратно) выбирается решение $x(s_i^c)$, представляющее собой пару « i -ситуация – j -решение», которая становится предпочтительней любого другого решения $x(s_j^c)$ из множества всех решений $\{X''\}$. Множество каталожных ситуаций состоит из подмножеств исходных предаварийных режимов использования судна, тактико-технических свойств судна, возможных режимов использования судовых комплексов, систем, подсистем, сложных агрегатов, оборудования, конструкций, трубопроводов, кабелей и т.п. Выбор эффективных решений осуществляется последовательно на всех уровнях иерархической организации системы управления, начиная с базового и далее вниз. Движение по цепи критических неблагоприятных событий осуществляется до выявления причины, следствием которой стала опасная ситуация на старшем уровне управления. Предпочтительное решение, устраняющее причину опасной ситуации, выбирается по принципу доступности и минимального времени реализации альтернатив выбора.

Такой подход не дает строго оптимального решения, но обеспечивает выбор рационального решения, учитывая динамическую изменчивость критериев оценок решения и необходимость ранжирования приоритетов решения. Так в общем, виде выглядит логика вывода в продукционной системе. По сути, это и есть механизм ситуационного выбора, лежащий в основе действия экспертной системы. Однако мы предлагаем не хранить в базе данных все пары возможных «ситуаций-решений», а вычислять процессы их возникновения с помощью логико-динамической имитационной модели комбинаторного типа, подобной [6], затем осуществлять их эвристическую фильтрацию в логико-лингвистические переменные и применять правило ситуационного выбора с помощью когнитивной компьютерной графики.

**Логико-динамическая имитационная модель комбинационного типа
для описания функционирования судовых структурно-сложных организационно-
технических систем в экстремальных условиях**

Поставим задачу моделирования потоков неблагоприятных событий повреждения и отказов техники в структурно-сложных системах в виде импульсных однородных процессов и выделения среди них *потока* наиболее опасных *критических событий*. Указанные потоки событий должны быть ранжированы по степени опасности или важности. Примем положение, что отказы на верхнем иерархическом уровне описания системы представляют максимальную *опасность* и, следовательно, имеют первостепенную *важность* для принятия мер по их недопущению или ликвидации, если они случились.

На рис. 1 представлен фрагмент четырехуровневой логико-динамической модели комбинационного типа для описания функционирования структурно-сложных иерархически организованных судовых систем (на примере системы электроснабжения).

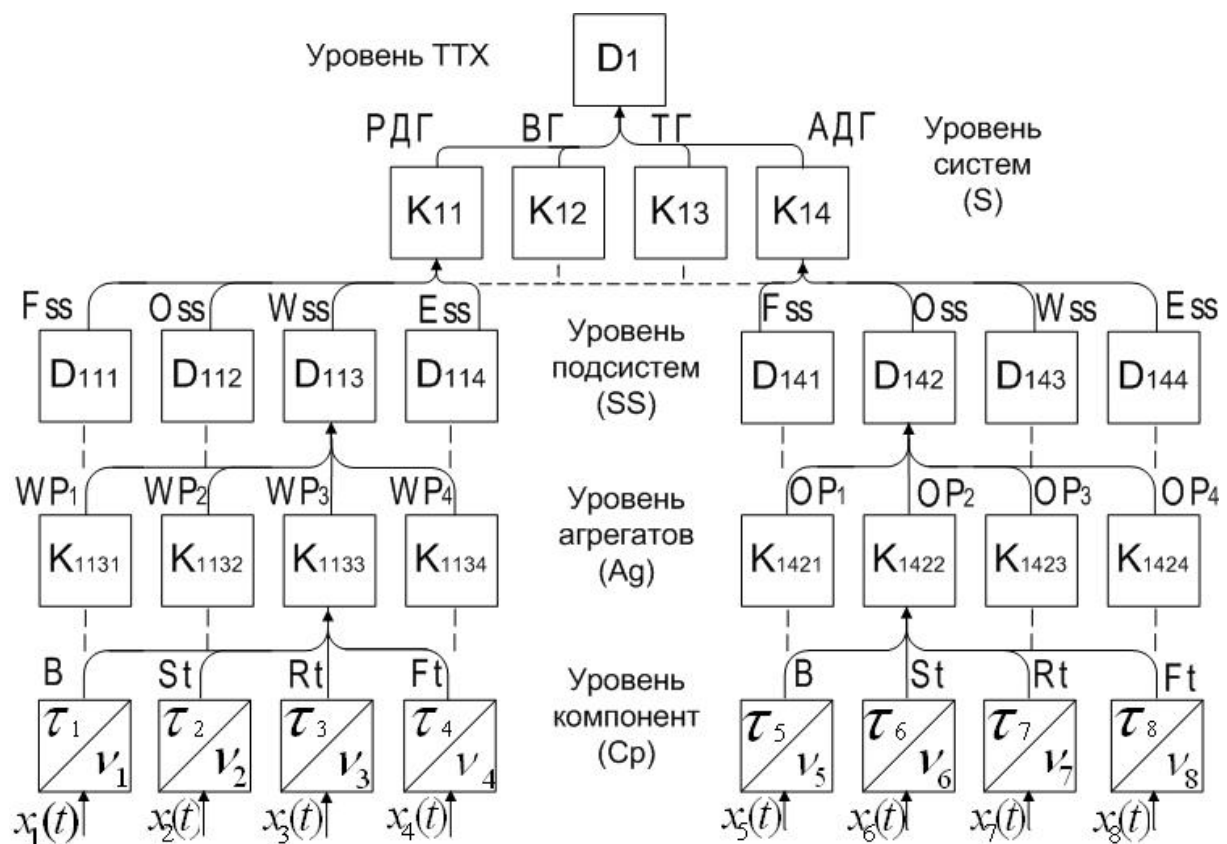


Рис. 1. Четырехуровневая логико-динамическая модель комбинационного типа

На рисунке использованы следующие обозначения: K – многоходовые конъюнкты; D – многоходовые дизъюнкты; $X(t)$ – входные импульсные процессы в виде потоков однородных случайных событий, отражающих временные последовательности интервалов работоспособного и неработоспособного состояний агрегатов, где τ – постоянная времени потери работоспособного состояния компонент; ν – постоянная времени восстановления компоненты в работоспособное состояние; $\lambda = 1/\tau$ – интенсивность отказов; $\mu = 1/\nu$ – интенсивность восстановления. Здесь также: РДГ – система резервных дизель-генераторов; АДГ – система аварийных дизель-генераторов; ВГ – валогенераторы; ТГ – турбогенераторы; FSS – топливная подсистема; OSS – масляная подсистема; WSS – подсистема водяного охлаждения; ESS – подсистема электроснаб-

жения; WP – насос охлаждения; OP – масляный насос; St – статор; Rt – ротор; B – подшипниковый узел; Ft – гидроарматура.

K – многовходовый конъюнктор – n -местный логический оператор, преобразующий воздействия $x_1(t), x_2(t), \dots, x_n(t)$ в реакцию $y(t)$ в соответствии с правилом n -местной конъюнкции:

$$y(t) = x_1 \& x_2 \& \dots \& x_n.$$

D – многовходовой дизъюнктор – m -местный логический оператор, преобразующий воздействия $x_1(t), x_2(t), \dots, x_m(t)$ в реакцию $y(t)$ в соответствии с правилом m -местной дизъюнкции:

$$y(t) = x_1 \vee x_2 \vee \dots \vee x_m.$$

Входные импульсные процессы однородных событий могут быть следствием как внутренних отказов в соответствии с характеристиками надежности техники, так и внешних воздействий, связанных с повреждениями от пожаров, затоплений, воздействия морской стихии и т.п. Выходные импульсные процессы событий-последствий на верхних уровнях иерархии системы определяются с помощью решения многомерных логических матриц в соответствии с [9, 10].

В общем случае функцию времени работоспособного состояния системы электрооборудования, в зависимости от интервалов времени работоспособного состояния систем, подсистем и агрегатов, можно записать в следующем виде:

$$\begin{cases} t_0 = \max_{j=1..4} [t_j^s] = \max_{j=1..4} \{ \min_{i=1..n} [t_{ij}^{ss}] \} \\ t_s = \min_{i=1..4} [t_i^{ss}] = \min_{i=1..n} \{ \max_{j=1..m} [t_{ij}^{Ag}] \} \end{cases}$$

Анализ системы сводится к поиску реакции логической схемы на задаваемые входные воздействия в дискретные моменты времени. Таким образом, можно зафиксировать потоки критических событий – отказов техники на всех уровнях описания от подчиненного до старшего. При этом цепочки событий отказов от нижнего до верхнего уровня, связанных причинно-следственными отношениями, представляют максимальную опасность и определяются как критические ситуации (сценарии).

Указанные уравнения позволяют перейти от методов булевой алгебры к бесконечнозначной логике. Теория бесконечнозначной логики разработана в [6] и вводит функции, областью определения которых является множество действительных чисел, в нашем случае – это временные характеристики неблагоприятных событий. Основные операции бесконечнозначной логики – n -местная конъюнкция и m -местная дизъюнкция – позволяют определять минимальные и максимальные значения временных интервалов событий на всех иерархически организованных уровнях структурно-сложных систем.

Блок-схема алгоритма процесса профессионального обучения судовых операторов

Блок-схема алгоритма процесса обучения командного состава экипажа по управлению процессом борьбы за живучесть судна представлена на рис. 2.

Здесь модули и блоки имеют следующее назначение.

- 1) *База данных* содержит совокупность данных из предметной области профессиональных знаний о судне, экипаже и мореплавании.
- 2) *Модуль формирования заданий обучения* осуществляет выбор учебных упражнений и генерацию исходных наборов аварийных событий.
- 3) *Инициализация первичных и последующих мероприятий по борьбе за живучесть судна*, когда обучающийся подает команду на управляющее воздействие, регист-

рирует полноту и время выполнения действий на постах и командных пунктах своего заведования.

4) *Имитационное численное моделирование процессов* развития, локализации и ликвидации аварийных событий (состояний сложной организационно-технической системы), в том числе:

- а) поражающих факторов пожаров, затоплений, загазованности помещений и т.п.;
- б) действий экипажа по борьбе за живучесть судна и условий жизнедеятельности на постах и в других помещениях;
- в) потока отказов и повреждений оборудования технических средств;
- г) изменения основных функций или важнейших свойств судна (хода, управляемости, энергоснабжения, плавучести, остойчивости, пожаробезопасности, обитаемости и т.п.).

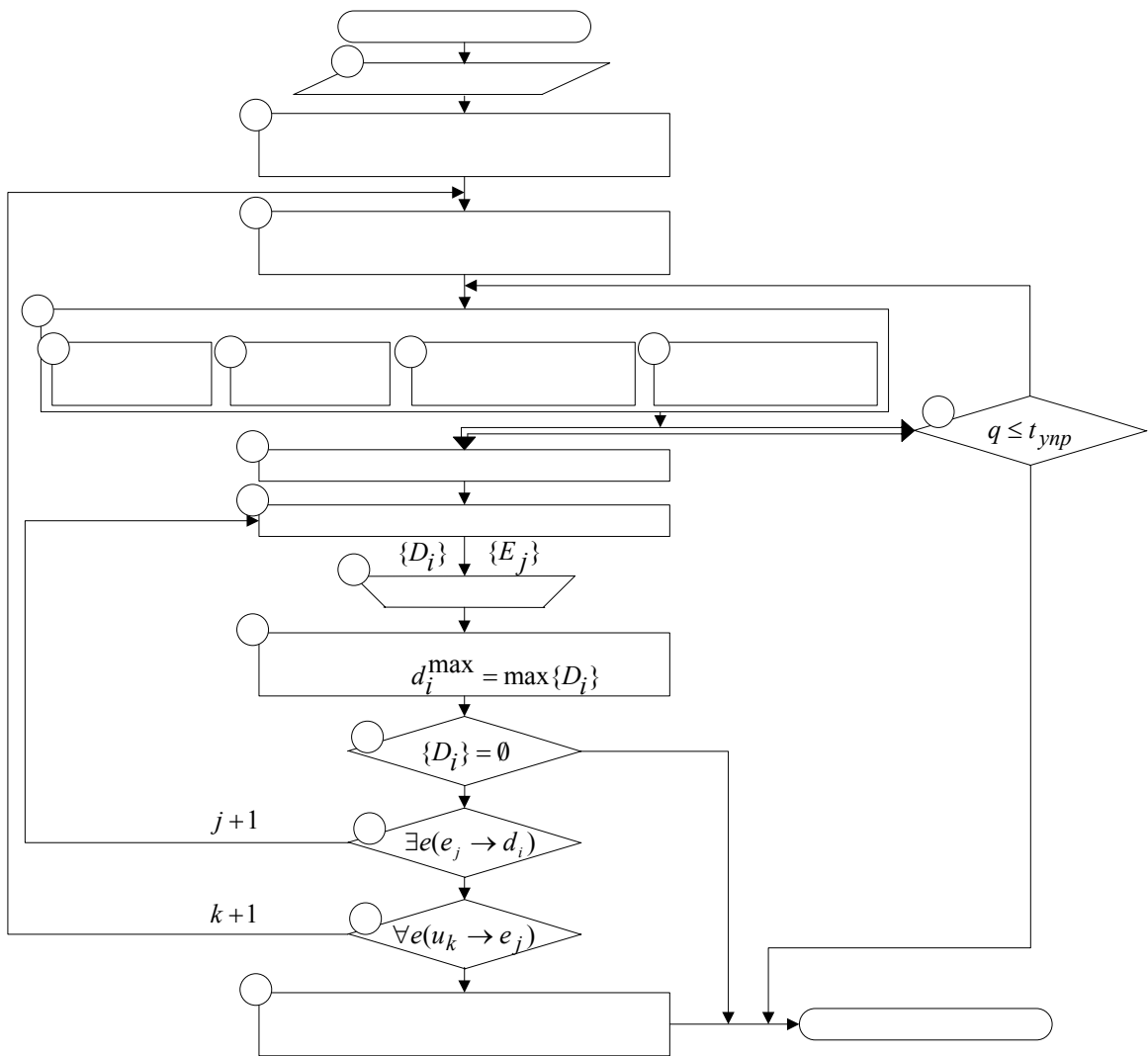


Рис. 2. Блок-схема алгоритма процесса обучения командного состава экипажа

Здесь решаются указанные задачи на основе математического описания процессов системой разностных уравнений в обобщенной матрично-векторной форме:

$$Y_{q+1} = F(Y_q, X_q, U_q, t_{q+1}),$$

где X_q – вектор входных переменных, Y_q – вектор выходных переменных, U_q – вектор управляющих воздействий, t_{q+1} – дискретный момент времени на $(q+1)$ интервале.

5) Эвристический фильтр для преобразования (редуцирования) точных численных характеристик результатов имитационного моделирования процессов в неточные (ка-

чественные) оценки «угроз» с помощью реляционных и булевских операторов, обеспечивающих переход из метрического в топологическое пространство:

$$C_q = \begin{cases} 1, & \text{if } G_q < X_q < H_q \\ 0, & \text{else} \end{cases} \quad \text{– вектор входных логических переменных в двоичной форме,}$$

$$D_q = \begin{cases} 1, & \text{if } A_q < Y_q < B_q \\ 0, & \text{else} \end{cases} \quad \text{– вектор выходных логических переменных в двоичной форме,}$$

где A_q, B_q, G_q, H_q – векторные константы чисел, характеризующие предельные значения переменных параметров процессов, при которых соответствующие утверждения с оценками угроз истинны или ложны.

6) Ситуационный анализ процессов развития, локализации и ликвидации аварийных событий с помощью логико-лингвистических отношений в понятиях «угроза-событие-команда». Здесь выделяются причинно-следственные логические связи между множествами неблагоприятных событий судовых объектов, а именно:

$\{E_{jq}\} \rightarrow \{D_{iq}\}$ – множество неблагоприятных событий $\{E_{jq}\}$ j -корабельных объектов, следствием которых являются переходы в опасные состояния (угрозы) i -объекта старшего уровня $\{D_{iq}\}$;

$\{U_{kq}\} \rightarrow \{E_{jq}\}$ – множество управляющих команд $\{U_{kq}\}$, следствием которых является устранение опасных событий $\{E_{jq}\}$ j -корабельных объектов.

7) ЛПР – лицо, принимающее решение (обучающийся), выделяет лидирующую угрозу.

8) Определение лидирующей угрозы из всего множества угроз.

9) Условный оператор, контролирующий наличие неустранимых угроз.

10) Условный оператор, реализующий логический вывод истинности или ложности составного суждения, включающего *критическое «событие»* (e_j) и выделенную *лидирующую «угрозу»* (d_i) в виде *квантора существования*. Он отражает заключение: если «существует по крайней мере одно *критическое «событие»* (e_j)», изменение времени которого влечет за собой, соответственно, изменение времени наступления или окончания выделенной *лидирующей «угрозы»* (d_i), то необходимо перейти к следующему оператору обработки выделенного критического события (e_j):

$$\exists e(e_j \rightarrow d_i).$$

В противном случае необходимо перейти к обработке критического события следующего уровня описания судна, т.е. организуется цикл ($j+1$).

11) Условный оператор, реализующий логический вывод истинности или ложности составного суждения, включающего *рекомендуемую «команду»* (U_k) и выделенное *критическое «событие»* (e_j) в виде *квантора общности*. Он отражает необходимость применения всех *рекомендуемых «команд»* (U_k) для выделенных *критических «событий»* (e_j), которые приводят к увеличению времени наступления и/или сокращения времени существования текущей *лидирующей «угрозы»* (d_j), т.е. организует цикл ($k+1$):

$$\forall e(U_k \rightarrow e_j).$$

В противном случае перейти к модулю выделения следующей лидирующей угрозы.

12) Модуль диагностики обучения с оценкой выполненных упражнений.

13) Условный оператор, контролирующий превышение установленного времени на выполнение заданного упражнения.

Имитационное численное моделирование процессов развития, локализации и ликвидации аварийных событий, а также изменения состояния сложной организационно-технической системы «судно-экипаж» является основой решателя задач. Моделирование обеспечивает возможность проведения многовариантных расчетов детерминированных и статистических характеристик процессов в ограниченной окрестности многомерного пространства состояний системы. При этом требуется точная детализация исходных условий и параметров имитируемых процессов. Имитация процессов развития аварийных происшествий включает в себя совместное алгоритмическое описание физических процессов на корабле и поведения корабельных специалистов на основе известных математических методов и приемов [4–9].

Ситуационный анализ результатов имитационного моделирования аварийных процессов включает в себя логико-лингвистические (качественные) оценки и правила логических выводов (эвристические процедуры) на множестве аварийных событий.

Таким образом, предлагается осуществлять прием итерационного изменения плана решения: выбор решения определять сначала по грубому эвристическому плану ситуационных моделей, а затем уточнять по исходному на базе точных численных оценок имитационного моделирования процессов.

Семантической основой ситуационных моделей решателя задач являются следующие понятия: «объект», «субъект», «угроза», «событие» и «команда». Определим их следующим образом.

Корабельные объекты и субъекты структурированы по функциональным связям и отношениям подчиненности (как части и целое) на трех иерархических уровнях описания корабля:

- важнейших свойств судна как целостного неструктурированного объекта;
- основных судовых систем, обеспечивающих сохранение важнейших свойств судна 1-го уровня описания;
- оборудования, агрегатов и конструкций, обеспечивающих функционирование основных судовых систем 2-го уровня описания.

«Угроза» – опасное состояние объекта, которое характеризуется несколькими уровнями опасности (в простейшем случае двумя). Угроза характеризуется также временем наступления и временем прекращения и может быть представлена как лингвистическая переменная «угроза (d)», имеющая, например, пять степеней опасности: 1 – нулевая (0); 2 – минимальная (min); 3 – средняя (mid); 4 – максимальная (max); 5 – недопустимая (∞).

Тогда прогнозируемая аварийная ситуация на судне может быть представлена совокупностью распределенных во времени угроз для различных судовых объектов. На этом множестве существуют наиболее опасные по уровню и времени сочетания угроз (критические пути или траектории), определение которых и поиск способов минимизации времени их существования, является формальной задачей решателя задач. Опасные состояния (угрозы) объектов старшего уровня являются следствием опасных состояний (угроз) объектов подчиненного уровня, т.е. связаны каузальным отношением как следствие и причина, что в логике предикатов отображается логической связью, именуемой *импликация*.

В общем случае момент возникновения «угрозы (d)» может зависеть не от одного, а некоторой совокупности «событий (e)» с различными моментами наступления. Тогда возникает задача поиска *критических* событий, воздействия на которые приводит к

увеличению времени наступления (удалению фронта угрозы) или уменьшению времени существования (сокращению временного интервала) угрозы. *Критическое «событие (e)»* – это прогнозируемое или планируемое событие, реализация которого влияет на общее время наступления или нахождения соответствующего *критического объекта* в состоянии *угрозы*. По мере локализации аварийной ситуации переменная «критическое событие» обновляется до тех пор, пока не закончится весь список событий, влияющих на соответствующую угрозу.

Определим понятие «команда» как задание (или направление) управляющего воздействия, которое устраняет событие, являющееся причиной опасного состояния объекта. Таким образом, лингвистическая переменная «команда (u)» включает список инверсированных лингвистических констант, представленных в переменной «событие(e)». Если среди всех событий определены критические, то соответствующие направления управляющих воздействий определим как *рекомендуемые команды*. Команды в зависимости от уровня иерархической организации управления сложной системой могут инициировать ряд дополнительных команд, которые отличаются большей степенью детализации и конкретизации. Подобно отношениям «событие (e)» и «угроза (d)», лингвистические переменные «команда (u)» и «событие (e)» так же связаны отношением, как причина и следствие, что отображается логической связью *импликация*.

Основные компоненты когнитивного пользовательского интерфейса в обучающих судовых тренажерах

Представленные на рис. 3 основные компоненты пользовательского интерфейса тренажера в основном окне дают в виде структуры граф-дерева текстовое описание технического состояния всех судовых комплексов, систем, подсистем и основного оборудования, значимых с точки зрения борьбы за живучесть судна. Во вспомогательном окне представлена диаграмма потенциальных угроз, связанная с качественными оценками состояния этих объектов, в прогнозном (часовом) интервале времени (с дискретным варьируемым шагом от 1 до 3 минут). При этом учитываются отношения включения судовых объектов (по типу дерева) и причинно-следственные связи опасных событий во времени с помощью визуальных образов, подобных используемым в «Проводнике».

Дерево состояний объектов судна позволяет обучаемому получать точную информацию в фиксированные моменты времени, а диаграмма потенциальных угроз – наблюдать общую картину поведения многоуровневого объекта во времени в терминах опасного состояния его компонентов. Управление временем наблюдения может раскрывать возможные события в будущем или происшедшие в прошлом (историю событий). Основное информационное поле диаграммы потенциальных угроз представляет расцветченные индикаторы-кнопки по вызову текстовой информации о состоянии судовых комплексов, систем и оборудования в выбранные оператором дискретные интервалы времени. Переменная цветовая раскраска индикаторов-кнопок соответствует требованиям ГОСТ «Цвета обозначения систем и выделения информации в образцах техники».

Диаграмма потенциальных угроз представляет результаты эвристического преобразования численных оценок аварийных процессов в качественные оценки опасного состояния соответствующих объектов судна. Можно сказать, что происходит переход от численной меры «больше-меньше» к оценкам «хуже-лучше» в смысле опасности. Это позволяет свести все разнообразные характеристики аварийных процессов к одной мере, имеющей знакомый зрительный образ цветов опасности/безопасности: красный, желтый и зеленый. Поиск эффективных решений человеком идет значительно быстрее, когда он имеет возможность дополнительно манипулировать изображениями образов реальных процессов. Предложенный когнитивный механизм поиска управленческих

решений не исключает, а лишь дополняет точную численную оценку найденных решений с помощью последующего имитационного моделирования их эффективности.

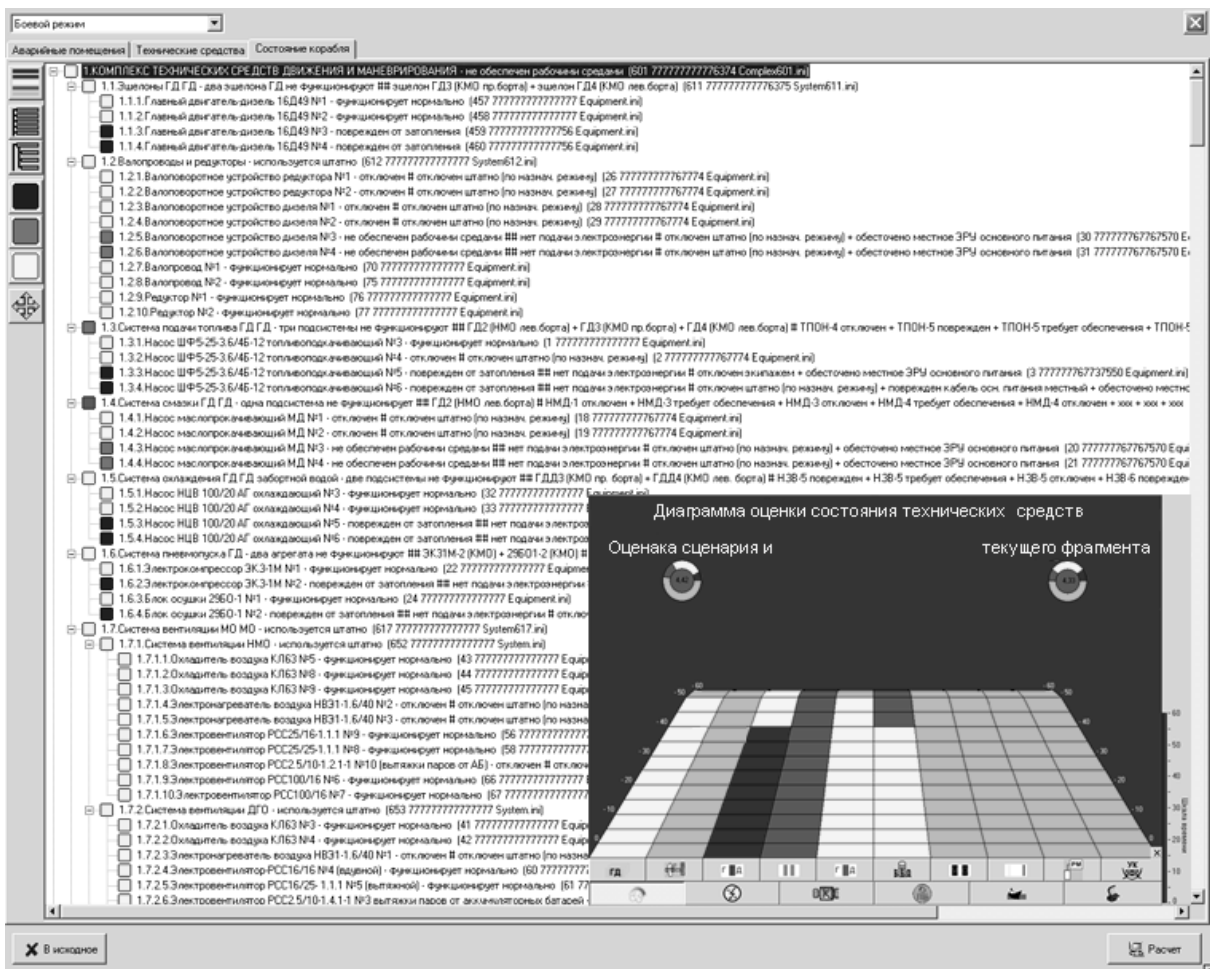


Рис. 3. Когнитивный пользовательский интерфейс

Диаграмма потенциальных угроз представляет лишь грубый сценарий прогнозируемых состояний аварийного судна. Размерность этого пространства существенно снижена, что позволяет оператору на первых этапах анализа, не вдаваясь в подробности, выделить наиболее опасные направления развития аварии и сконцентрировать на них внимание ЛПР.

Заключение

Рассмотренные алгоритмы имитационного моделирования и ситуационного анализа процессов противоаварийного управления сложными человеко-машинными системами реализованы в опытно-конструкторской разработке обучающего судового тренажера в 2006 г. Комплекс прикладных программ разработан для ПК офисного класса с характеристиками не хуже Pentium IV (3 ГГц, RAM 512 Mb, HDD 20Gb), монитор с разрешением 1280x1024 пикселей. Операционная система ПК – Windows XP. Инструментально-программная среда разработки – Borland Delphi 7.

Литература

1. A Nautical Institute project sponsored by Lloyd's Register. The international Maritime Human Element Bulletin «Alert» // №1–3, 2003–2004. (Бюллетень по учету человеческого фактора в морской индустрии).

2. Гузанов А.Н., Ковтун Л.И. Морские компьютерные системы управления техническими средствами судов: Учебное пособие // СПб: СПбГУВК, 2001. 127 с.
3. Поспелов Д.А. Ситуационное управление: теория и практика. М.: Наука, 1986. 288 с.
4. Захаров И.Г. Обоснование выбора. Теория практики. СПб: Судостроение, 2006. 528 с.
5. Казиев В.М. Введение в анализ, синтез и моделирование систем. // www.intuit.ru
6. Левин В.И. Структурно-логические методы исследования сложных систем с применением ЭВМ. // М.: Наука, 1987. 304 с.
7. Ковтун Л.И. Современные информационные технологии в процессах управления борьбой за живучесть корабля. // Судостроение. 2002. № 3. С. 20–24.
8. Ковтун Л.И., Семенов С.В. Интеллектуальная технология автоматизированного обучения управлению аварийным кораблем на морских компьютерных тренажерах. // Судостроение. 2004. № 6. С. 23–27.
9. Рябинин И.А. Основы теории и расчета надежности судовых электроэнергетических систем. Л.: Судостроение, 1967. 362 с.
10. Супрун Г.Ф. Синтез систем электроэнергетики судов. Л.: Судостроение, 1972. 326 с.

МАРКОВСКИЕ ПРОЦЕССЫ И ПРОСТЕЙШИЕ МОДЕЛИ ТЕОРИИ МАССОВОГО ОБСЛУЖИВАНИЯ. КОМПЬЮТЕРНОЕ МОДЕЛИРОВАНИЕ ПРОСТЕЙШИХ МОДЕЛЕЙ МАССОВОГО ОБСЛУЖИВАНИЯ

А.Я. Ремицкая

Научный руководитель – к.ф.-м.н., доцент И.А. Суслина

Целью работы является моделирование марковских процессов и систем массового обслуживания. Были реализованы простейшие модели теории массового обслуживания. Ценность работы – в возможности использования полученных результатов для выяснения оптимальности реальных систем массового обслуживания, обладающих свойством марковости.

Введение

На практике часто приходится сталкиваться с системами, предназначенными для многократного использования при решении однотипных задач. Возникающие при этом процессы получили название процессов обслуживания, а системы – систем массового обслуживания. Примерами таких систем являются телефонные системы, ремонтные мастерские, вычислительные комплексы, билетные кассы, магазины, парикмахерские и т.п.

В данной работе на базе марковских процессов моделируются системы массового обслуживания, при этом используются процессы рождения и гибели, учитываются вероятностные характеристики рассматриваемых процессов.

Теоретически системы массового обслуживания уже достаточно хорошо изучены. Одной из первых и наиболее известных задач в теории массового обслуживания является «Задача о вырождении фамилий» (более научное название – «Парадокс ветвящихся процессов»). В первой половине прошлого века было замечено следующее интересное явление: некоторые знаменитые аристократические фамилии постепенно исчезали. Эту проблему с математической точки зрения изучали еще в 1845 г. И.Ж. Бьенеме и в 1873 г. де Коделье. Ветвящиеся цепочки фамилий стали первым примером случайного ветвящегося процесса. Впоследствии теория массового обслуживания была применена к решению и других интересных вопросов. Сегодня важной темой изучения является управление и анализ сетей массового обслуживания с переменной структурой. Например, исследованиями в этой области занимается И.Е. Тананко, кандидат физ.-мат. наук, доцент кафедры системного анализа и автоматического управления Саратовского государственного университета.

Особой задачей является реализация некоторых моделей систем массового обслуживания в программном виде. Моделирование выполняется в среде MATLAB. Модели сопровождаются графическим отображением результатов. Представляет большой интерес практическое применение моделей. В наше время тематика компьютерного моделирования систем массового обслуживания чрезвычайно актуальна. Такие модели существенно упрощают исследование сетей массового обслуживания, достаточно ясно дают представление о состоянии рассматриваемой системы, позволяют оптимизировать их работу.

Теория массового обслуживания

Предметом теории массового обслуживания является построение математических моделей, связывающих заданные условия работы систем массового обслуживания (число каналов, их производительность, характер потока заявок и т.п.) с показателями эффективности системы, описывающими ее способность справляться с потоками заявок. При моделировании таких систем мы работаем с процессами, обладающими марковским свойством. Это значит, что прошлая история показательно распределенной

случайной величины при фиксированном настоящем не играет никакой роли в предсказании ее будущего. Это свойство также называется свойством отсутствия последствия.

В моделируемую систему в случайные моменты поступают заявки, такие поступления независимы и показательно распределены. Поток этих заявок называется простейшим. Общепринято, что функционирование систем массового обслуживания определяют 4 следующих компонента.

(1) Входящий поток требований. Заявки поступают в систему обычно не регулярно, а случайно, образуя так называемый случайный поток заявок (требований). Требования могут поступать в систему как по одному, так и по группам. Характеризует такой поток величина λ , называемая интенсивностью поступления требований в систему.

(2) Время обслуживания каждого требования. Это случайное время, которое тратится на обслуживание заявки. Поток характеризуется величиной μ , называемой интенсивностью обслуживания.

Случайный характер потока заявок и времени обслуживания приводит к тому, что система оказывается загруженной неравномерно: в какие-то периоды времени скапливается очень большое количество заявок (они либо становятся в очередь, либо покидают систему необслуженными), в другие же периоды система работает с недогрузкой или простаивает.

(3) Число обслуживающих устройств. Каждая система массового обслуживания состоит из определенного числа обслуживающих единиц (приборов, устройств, пунктов, станций и др.), которые будем называть каналами обслуживания. Каналами могут быть линии связи, рабочие точки, вычислительные машины, продавцы и др. По числу каналов системы подразделяют на одноканальные и многоканальные.

(4) Алгоритм обслуживания. В понятие алгоритма (организации) обслуживания входят:

- (а) возможность образования очереди в системе;
- (б) ограничение на длину очереди (объем накопителя);
- (в) очередность (дисциплина) обслуживания (т.е. правила, в соответствии с которыми выбирается из очереди требование для обслуживания);
- (г) правило, в соответствии с которым требования занимают свободные приборы (если их больше одного);
- (д) некоторые особенности работы обслуживающих приборов (отдыхающие приборы, надежные приборы и др.).

Многообразие вариантов выполнения программ и разработки алгоритмов обслуживания порождает чрезвычайно много различных видов систем массового обслуживания. Их делят на несколько основных типов (классов): это системы массового обслуживания с ожиданием (очередью), с отказами (потерями), с ограниченными очередями и, наконец, открытые и замкнутые системы массового обслуживания. Например, в системе массового обслуживания с отказами заявка, поступившая в момент, когда все каналы заняты, получает отказ, покидает систему и в дальнейшем процессе обслуживания не участвует. В системе массового обслуживания с ожиданием заявка, пришедшая в момент, когда все каналы заняты, не уходит, а становится в очередь на обслуживание.

Для обозначения различных систем обслуживания принята кодировка Кендалла: $A|B|m|n$. В этой записи символ A – код функции распределения случайной длины интервала между двумя последовательно поступившими требованиями входного потока; B – код функции распределения случайного времени обслуживания каким-либо прибором в системе; m – число обслуживающих приборов в системе; n – число мест в системе для ожидания. Если число мест в системе для ожидания не ограничено ($n=\infty$), то кодировка Кендалла записывается в виде $A|B|m$. Для кодировки систем массового об-

служивания замкнутого типа употребляют еще пятый символ S , означающий число возможных источников требований, поступивших в систему [1].

В качестве показателей эффективности систем массового обслуживания используются: среднее число заявок, обслуживаемых в единицу времени; среднее число заявок в очереди; среднее время ожидания обслуживания; вероятность отказа в обслуживании без ожидания; вероятность того, что число заявок в очереди превысит определенное значение, и т.п.

Вероятностные процессы в системах массового обслуживания

Процесс работы системы массового обслуживания представляет собой случайный процесс.

Процесс называется процессом с дискретными состояниями, если его возможные состояния S_1, S_2, S_3, \dots можно заранее перечислить, а переход системы из состояния в состояние происходит мгновенно (скачком). Процесс называется процессом с непрерывным временем, если моменты возможных переходов системы из состояния в состояние не фиксированы заранее, а случайны.

Процесс работы системы массового обслуживания представляет собой случайный процесс с дискретными состояниями и непрерывным временем. Это означает, что состояние системы меняется скачком в случайные моменты появления каких-то событий (например, прихода новой заявки, окончания обслуживания и т.п.). Математический анализ работы систем массового обслуживания существенно упрощается, если процесс этой работы – марковский. Случайный процесс называется марковским или случайным процессом без последействия, если для любого момента времени t_0 вероятностные характеристики процесса в будущем зависят только от его состояния в данный момент t_0 и не зависят от того, когда и как система пришла в это состояние [2].

При анализе случайных процессов с дискретными состояниями удобно пользоваться геометрической схемой – так называемым графом состояний. Под состоянием системы обслуживания понимают значения различных числовых характеристик, связанных с вызовами, находящимися в данный момент времени в системе (т.е. поступившими в систему, но еще не обслуженными). Изучение случайных процессов, описывающих поведение во времени этих характеристик, составляет основное содержание математической теории систем с очередями.

В теории массового обслуживания широко распространен специальный класс случайных процессов – так называемые процессы гибели и рождения. Предназначенные первоначально для описания эволюции численности биологических популяций, процессы рождения и гибели оказались хорошей моделью явлений, происходящих в простейших (марковских) системах обслуживания. В случайные моменты времени в систему массового обслуживания приходят требования (рождение), и в случайные моменты времени заканчивается обслуживание прибывших ранее требований (гибель). Граф состояний процесса гибели и рождения имеет вид, показанный на рис. 1.

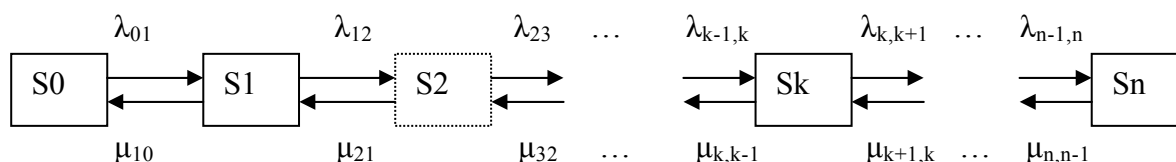


Рис. 1. Граф состояний процесса гибели и рождения

Здесь: S_k (где $k=0, 1, \dots$) – состояния процесса; $\lambda_{k-1,k}$ – интенсивность рождения, $\mu_{k,k-1}$ – интенсивность гибели.

Вернемся к рассматриваемым процессам. Для того чтобы определить представляющие интерес случайные процессы, рассмотрим систему обслуживания общего вида $G|G|m|n$ (где G (General) – общее распределение и возможно $n=0$ или $n=\infty$). Многие важные проблемы, возникающие при анализе систем массового обслуживания, связаны со следующими процессами:

(1) $Q(t)$ – число требований, находящихся в системе в момент времени t . При $m < \infty$ в $Q(t)$ входят как обслуживаемые, так и ожидающие в очереди требования. Если $m = \infty$, то $Q(t)$ – число занятых обслуживающих приборов.

(2) $q(t) = \max(0, Q(t) - m)$ – число требований, ожидающих в очереди в момент времени t .

(3) $W(t)$ – виртуальное (возможное) время ожидания требования, пришедшего в момент времени t .

Кроме этих основных процессов, важными являются также последовательности случайных величин, характеризующие процесс обслуживания:

(1) W_k – случайное время ожидания (в очереди) k -го требования, $k=1, 2, \dots$

(2) $T_k = W_k + v_k$, $k=1, 2, \dots$ – случайное время пребывания в системе (очередь плюс обслуживание).

Определенные выше случайные процессы с дискретным и непрерывным временем описывают возможные состояния системы массового обслуживания. Знание их полного вероятностного описания желательно, но не обязательно и к тому же связано с решением сложных математических вопросов. На практике часто ограничиваются вычислением функционалов (числовых характеристик) от этих процессов и моментов (средних, дисперсий и т.д.) случайных величин – показателей эффективности систем массового обслуживания. Например, для систем с ожиданием такими показателями являются: \bar{Q} , \bar{q} – среднее число требований и средняя длина очереди в системе, соответственно; \bar{W} , \bar{T} – среднее время ожидания и среднее время пребывания в системе; \bar{v} – среднее время обслуживания.

Эти средние характеристики систем массового обслуживания, вообще говоря, зависят от времени t . В практических задачах интересуются значениями этих характеристик при $t \rightarrow \infty$, что соответствует стационарному режиму работы системы. В таком режиме начальное состояние системы уже не оказывает никакого влияния на поведение системы, и ее показатели не зависят от времени t . Определение характеристик систем массового обслуживания для стационарного режима упрощает исследование и вместе с тем позволяет выбрать способы разумной организации систем.

Кроме перечисленных выше характеристик систем, в стационарном режиме большую роль играют следующие показатели: p_0 – вероятность застать систему свободной; p_k – вероятность того, что в системе находится k требований; p_q – вероятность того, что система занята (имеется очередь); $\rho = \lambda/\mu$ – загрузка системы; ρ/m – загрузка СМО, приходящаяся на один прибор.

Для СМО с отказами важными характеристиками являются: P_k – вероятность того, что в системе находится k требований ($k \leq m$), т.е. k приборов заняты обслуживанием; P_m – вероятность отказа; \bar{Q} – среднее число приборов, занятых обслуживанием; \bar{Q}/m – коэффициент использования прибора; $1 - \bar{Q}/m$ – коэффициент простоя системы; P_0 – вероятность простоя системы; $(1 - p_m)$ – относительная пропускная способность; $\lambda(1 - p_m)$ – абсолютная пропускная способность; $\rho = \lambda/\mu$ – загрузка системы; ρ/m – загрузка СМО, приходящаяся на один прибор.

Вычисления таких характеристик бывают очень сложными и громоздкими [3]. Однако их знание – необходимая часть оценки работы системы.

Марковские процессы

В работе проводится моделирование систем массового обслуживания на основе вероятностных свойств марковского процесса, который описывается начальным распределением вероятностей и матрицей переходных вероятностей.

Вероятностные свойства однородного марковского процесса описываются начальным распределением вероятностей $p_i, i=0, \pm 1, \dots$ и матрицей переходных вероятностей

$$P(t) = \|p_{ij}(t_0)\| = \begin{pmatrix} \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & p_{11}(t) & p_{12}(t) & \cdots & p_{1j}(t) & \cdots \\ \cdots & p_{12}(t) & p_{22}(t) & \cdots & p_{2j}(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & p_{i1}(t) & p_{i1}(t) & \cdots & p_{ij}(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Матрица $P(t)$ называется стохастической, а ее элементы удовлетворяют условиям

$$p_{ij}(t) \geq 0, \quad p_{ij}(0) = \begin{cases} 1, & i = j, \\ 0, & i \neq j, \end{cases} \quad (1)$$

$$\sum_j p_{ij}(t) = 1, \quad i = 0, \pm 1, \dots \quad (2)$$

Уравнение Колмогорова-Чепмена

Переходные вероятности $p_{ij}(t)$ удовлетворяют уравнению Колмогорова-Чепмена

$$p_{ij}(t+s) = \sum_k p_{ik}(s)p_{kj}(t).$$

Это уравнение отражает тот факт, что марковская система, переходя из состояния i в состояние j за время $t+s$, сначала за время s из состояния i переходит в некоторое промежуточное состояние k , а затем за время t из состояния k переходит в состояние j , причем вероятность второго перехода не зависит от того, каким образом было достигнуто состояние k .

Дифференциальные уравнения Колмогорова

Определим матрицу $\Lambda = \|\lambda_{ij}\|$:

$$\Lambda = \lim_{s \rightarrow 0} \frac{P(s) - I}{s},$$

где $p(s)$ – матрица переходных вероятностей, I – единичная матрица с элементами

$$\lambda_{ij} = \lim_{s \rightarrow 0} \frac{p_{ij}(s)}{s}, \quad \text{если } i=j,$$

$$\lambda_{ii} = \lim_{s \rightarrow 0} \frac{p_{ii}(s) - 1}{s}, \quad \text{если } i \neq j.$$

Величины λ_{ij} называются инфинитезимальными коэффициентами, а Λ – инфинитезимальной матрицей. По другой терминологии λ_{ij} – интенсивность (плотность) перехода цепи Маркова из состояния i в состояние j ; $\lambda_i = -\lambda_{ii}$ – интенсивность (плотность) выхода из состояния i . По определению $\lambda_{ij} \geq 0$ для $i \neq j$ и $\lambda_{ii} < 0$.

Очевидно, что $\sum \lambda_{ij} = 0, i=0, +1, \pm 2, \dots$

Уравнение Колмогорова-Чепмена в матричной форме записывается так:

$$P(t+s)=P(t)P(s)=P(s)P(t).$$

В соответствии с этим, при $s>0$

$$\frac{P(t+s)-P(t)}{s} = P(t) \frac{P(s)-I}{s}. \quad (3)$$

Отсюда при $s \rightarrow 0$ получаем

$$P'(t) = P(t)\Lambda \quad (4)$$

– прямое дифференциальное уравнение Колмогорова в матричной форме. Аналогичным образом из (3) получается и обратное дифференциальное уравнение Колмогорова в матричной форме

$$P'(t) = \Lambda P(t), \quad (5)$$

при этом $P(0)=I$. Каждое из уравнений имеет единственное решение

$$P(t) = e^{\Lambda t},$$

где через $e^{\Lambda t}$ обозначена матричная функция

$$e^{\Lambda t} = I + \sum_{k=1}^{\Gamma} t^k \Lambda^k / k!$$

В координатной форме уравнения (4) и (5) записываются в виде соответствующих систем дифференциальных уравнений Колмогорова

$$p'_{ij}(t) = \sum_k p_{ik}(t) \lambda_{kj}$$

и

$$p'_{ij}(t) = \sum_k \lambda_{ik} p_{kj}(t).$$

Аналогичными методами выводится система дифференциальных уравнений для вероятностей $P_i(t)$:

$$p'_j(t) = \sum_k p_k(t) \lambda_{kj}$$

с начальным распределением $p_i(0) = P\{X(0)=i\}$, $i=0, \pm 1, \pm 2, \dots$

Таким образом, мы получаем дифференциальную систему уравнений Колмогорова для счета вероятностей состояний [4, 5]. Расчеты «вручную» таких систем достаточно сложны, но программы, смоделированные на их основе, значительно упрощают работу с системами массового обслуживания.

Моделирование систем массового обслуживания

Одна из задач работы – компьютерная реализация простейших моделей систем массового обслуживания. Методика, разработанная в программах, может быть использована при анализе и оценке эффективности систем массового обслуживания, обладающих свойством марковости. Значительно упрощается расчет вероятностных процессов, задаваемых громоздкими формулами. Графическое отображение состояния систем и их вероятностных характеристик позволяет удобно изучать необходимые вопросы.

В ходе работы проводилось компьютерное моделирование простейших моделей систем массового обслуживания в среде MATLAB.

Среда работы MATLAB

MATLAB (МАТричная ЛАБОратория) – современная математическая среда для профессиональной работы в широком диапазоне научных, инженерных, финансово-экономических и иных приложений. Языком MATLAB являются простые команды, вызывающие готовые вычислительные алгоритмы. Вместе с тем язык MATLAB максимально приближен к обычной форме записи математических выражений, что существенно облегчает его освоение и последующую работу. MATLAB имеет мощные

средства для процедурного, объектно-ориентированного программирования, отладки программ и разработки пользовательского интерфейса. Встроенная библиотека объединяет более 500 математических функций с эффективными средствами двух- и трехмерной визуализации. В дополнение к этому пользователям предоставлена возможность свободно дополнять эту библиотеку своими функциями – специализированными программами, написанными на язык MATLAB, C или Fortran посредством API (Application Programm Interface). Коды MATLAB тщательно оптимизированы, что обуславливает его высокие вычислительные характеристики, особенно при работе с массивами данных.

MATLAB – это интерактивная система, в которой основным элементом данных является массив, это позволяет решать различные задачи, связанные с техническими вычислениями, особенно те, в которых используются матрицы и вектора, в несколько раз быстрее, чем при написании программ с использованием «скалярных» языков программирования, таких как C или Fortran.

Система MATLAB – это набор инструментов и приспособлений, с которыми работает пользователь или программист MATLAB. Она включает в себя средства для управления переменными в рабочем пространстве MATLAB, вводом и выводом данных, а также создания, контроля и отладки М-файлов и приложений MATLAB [6].

Исходя из преимуществ математической среды MATLAB, был выбран именно этот программный продукт, обеспечивающий оптимальную базу вычислительных алгоритмов для моделирования систем массового обслуживания. Блок управляемой графики MATLAB незаменим для анализа полученных результатов путем построения изображений и анимации. Пользовательский интерфейс системы прост в освоении и удобен для реализации проекта.

Результаты

Преимущества, полученные таким образом, очевидны: во-первых, как уже упоминалось, в программах производится расчет громоздких формул. Ниже приведены формула для счета среднего числа приборов, занятых на данный момент времени в процессе обслуживания,

$$a_{m,n} = \sum_{k=1}^m k p_k + m \sum_{k=m+1}^{m+n} p_k = (\rho \sum_{k=0}^{m-1} \rho^k / k! + \rho^m / (m-1)!) \sum_{j=1}^n (\rho/m)^j (\rho \sum_{k=0}^m \rho^k / k! + \rho^m / m!) \sum_{j=1}^n (\rho/m)^j)^{-1}$$

и формула, дающая возможность вычислять характеристики, связанные со временем ожидания.

$$P\{W < t\} = \sum_{k=0}^{m-1} p_m + m \mu \int_0^t e^{-\mu u} \left(\sum_{k=m}^{\infty} p_k \frac{(m\mu)^{k-m} e^{-m\mu}}{(k-m)!} \right) du = p_0 \sum_{k=0}^{m-1} \rho^k / k! + p_0 m \mu \int_0^t e^{-m\mu u} \left(\sum_{k=m}^{\infty} \frac{\rho^k}{m! m^{k-m}} \frac{(m\mu)^{k-m}}{(k-m)!} \right) du.$$

Аналитические модели требуют индивидуальных расчетов для различных типов систем и даже одинаковых систем с различными параметрами. Соответствующие выкладки достаточно громоздки и объемны. При компьютерном моделировании этого удастся избежать.

Исходными данными для программ служат параметры задаваемой системы. На выходе мы получаем графики процессов, проходящих в системе ((а) график общего состояния системы, (б) процесс образования очереди, (в) поток поступающих заявок, (г) поток обслуживания на рис. 2) и теоретические оценки такой системы.

В систему в случайные моменты времени поступают заявки. В дальнейшем поведение заявки зависит от наличия свободных приборов. В случае, когда в системе есть хотя бы 1 свободный прибор, заявка поступает на обслуживание. Если же в системе все приборы заняты, то заявка становится в очередь либо покидает систему (опять же,

в зависимости от того, с какой системой мы имеем дело). Для наглядного примера на рис. 2, 3 приведены графики двух возможных вариантов работы программы:

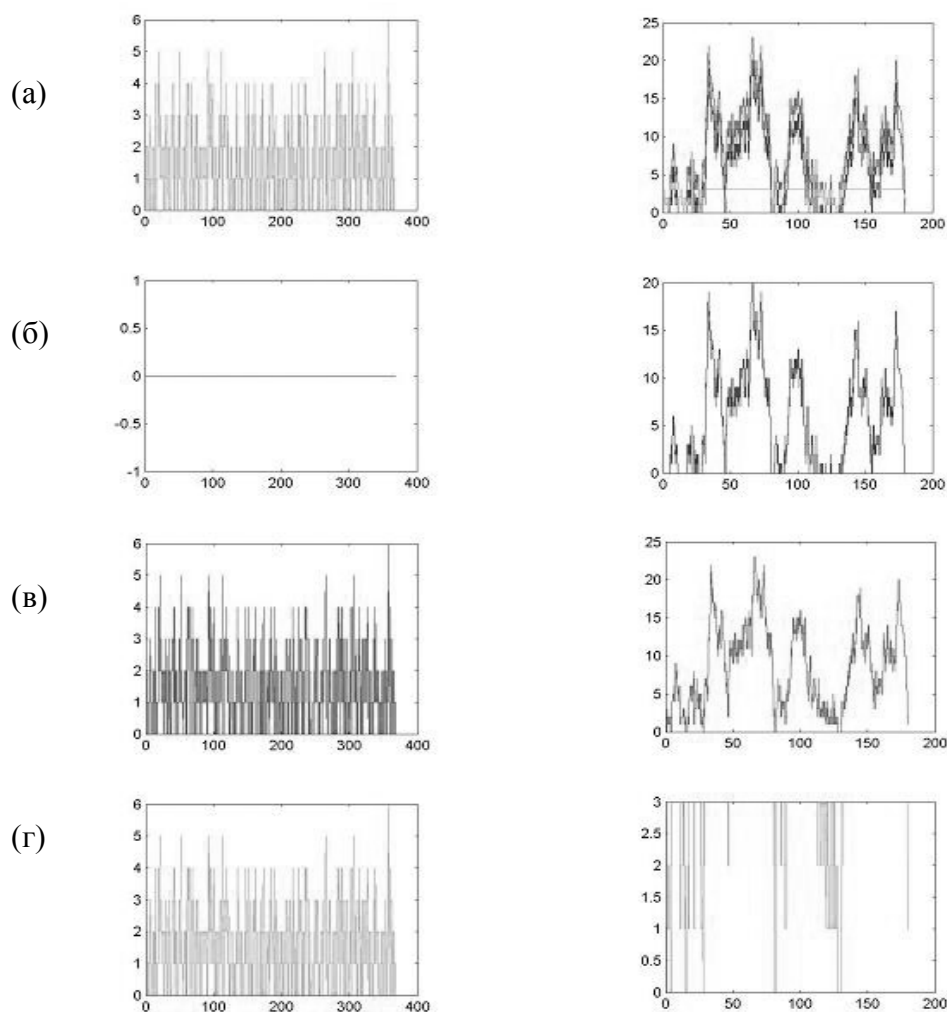


Рис. 2. Оптимальная работа системы Рис. 3. Работа системы с перегрузкой

На приведенных графиках вертикальная ось координат показывает количество приборов, работающих в системе, горизонтальная – количество итераций.

На рис. 2 работа системы положительна, т.е. в ней не образуется очереди, соответственно, не происходит потеря заявок, и система не перегружается.

Об оптимальности работы системы рис. 3 говорить не приходится. Здесь происходит перегрузка (в системе образуется очередь).

Второе преимущество, реализованное в работе, – быстрое и легкое оценивание поведения работающей системы в момент времени t . В заданный момент времени мы можем наглядно оценить вероятностные процессы, проходящие в системе. Другая часть программы позволяет нам увидеть, как с течением времени изменяется вероятность того, что в системе имеется k заявлений.

Рассмотрим систему массового обслуживания с ожиданием. На рис. 4.1 все вероятности, по мере увеличения количества итераций, стабилизируются – стремятся к фиксированному числу. Рис. 4.2 показывает вероятность того, что заявка, пришедшая в систему, окажется в очереди. Как видим, она увеличивается с течением времени. На рис. 4.1: вертикальная ось координат показывает вероятность, горизонтальная – количество итераций. На рис. 4.2 вертикальная ось координат показывает вероятность, горизонтальная – течение времени. На рис. 5 представлена замкнутая система.

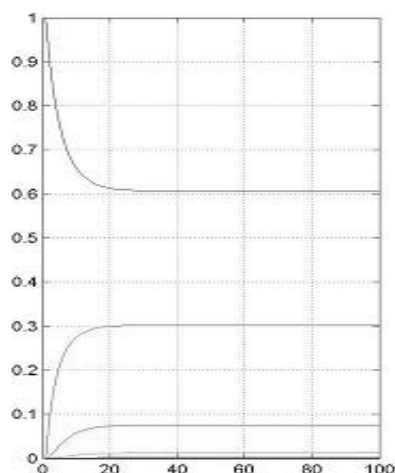


Рис. 4.1. Система с ожиданием

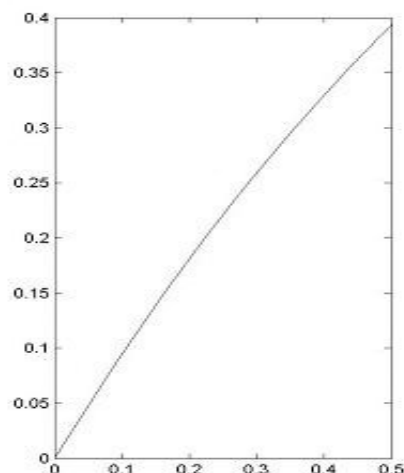


Рис. 4.2. Система с ожиданием

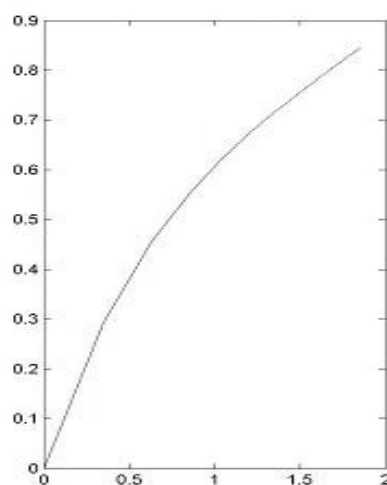
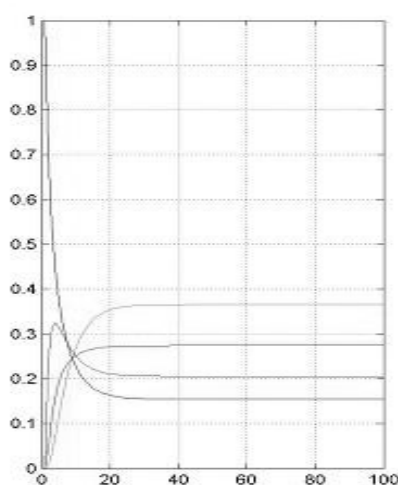


Рис. 5. Замкнутая система

Сравнивая и оценивая поведение разных систем, мы можем прийти к выводу о наибольшей выгоде, получаемой при работе различных систем в заданных условиях.

Заключение

Итак, мы промоделировали системы массового обслуживания на основе вероятностных свойств марковских процессов. При компьютерной реализации систем массового обслуживания использовался математический пакет MATLAB 6.5. Практическая ценность такой работы состоит в возможности ее использования для выяснения оптимальности реальных систем массового обслуживания, обладающих свойством марковости. Преимуществом является быстрое и наглядное получение данных, простые расчеты аналогичных математических формул. В процессе работы получены наглядные результаты функционирования некоторых систем массового обслуживания.

Литература

1. Бочаров П.П., Печенкин А.В. Теория вероятностей и математическая статистика. М.: Гардарики, 1998.
2. Калинина В.Н., Панкин В.Н. Математическая статистика. М.: Высшая школа, 1998.

3. Методические указания для студентов технических специальностей всех форм обучения. СПб, 1998.
4. Кремер Н.Ш. Теория вероятностей и математическая статистика: Учебник для вузов. / 2-е изд., перераб. и доп. М.: ЮНИТИ-ДАНА, 2004.
5. Колемаев В.А., Калинина В.Н. Теория вероятностей и математическая статистика. М.: ИНФРА-М, 1997.
6. Потемкин В.Г. Система MATLAB, М.: ДИАЛОГ-МИФИ, 1997.
7. Так же использованы статьи и материалы среды Интернет.

ИЗУЧЕНИЕ ЭФФЕКТИВНОСТИ ОБНАРУЖЕНИЯ СИГНАЛА ДЛЯ МНОГОКАНАЛЬНОЙ СИСТЕМЫ

Н.В. Дьякова

Научный руководитель – к.ф.-м.н., доцент И.А. Суслина

Целью работы является исследование качества линейных тестов и сравнение их свойств с оптимальным тестом в неасимптотическом случае для задачи проверки гипотез в многоканальной системе. Проведено моделирование и обнаружение сигнала в многоканальной системе. Оценена эффективность асимптотически минимаксных тестов.

Введение

В статье изучается задача обнаружения сигнала в многоканальной системе, каждый из каналов которой может содержать или не содержать сигнал $f(t)$, $t \in (0,1)$, наблюдаемый в гауссовском белом шуме. Задача состоит в обнаружении факта наличия сигнала в наблюдаемых данных.

Рассматривается модель гауссовского белого шума, т.е. наблюдаются n независимых случайных процессов $X_i(t)$, $t \in (0,1)$, каждый из которых соответствует i -му каналу системы и удовлетворяет стохастическому дифференциальному уравнению

$$dX_i(t) = \xi_i f(t) + \varepsilon dW_i(t), t \in (0,1), i = 1, \dots, n.$$

Рассматривается случай сигнала известной формы $f(t) = c\phi(t)$, $\phi \in L_2(0,1)$, $\|\phi\| = 1$; где $L_2(0,1)$ – пространство измеримых функций на $(0,1)$ и суммируемых со степенью 2. Здесь ϕ – известная функция, определяющая форму сигнала, а величина $c = c_n = \|\mathbf{f}\| > 0$ определяет уровень сигнала.

Цель работы – исследование качества линейных тестов и сравнение их свойств с оптимальным тестом в неасимптотическом случае, т.е. в случае небольшого числа каналов, для задачи проверки гипотез в многоканальной системе.

В процессе работы было произведено:

- (а) моделирование статистик $L(X)$;
- (б) построение кривых обнаружения для байесовского теста и линейного теста;
- (в) сравнение тестов на базе кривых обнаружения.

При моделировании многоканальной системы использовался математический пакет MATLAB 6.5.

Полученные результаты можно использовать всюду, где идет обработка многоканальных систем (например, техническая и медицинская диагностика, прием и передача информации, радиотехника и т.д.).

1. Обнаружение сигнала в многоканальной системе

1.1. Постановка задачи

При моделировании многоканальной системы используются результаты статьи [1]. В [1] изучается задача обнаружения сигнала в многоканальной системе, каждый из каналов которой может содержать или не содержать сигнал $f(t)$, $t \in (0,1)$, наблюдаемый в гауссовском белом шуме. Задача состоит в обнаружении факта наличия сигнала в наблюдаемых данных.

Задачи такого типа возникают при разработке и исследовании различных систем технической или медицинской диагностики, радиолокационных систем, систем передачи информации и т.д.

Рассматривается случай, когда число каналов n велико, и поэтому рассматривается асимптотика при $n \rightarrow \infty$.

Точная постановка задачи состоит в следующем. Рассматривается модель гауссовского белого шума, т.е. наблюдаются n независимых случайных процессов $X_i(t), t \in (0,1)$, каждый из которых соответствует i -му каналу системы и удовлетворяет стохастическому дифференциальному уравнению

$$dX_i(t) = \xi_i f(t) + \varepsilon dW_i(t), t \in (0,1), i = 1, \dots, n. \quad (1.1)$$

Здесь $W_i(t)$ – независимые винеровские процессы, $\varepsilon = \varepsilon_n > 0$ – известный уровень шума, $\xi_i \in \{0,1\}$ – параметры, соответствующие наличию сигнала $f \in L_2(0,1)$ в i -ом канале ($\xi_i = 1$) или отсутствию сигнала ($\xi_i = 0$). Рассматривается случай сигнала известной формы $f(t) = c\phi(t)$, $\phi \in L_2(0,1), \|\phi\| = 1$; здесь ϕ – известная функция, определяющая форму сигнала, а величина $c = c_n = \|f\| > 0$ определяет уровень сигнала.

Обозначим $\bar{f} = (\xi_1 f, \dots, \xi_n f) \in L_2^n(0,1)$ набор функций, характеризующих наличие или отсутствие сигнала f в каналах. Величина $K(\bar{\xi}) = \sum_{i=1}^n \xi_i$ есть число каналов, содержащих сигналы; $0 \leq K(\bar{\xi}) \leq n$.

Пусть $F_n = \{\bar{f}\} \subset L_2^n(0,1)$ – заданное множество наборов сигналов в каналах. Задачу обнаружения сигнала можно рассматривать как задачу проверки гипотезы $H_0: \bar{f} = 0$ против альтернативы $H_{1,F_n}: \bar{f} \in F_n$.

Изучается эта задача с точки зрения минимаксного подхода. Именно, пусть $\psi_n = \psi_n(X_1, \dots, X_n)$ – тест, т.е. измеримая функция от наблюдений со значениями в интервале $[0, 1]$. Пусть $\alpha_n(\psi_n) = E_{n,0}(\psi_n)$ есть вероятность ошибок I рода и $\beta_n(\psi_n, \bar{f}) = E_{n,\bar{f}}(1 - \psi_n)$ есть вероятность ошибок II рода теста n (здесь и ниже $E_{n,\bar{f}}$ есть математическое ожидание по мере $P_{n,\bar{f}}$, соответствующей наблюдениям (1.1); мера $P_{n,0}$ соответствует отсутствию сигналов в системе). Обозначим $\gamma_n(\psi_n, F_n)$ сумму вероятности ошибок I рода и максимальной вероятности ошибок II рода $\gamma_n(\psi_n, F_n) = \alpha_n(\psi_n) + \max_{\bar{f} \in F_n} \beta_n(\psi_n, \bar{f})$. Положим $\gamma_n = \gamma_n(F_n) = \inf_{\psi_n} \gamma_n(\psi_n, F_n)$. Величину $\gamma_n(F_n)$ будем называть минимаксной ошибкой обнаружения для альтернативы $H_1 = H_{1,F_n}$. Легко видеть, что $0 \leq \gamma_n \leq 1$.

Будем говорить, что тест n является минимаксным для альтернативы H_{1,F_n} , если $\gamma_n(\psi_n, F_n) = \gamma_n(F_n)$.

Рассматриваемую задачу можно представить в несколько другой форме. Рассмотрим статистику

$$x_i = \varepsilon^{-1} \int_0^1 \phi(t) dX_i(t), i = 1, \dots, n. \quad (1.2)$$

Легко видеть, что это достаточные статистики в рассматриваемой задаче, и они независимы как при гипотезе, так и при альтернативе, соответствующей любому набору $\bar{\xi}_n = (\xi_1, \dots, \xi_n)$. Переход к статистикам x_i вида (1.2) в радиотехнике часто называют фильтрацией сигналов в каналах, согласованной с известной формой сигнала. При ги-

позете H_0 статистики x_i имеют стандартное гауссовское распределение $N(0,1)$. Пусть $b=c/\varepsilon$ есть отношение «сигнал-шум». Тогда при альтернативе, соответствующей набору $\bar{\xi}_n = (\xi_1, \dots, \xi_n)$, статистики x_i имеют гауссовское распределение $N(\nu_i, 1)$, $\nu_i = b\xi_i$. Переходя к статистикам (1.2), получаем вектор наблюдений вида $x = \nu + \eta$, $\nu \in \mathbb{R}^n$, где $\eta = (\eta_1, \dots, \eta_n)$ – стандартный n -мерный нормальный вектор, а $\nu = (b\xi_1, \dots, b\xi_n)$ – n -мерный вектор, соответствующий наличию или отсутствию сигнала уровня c в каналах. Получается эквивалентная задача проверки гипотез о среднем n -мерного нормального распределения: $H_0: \nu = 0$, $H_{1, V_n}: \nu \in V_n$, где $V_n \subset \mathbb{R}^n$ – множество, соответствующее множеству $F_n \subset L_2^n(0, 1)$. Величины $\beta_n(\psi_n, \nu)$, $\gamma_n(\psi_n) = \gamma_n(\psi_n, V_n)$, $\gamma_n = \gamma_n(V_n)$ в этой задаче определяются аналогично.

Пусть число каналов k , $1 \leq k \leq n$, содержащих сигналы, и отношение сигнал-шум $b < 0$ известны. Полагается

$$V_n(b, k) = \{ \nu = (\xi_1 b, \dots, \xi_n b), \xi_i \in \{0, 1\}, \sum_{i=1}^n \xi_i = k \}, \quad (1.3)$$

и рассматривается эквивалентная задача проверки гипотез о среднем n -мерного нормального распределения $H_0: \nu = 0$ против альтернативы $H_{1, V_n}(b, k): \nu \in V_n(b, k)$. Нетрудно указать вид минимаксного теста в этой задаче. Именно, пусть π – равномерное априорное распределение на дискретном множестве $V_n(b, k)$:

$\pi(\nu) = M^{-1}$, $\forall \nu \in V_n(b, k)$; $M = (V_n(b, k)) = C_k^n$. Рассматривается байесовская задача проверки гипотезы $H_0: P = P_{n,0}$ против альтернативы $H_1: P = P_\pi$ о распределении P наблюдений x ; здесь $P_\pi = \int_{\mathbb{R}^n} P_{n,\nu} \pi(d\nu)$ – смесь по априорному распределению. Пусть ψ_n –

байесовский тест, минимизирующий $\gamma(\psi, P_\pi) = E_{n,0} \psi + E_{P_\pi} (1 - \psi)$. По лемме Неймана-Пирсона он имеет вид $\psi_\pi = 1_{\{L(x) > 1\}}$, где $L(x)$ есть байесовское отношение правдоподобия:

$$L(x) = \frac{dP_\pi}{dP_{n,0}}(x) = \int \exp(-|\nu|^2 / 2 + (x, \nu)) \pi(d\nu) = (C_k^n)^{-1} e^{-kb^2/2} \sum_{A_{n,k}} \exp(b \sum_{i \in A_{n,k}} x_i),$$

где суммирование ведется по всевозможным подмножествам $A_{n,k} \subset \{1, \dots, n\}$, содержащим k элементов. Из соображений инвариантности следует, что есть минимаксный тест для альтернативы $V_n(b, k)$, т.е. $\psi(V_n(b, k)) = \gamma(\psi_\pi, V_n(b, k))$. Это означает, что есть наиболее благоприятное априорное распределение. Однако структура статистики $L(x)$ достаточно сложна при $k < n$, и трудно исследовать вероятности ошибок минимаксного теста. Кроме того, статистика $L(x)$ существенно зависит от параметров k, b .

В этой связи мы изучаем асимптотический вариант задачи при $n \rightarrow \infty$. Будем говорить, что последовательность тестов ψ_n является асимптотически минимаксной для альтернативы H_{1, V_n} , если $\gamma_n(\psi_n, V_n) = \gamma_n(V_n) + o(1)$, и состоятельной для альтернативы H_{1, V_n} , если $\gamma_n(\psi_n, V_n) \rightarrow 0$. Здесь и далее пределы рассматриваются при $n \rightarrow \infty$.

1.2. Основные результаты

Пусть $k = k_n$, $b = b_n$ – заданные последовательности, $b_n > 0$, $1 \leq k_n \leq n$.

Нас интересуют следующие вопросы:

(1) какой вид имеют условия различимости в задаче, т.е. насколько большими должны быть отношения сигнал-шум b_n в зависимости от общего числа каналов n и

числа каналов k_n , содержащих сигналы, чтобы минимаксные ошибки обнаружения γ_n стремились к 0.

(2) как построить возможно более простые асимптотически минимаксные и состоятельные тесты ψ_n при выполнении соответствующих условий различимости, и как их структура зависит от параметров b_n, k_n .

Введем величину $\delta_n = (\log k_n) / (\log n) \in [0, 1]$, т.е. $k_n = n^{\delta_n}$. Положим $u_n^2 = u_n(k_n, b_n) = n^{-1} k_n^2 (e^{b_n^2} - 1)$, $\tilde{u}_n = \tilde{u}_n(k_n, b_n) = n^{-1/2} k_n b_n$, $T_n = \sqrt{2 \log n}$. Будем обозначать $\Phi(t)$ функцию распределения стандартного нормального закона.

Теорема 1.1. Пусть $V_n = V_n(b, k)$ – множества вида (1.3) и $\liminf \delta_n > 1/4$.

Тогда справедливы следующие утверждения.

(1) Справедливо неравенство: $\gamma_n \geq 2\Phi(-u_n/2) + o(1)$.

(2) Рассмотрим тесты вида $\hat{\psi}_{n, b_n} = \hat{\psi}_{n, b_n, H_n} = \max(\hat{\psi}_{n, b_n, H_n}, \hat{\psi}_n^{thr})$. Здесь $\hat{\psi}_{n, b_n, H_n} = 1_{t_{n, b_n} > H_n}$ есть тесты, основанные на статистиках $t_{n, b_n} = A_n \sum_{i=1}^n v(x_i, b_n)$, $A_n = (n(e^{b_n^2} - 1))^{-1/2}$, $v(t, b) = (e^{-b^2/2+tb} - 1)$ и пороге $H_n = u_n/2$, а $\hat{\psi}_n^{thr}$ есть тесты порогового типа

$$\hat{\psi}_n^{thr} = 1_{\chi_n}, \chi_n = \{x \in \mathbb{R}^n : \max_{1 \leq i \leq n} x_i > T_n\} \quad (1.4)$$

Тогда $\gamma_n \leq \gamma_n(\hat{\psi}_{n, b_n, H_n}) \leq 2\Phi(-u_n/2) + o(1)$.

(3) Рассмотрим тесты, основанные на линейных статистиках $\hat{\psi}_n^{lin} = 1_{t_n^{lin} > H_n}$, $t_n^{lin} = n^{-1/2} \sum_{i=1}^n x_i$, и пусть $H_n = \tilde{u}_n/2$. Тогда $\gamma_n \leq \gamma_n(\hat{\psi}_n^{lin}) = 2\Phi(-\tilde{u}_n/2)$. Кроме того, $\gamma_n(\hat{\psi}_n^{lin}) \rightarrow 0$ при $\tilde{u}_n \rightarrow \infty$ и таких $H_n \rightarrow \infty$, что $\tilde{u}_n - H_n \rightarrow \infty$.

Теорема 1.1 описывает точную асимптотику гауссовского типа минимаксных вероятностей ошибок в рассматриваемой задаче. Утверждения п.п. (1), (2) означают, что если $\liminf \delta_n > 1/4$, то тесты $\hat{\psi}_{n, b_n, H_n}$ являются асимптотически минимаксными.

Имеется соотношение $\tilde{u}_n \leq u_n$, и при условии $nk_n^{-2} = O(1)$ соотношение $\tilde{u}_n \rightarrow \infty$ имеет место тогда и только тогда, когда $u_n \rightarrow \infty$; кроме того, при условиях $nk_n^{-2} = o(1)$ и $u_n = O(1)$ мы имеем $b_n = o(1)$ и, следовательно, $\tilde{u}_n \sim u_n$ в этом случае. Поэтому утверждение п. (3) означает, что тесты $\hat{\psi}_n^{lin}$ также являются асимптотически минимаксными в случае $nk_n^{-2} = o(1)$, а при $nk_n^{-2} = O(1)$ они обеспечивают состоятельную различимость, если состоятельные тесты существуют, т.е. при $u_n \rightarrow \infty$. Кроме того, нетрудно видеть, что если $k_n = n$, то тесты $\hat{\psi}_n^{lin}$ являются минимаксными (не асимптотически). Статистики тестов $\hat{\psi}_n^{lin}$, не зависят от параметров b_n и k_n .

Используя теорему 1.1 и эти замечания, можно получить простые условия различимости и неразличимости в задаче (т.е. условия того, что $\gamma_n \rightarrow 0$ или $\gamma_n \rightarrow 1$).

Следствие.

(1) Пусть $nk_n^{-2} = o(1)$. Тогда $\gamma_n \rightarrow 1$ при $\tilde{u}_n \rightarrow 0$ и $\gamma_n(\psi_n^{lin}) \rightarrow 0$ при $\tilde{u}_n \rightarrow \infty$.

(2) Пусть $k_n = n_n^\delta$, $nk_n^{-2} \rightarrow \infty$ и $\delta_n \rightarrow \delta \in (1/4, 1/2]$. Положим

$$b_n^*(k_n) = \sqrt{\log(nk_n^{-2})} = \sqrt{(1-2\delta_n)\log n}. \quad (1.5)$$

Тогда, если $\limsup b_n / b_n^*(k_n) < 1$, то $\gamma_n \rightarrow 1$. Если же $\liminf b_n / b_n^*(k_n) > 1$, то

$$\gamma_n(\hat{\psi}_{n, b_n, H_n}) \rightarrow 0.$$

Если $\delta_n \rightarrow \delta \in (0, 1/4)$, то точная асимптотика минимаксных вероятностей ошибок в задаче описывается в терминах специальных безгранично делимых распределений без гауссовской компоненты. Однако можно получить другую точную границу между условиями различимости и неразличимости при $\liminf \delta_n \leq 1/4$.

Положим при $\delta_n \in [0, 1/4]$

$$b_n^*(k_n) = (1 - \delta_n^{1/2}) \sqrt{2 \log n}. \quad (1.6)$$

Теорема 1.2. Пусть $\delta_n \rightarrow \delta \in [0, 1/4]$. Тогда справедливы утверждения.

(1) Если $\limsup b_n / b_n^*(k_n) < 1$, то $\gamma_n \rightarrow 1$.

(2) Если $\liminf b_n / b_n^*(k_n) > 1$, то $\gamma_n^{thr}(\psi_n^{thr}) \rightarrow 0$, где тесты ψ_n^{thr} определены (1.4).

Более того, это справедливо при $\delta_n \rightarrow \delta$ для любого $\delta_n \in [0, 1/2)$.

Заметим, что тесты ψ_n^{thr} , являющиеся состоятельными при условиях $\liminf b_n / b_n^*(k_n) > 1$ и $\liminf b_n / b_n^*(k_n) < 1/2$, где определены равенствами (1.5), (1.6), не зависят от параметров k_n и b_n .

2. Моделирование многоканальной системы в среде Matlab

2.1 Исходные данные

Для выбора исходных данных моделирования системы были учтены особенности эксперимента:

- необходимость построения оптимальных тестов для анализа обнаружения сигнала;
- определение более точной оценки близости (расходимости) кривых обнаружения для различных тестов;
- вычислительные мощности используемой для осуществления эксперимента техники.

После анализа всех нюансов эксперимента были выбраны следующие исходные данные: количество каналов системы: $n=50$; количество каналов с сигналами: $k=49; 48; 47; 46$; отношение сигнал-шум: $b = c\sqrt{n}/k$, где $c = 1$.

2.2 Создание статистик L(X)

Сигнал многоканальной системы имеет вид: $x = \nu + \eta$, $\nu \in \mathbb{R}^n$.

Для проверки гипотезы с альтернативой вида: $H_0: \nu = 0$, $H_{1, \nu_n}: \nu \in V_n$ необходимо построить на базе леммы Неймана-Пирсона байесовский критерий при наиболее неблагоприятном априорном распределении π , он базируется на статистике $L(x)$, где $\bar{x} = (x_1, \dots, x_n)$, $x_i \sim N(0, 1)$:

$$L(x) = \frac{dP_{\pi}(x)}{dP_{n,0}} = \int \exp(-|v|^2/2 + (x, v)) \pi(dv) = (C_k^n)^{-1} e^{-kb^2/2} \sum_{A_{n,k}} \exp(b \sum_{i \in A_{n,k}} x_i).$$

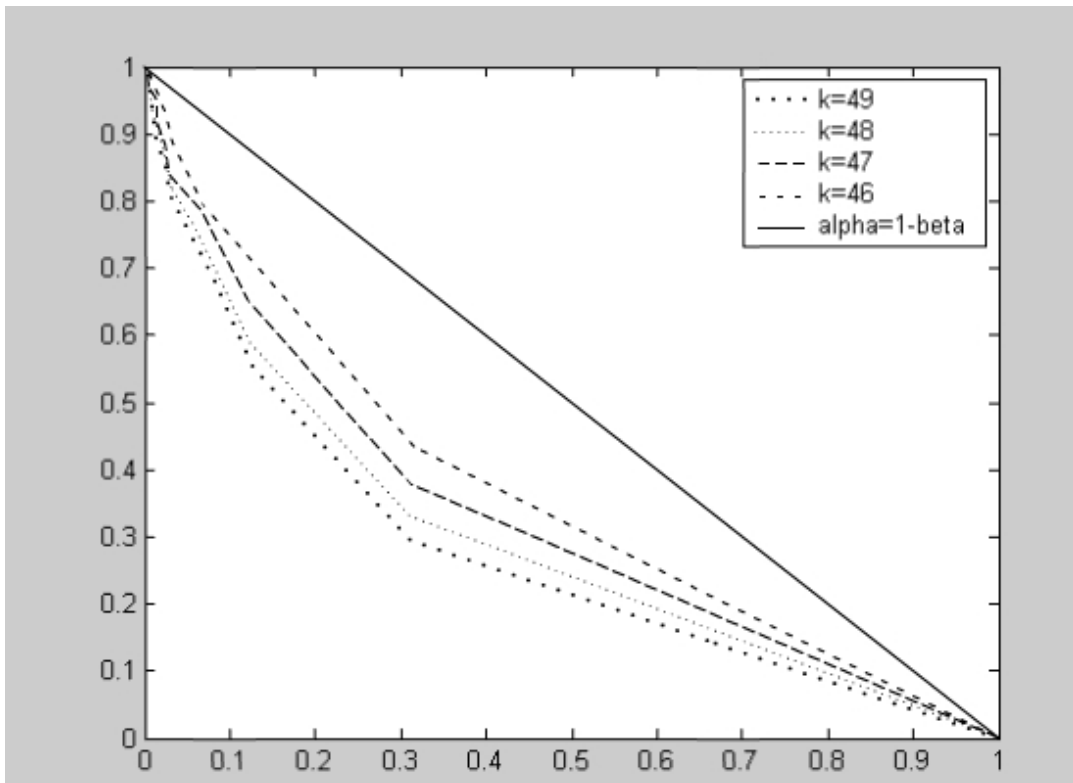


Рис. 2.1. Кривые обнаружения для байесовского теста ($n=50, c=1$) в динамике по количеству каналов с сигналами

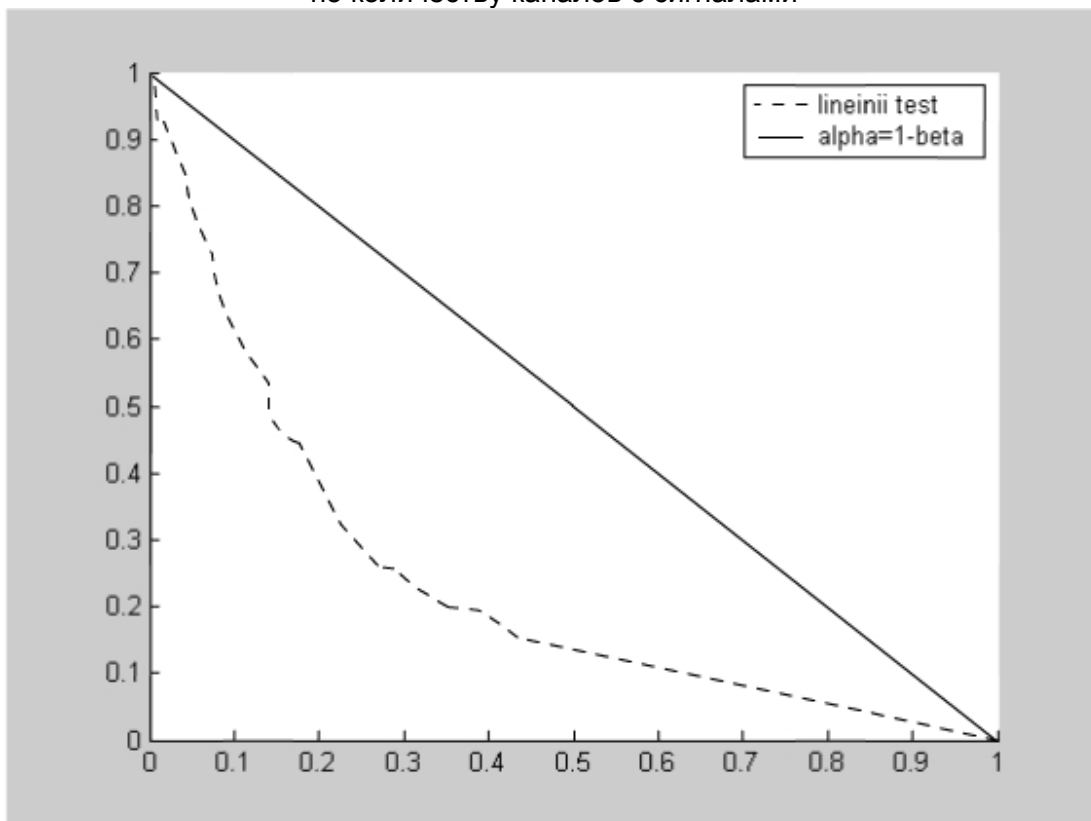


Рис. 2.2. Кривая обнаружения для линейного теста ($c=1$)

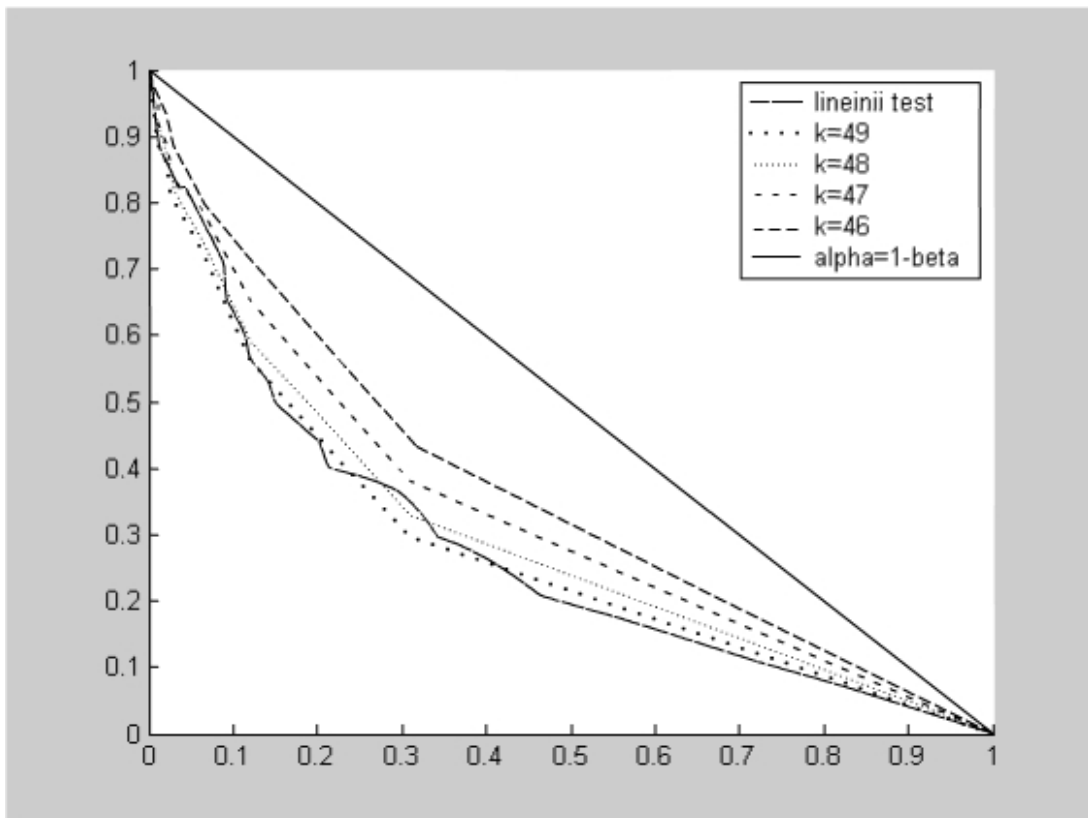


Рис. 2.3. Кривые обнаружения для байесовского теста ($n=50, c=1$) в динамике по количеству каналов с сигналами и кривая обнаружения для линейного теста

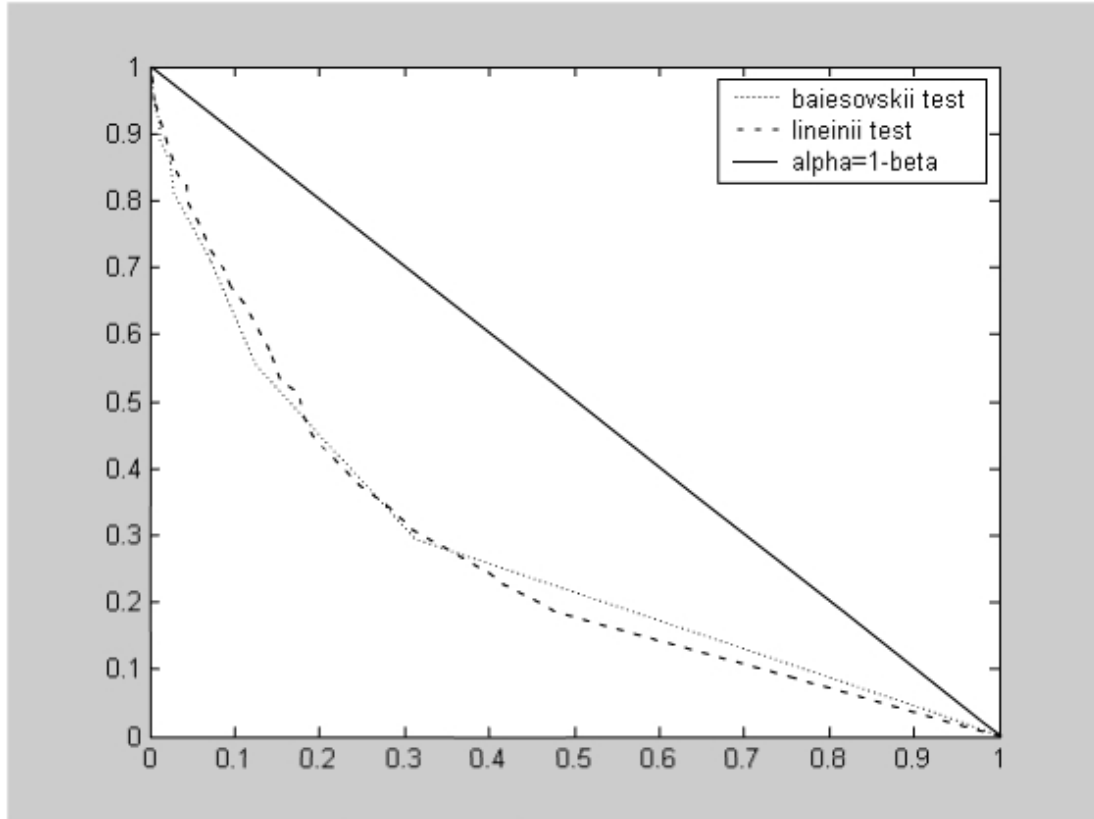


Рис. 2.4. Кривые обнаружения для байесовского и линейного тестов ($n=50, k=49$)

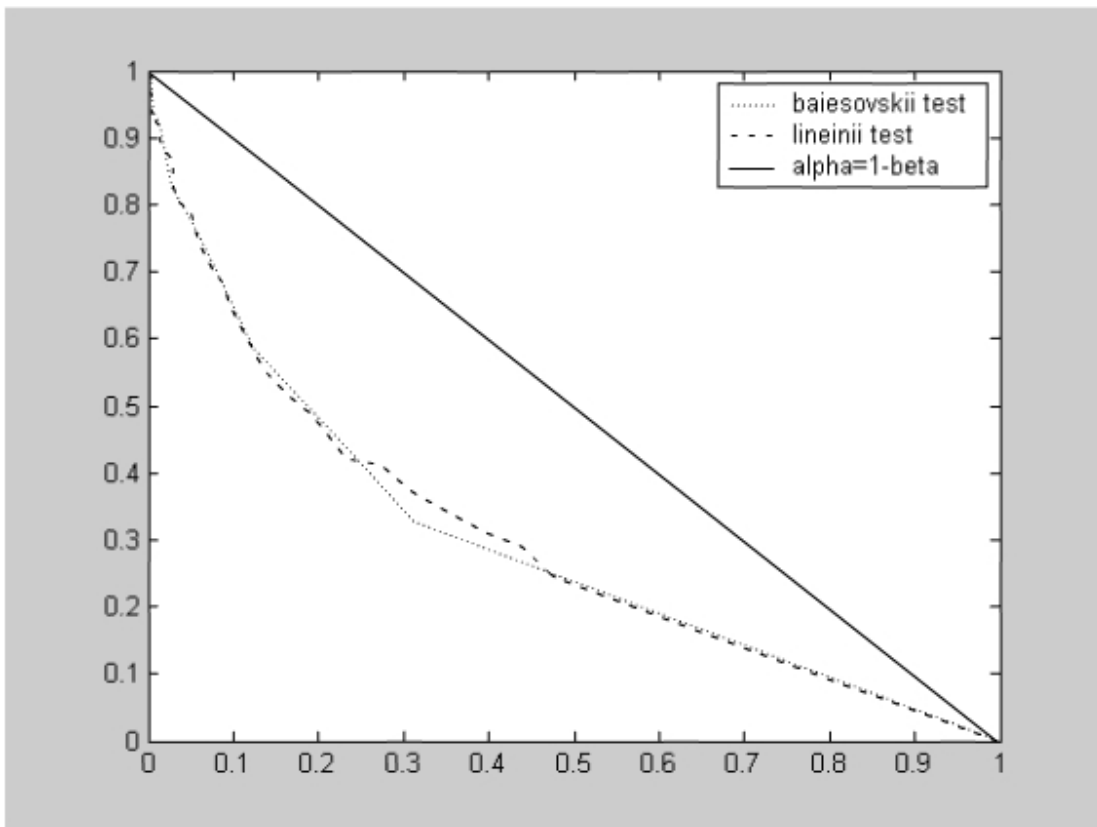


Рис. 2.5. Кривые обнаружения для байесовского и линейного тестов ($n=50, k=48$).

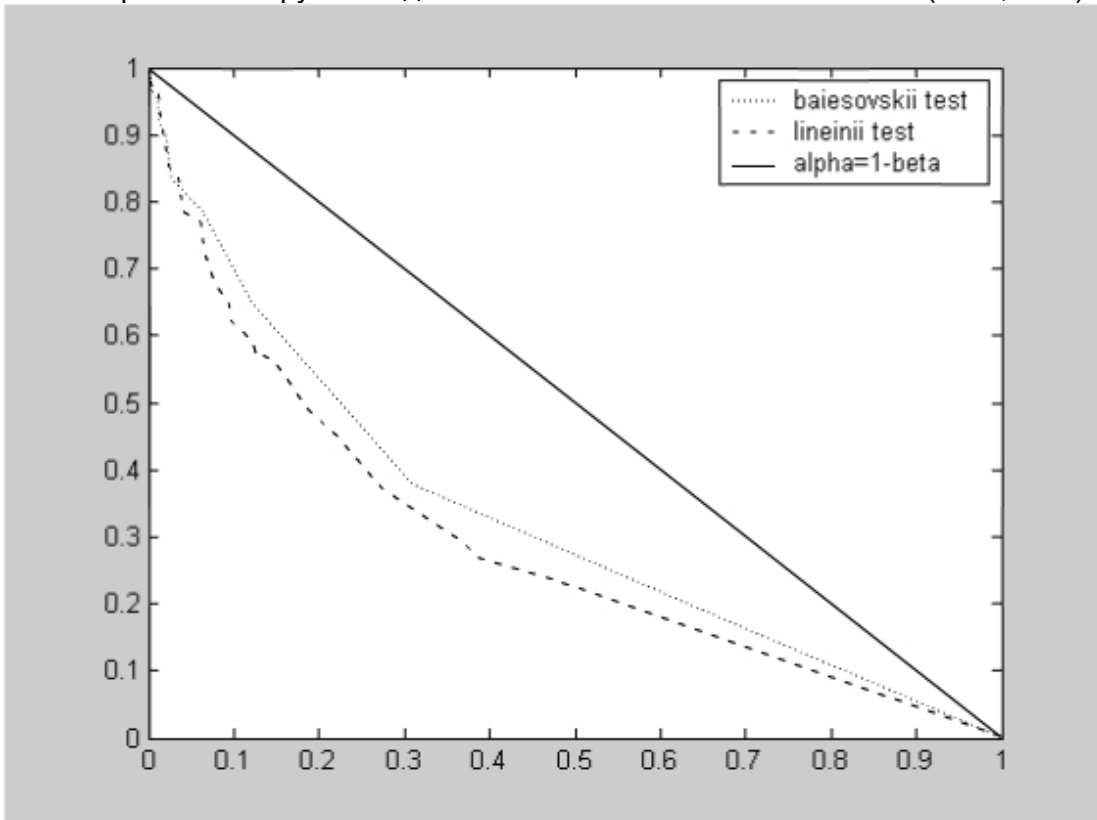


Рис. 2.6. Кривые обнаружения для байесовского и линейного тестов ($n=50, k=47$).

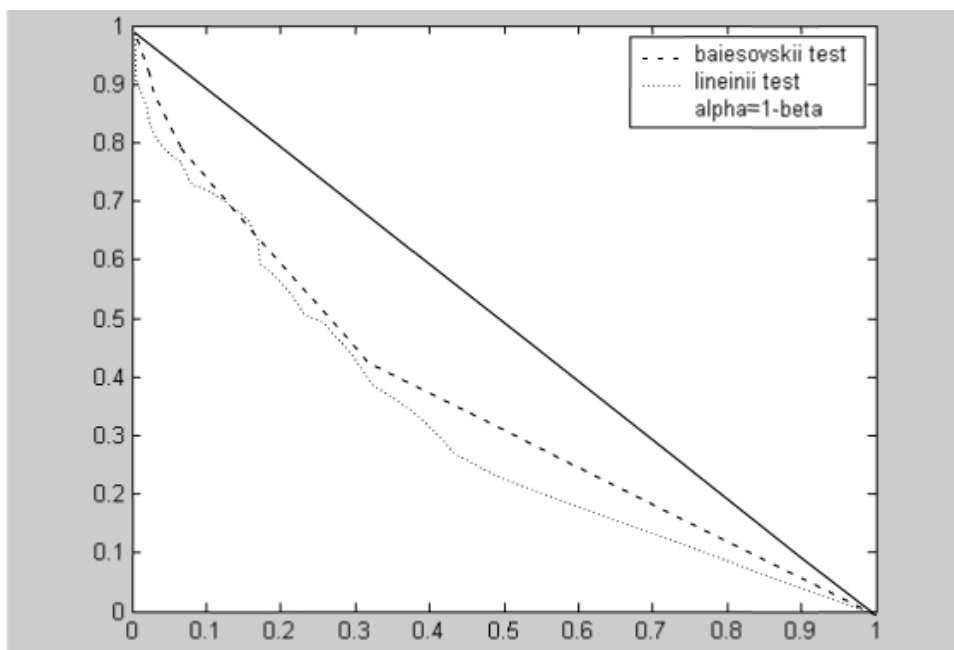


Рис. 2.7. Кривые обнаружения для байесовского и линейного тестов ($n=50, k=46$)

3. Выводы

В ходе эксперимента, основываясь на лемме Неймана-Пирсона, был построен байесовский критерий при наиболее неблагоприятном априорном распределении π на базе статистик $L(x)$ и $L_1(x)$. Этот критерий из-за сложности моделирования статистики неудобен в применении. Было проведено сравнение этого критерия с асимптотически минимаксным (линейным) с помощью кривых обнаружения. Анализ графиков показал состоятельность линейного теста для 50 каналов. Кривая обнаружения для линейного теста имеет сравнимую с кривыми обнаружения для байесовского теста различимость при уровне сигнала $s=1$ и количестве каналов с сигналами $k=49, 48, 47, 46$. Благодаря полученным результатам облегчается обработка многоканальных систем и при небольшом числе каналов.

Заключение

В ходе работы были проведены эксперименты по моделированию многоканальной системы со сравнительно малым числом каналов ($n=50$) и выполнено сравнение точных байесовских тестов и асимптотически минимаксных (линейных) на базе кривых обнаружения. В результате анализа экспериментов сделаны выводы о состоятельности линейных тестов и для малого числа каналов. Эти тесты удобны в применении, в отличие от байесовских тестов, которые строятся на базе леммы Неймана-Пирсона.

Литература

1. Ингстер И.Ю., Суслина И.А. Об обнаружении сигнала известной формы в многоканальной системе // Записки научных семинаров ПОМРАН, 2002. Т. 294. С.88–112.
2. Бодрова Н.А., Родина Т.В., Суслина И.А. Элементы теории вероятностей и математической статистики (Под редакцией В.П. Смирнова). СПб: СПбГУ ИТМО, 2001.
3. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики // СПб: Лань, 1998.
4. Ингстер Ю.И. Асимптотические методы в статистике // СПб: ПГУПС, 2000.
5. Тюрин Ю.Н., Макаров А.А. Статистический анализ данных на компьютере // М.: Имфра-М, 1998.

МОДЕЛИРОВАНИЕ ТЕСТОВ ОБНАРУЖЕНИЯ СИГНАЛА В МНОГОКАНАЛЬНОЙ СИСТЕМЕ И СРАВНЕНИЕ ИХ СВОЙСТВ

К.В. Дехтерева

Научный руководитель – к.ф.-м.н, доцент И.А. Суслина

Цель работы – изучение задачи обнаружения сигнала в многоканальной системе, каждый из каналов которой может содержать или не содержать сигнал определенного вида. Проведено моделирование многоканальной системы, оценена эффективность оптимальных по порядку различимости тестов. По результатам экспериментов проведено сравнение эффективности алгоритмов обнаружения на базе полученных кривых обнаружения.

Введение

В работе изучается задача обнаружения сигнала в многоканальной системе, каждый из каналов которой может содержать или не содержать сигнал определенного вида, наблюдаемый в гауссовском белом шуме. Задача состоит в обнаружении факта наличия сигнала в наблюдаемых данных. Задачи такого типа возникают при разработке и исследовании различных систем технической или медицинской диагностики, радиолокационных систем, систем передачи информации и т.д. В ходе работы нужно смоделировать в среде MATLAB 6.5 многоканальную систему с наличием (или отсутствием) сигнала в каждом канале. На базе леммы Неймана-Пирсона в этой многоканальной системе со сравнительно малым числом каналов ($n = 47$) нужно построить байесовский критерий при наиболее неблагоприятном априорном распределении, провести эксперименты для разного числа включенных каналов и для разных уровней сигнала, по результатам экспериментов сравнить эффективность алгоритмов обнаружения на базе полученных кривых обнаружения.

1. Проверка статистических гипотез

1.1. Гипотеза и альтернатива [2]

Часто на основе данных наблюдения из множества наблюдений χ нужно проверить те или иные предположения о распределении вероятностей экспериментальных данных $P = P_\theta$, например, что распределение совпадает с заданным заранее ($\theta = \theta_0$), что оно имеет заданные характеристики (среднее, дисперсию и т.д., $g(\theta) = g_0$) или принадлежит заданному классу распределений (является равномерным, нормальным и т.д.). Любое такое предположение называется *статистической гипотезой* H и выражается соотношением $H : \theta \in \Theta_H$. Здесь Θ_H – некоторое подмножество Θ . Если множество Θ_H состоит из одного элемента $\Theta_H = \{\theta_H\}$, то гипотеза H называется *простой*: $H : \theta \in \theta_H$. Если множество Θ_H содержит больше элементов, то гипотеза H называется *сложной*. Мы будем рассматривать двухальтернативные задачи проверки гипотез. В этом случае одна из гипотез $H_0 : \theta \in \Theta_0$ называется *основной (нулевой) гипотезой*, а другая, $H_1 : \theta \in \Theta_1$, – *альтернативой* $\Theta_0, \Theta_1 \subset \Theta, \Theta_0 \cap \Theta_1 = \emptyset$. Часто в задачах проверки гипотез задается лишь основная гипотеза $H_0 : \theta \in \Theta_0$. При этом обычно считают, что $H_1 : \theta \notin \Theta_0$, т.е. $\Theta_1 = \Theta \setminus \Theta_0$ – дополнение множества Θ_0 в Θ . Такие задачи называются задачами проверки согласия (с гипотезой H_0).

1.2. Тесты проверки гипотез

Правило принятия или отклонения основной гипотезы называют *тестом проверки гипотез* (его называют также *критерием проверки гипотез*). Тест можно рассматривать как функцию наблюдений $\psi(X), X \in \chi$, принимающую значения $\psi(X) = 0$ (это

соответствует *принятию основной гипотезы*) или $\psi(X) = 1$ (это соответствует *отклонению основной гипотезы, т.е. принятию альтернативы*). Тест однозначно определяется одним из двух непересекающихся и дополняющих друг друга подмножеств $\chi_0 = \{X \in \chi : \psi(X) = 0\}$; $\chi_1 = \{X \in \chi : \psi(X) = 1\}$; $\chi_0 = \chi \setminus \chi_1$.

Множество χ_0 называется *допустимым*, а множество χ_1 – *критическим*. Если наблюдаемые данные x попадают в χ_1 , то основная гипотеза H_0 отвергается (принимается альтернатива H_1); в противном случае ($x \in \chi_0$) основная гипотеза H_0 принимается. Обычно критерий задается с помощью статистики критерия $L = L(X)$, $X \in \chi$ и числового порога критерия T :

$$\psi(X) = \begin{cases} 0 & \text{при } L(X) < T, \\ 1 & \text{при } L(X) \geq T. \end{cases}$$

Алгоритм принятия решения состоит из двух этапов:

- 1) вычисление по наблюдениям X значения статистики $L = L(X)$;
- 2) сравнение $L = L(X)$ с порогом T .

Основная гипотеза принимается при $L < T$, отвергается при $L \geq T$.

1.3. Ошибки I и II рода и их вероятности

Решения, принимаемые на основе случайных данных с помощью того или иного теста ψ , могут быть как правильными, так и ошибочными. В задачах проверки гипотез различают ошибки двух типов.

1. Ошибки I рода: отклонение основной гипотезы H_0 , в то время как она справедлива, т.е. $\psi(X) = 1$ при $\theta \in \Theta_0$.

2. Ошибки II рода: принятие основной гипотезы H_0 , в то время как имеет место альтернатива, т.е. $\psi(X) = 0$ при $\theta \in \Theta_1$.

Количественной характеристикой ошибок I и II рода являются их вероятности. Они характеризуют достоверность решений, принимаемых с помощью того или иного теста проверки гипотез.

Вероятность ошибок I рода теста ψ обозначается $\alpha(\psi, \theta)$ и зависит от значения параметра $\theta \in \Theta_0$: $\alpha(\psi, \theta) = P_\theta(\chi_1)$, $\theta \in \Theta_0$.

Уровнем значимости $\alpha(\psi)$ теста ψ называется величина максимальной вероятности ошибок I рода: $\alpha(\psi) = \sup_{\theta \in \Theta_0} \alpha(\psi, \theta)$.

Вероятность ошибок II рода теста ψ обозначается $\beta(\psi, \theta)$ и зависит от значения параметра $\theta \in \Theta_1$: $\beta(\psi, \theta) = P_\theta(\chi_0)$, $\theta \in \Theta_1$.

Часто рассматривают также *функцию мощности* $\gamma(\psi, \theta)$ теста ψ :

$$\gamma(\psi, \theta) = P_\theta(\chi_1) = 1 - \beta(\psi, \theta), \quad \theta \in \Theta_1.$$

Функция мощности $\gamma(\psi, \theta)$ отличается от вероятности ошибок I рода $\alpha(\psi, \theta)$ лишь областью изменения аргумента θ . Тест называется *несмещенным*, если $\gamma(\psi, \theta) \geq \alpha(\psi, \theta)$ для всех $\theta_0 \in \Theta_0$, $\theta_1 \in \Theta_1$.

Для теста ψ , заданного с помощью статистики L и порога T ,

$$\alpha(\psi, \theta_0) = 1 - G_{\theta_0}(T), \quad \theta_0 \in \Theta_0;$$

$$\beta(\psi, \theta_1) = G_{\theta_1}(T), \quad \gamma(\psi) = 1 - G_{\theta_1}(T), \quad \theta_1 \in \Theta_1,$$

где $G_\theta(T) = P_\theta(L < T)$ есть значение функции распределения статистики L для порога T , вычисленное при значении параметра $\theta \in \Theta$.

1.4. Подход Неймана-Пирсона [3]

В задачах проверки гипотез желательно построить такие тесты, у которых вероятность ошибок как I рода, так и II рода были бы минимальны. Это требование, однако, противоречиво: обычно уменьшение вероятностей ошибок I рода влечет увеличение вероятностей ошибок II рода (уменьшение мощности), и наоборот. Например, для теста, основанного на некоторой статистике L , уменьшение вероятностей ошибок I рода требует увеличения порога T , а уменьшение вероятностей ошибок II рода требует уменьшения порога T .

В этой связи обычно используют *подход Неймана-Пирсона*, состоящий в следующем. задается малая величина $\alpha \in (0, 1)$ – максимально допустимая величина вероятностей ошибок I рода (ее называют также *допустимым уровнем значимости*), и рассматриваются только тесты $\psi = \psi_\alpha$, для которых $\alpha(\psi) \leq \alpha$. Выбор величины α зависит от конкретной задачи. Часто выбирают $\alpha = 0.1, 0.05$ или 0.01 ; если ошибочное отклонение основной гипотезы сопряжено с большими потерями, выбирают $\alpha = 0.001$ и менее.

В теории проверки гипотез разрабатываются методы построения тестов проверки различных гипотез при различных альтернативах и изучаются различные характеристики достоверности принимаемых решений, то есть вероятности ошибок и мощность тестов. Наиболее общие результаты удается получить в рамках асимптотического подхода.

1.5. Оптимальная и асимптотически минимаксная проверка гипотез. О сравнении качества оценок [4]

1.5.1. Байесовский подход

Зададим борелевскую σ -алгебру в пространстве вероятностных мер P , и пусть P_0 и P_1 – борелевские подмножества P . Зададим две вероятностные меры π_0 и π_1 с носителями, соответственно, P_0 и P_1 , которые будем называть априорными распределениями на P_0, P_1 . Рассмотрим *усредненные* ошибки для данных распределений.

$$F_1(\psi) = \alpha(\psi, \pi_0) = E_{\pi_0} \alpha(\psi, P), \quad F_2(\psi) = \beta(\psi, \pi_1) = E_{\pi_1} \beta(\psi, P).$$

Затем определим

$$\gamma_1(\psi, \pi_0, \pi_1) = t\alpha(\psi, \pi_0) + \beta(\psi, \pi_1), \quad \gamma(t, \pi_0, \pi_1) = \inf_{\Psi} \gamma_1(\psi, \pi_0, \pi_1), \quad (1.1)$$

$$\beta(\alpha, \pi_0, \pi_1) = \inf_{\Psi_\alpha} \beta(\psi, \pi_1), \quad (1.2)$$

где для $\alpha \in (0, 1)$ множество тестов Ψ_α имеет вид

$$\Psi_\alpha = \Psi_{\alpha, \pi_0} = \{\psi \in \Psi : \alpha(\psi, \pi_0) \leq \alpha\}.$$

Тест $\psi^{(t)}$ или ψ_α называется *байесовским*, если удовлетворяет условиям (1.1) или (1.2). Рассмотрим смеси P_{π_0} и P_{π_1} , заданные равенствами вида $P_\pi(A) = E_\pi P(A)$, $A \in \mathcal{A}$, $\pi = \pi_0$ или $\pi = \pi_1$, $i = 0, 1$. Тогда задача байесовской проверки гипотез сводится к проверке простой гипотезы $P_0 = P_{\pi_0}$ против простой альтернативы $P_0 = P_{\pi_1}$, так как $\alpha(\psi, \pi_0) = \alpha(\psi, P_{\pi_0})$, $\beta(\psi, \pi_1) = \beta(\psi, P_{\pi_1})$. Тогда, используя лемму Неймана-Пирсона, получаем структуру байесовского теста. В целом, если P_{π_0} доминирует P_{π_1} , то байесовские тесты основаны на статистике $L = dP_{\pi_1} / dP_{\pi_0}$ и $\gamma = \gamma(1) = 1 - \frac{1}{2} E_{P_{\pi_0}} |L - 1|$.

Байесовский подход очень важен с теоретической точки зрения. Но, к сожалению, возникают две основные трудности в применении этого подхода. Во-первых, для большинства представляющих интерес моделей довольно сложно изучать оптимальные тесты для гипотез, относящихся к смесям мер.

Вторая трудность заключается в том, что обычно не очень понятно, каким образом выбирать априорные распределения. В принципе они должны представлять информацию, не основывающуюся на эксперименте. Однако часто либо отсутствует такая информация, либо ее трудно представить в форме априорного распределения.

1.5.2. Минимаксный подход

Рассмотрим не средние (усредненные), а максимальные ошибки

$$F_1(\psi) = \alpha(\psi) = \alpha(\psi, P_0) = \sup_{P \in P_0} \alpha(\psi, P),$$

$$F_2(\psi) = \beta(\psi) = \alpha(\psi, P_1) = \sup_{P \in P_1} \beta(\psi, P).$$

Иногда удобно считать, что $\alpha(\psi, P_0) = 0$, если $P_0 = \emptyset$ и $\beta(\psi, P_1) = 0$, если $P_1 = \emptyset$. Для $\alpha \in (0, 1)$ положим $\Psi_\alpha = \Psi_{\alpha, P_0} = \{\psi \in \Psi : \alpha(\psi, P_0) \leq \alpha\}$.

Можно определить величины

$$\gamma_t(\psi) = \gamma_t(\psi, P_0, P_1) = t\alpha(\psi) + \beta(\psi), \quad \gamma(t) = \gamma(t, P_0, P_1) = \inf_{\Psi} \gamma_t(\psi), \quad (1.3)$$

$$\beta = \beta(\alpha) = \beta(\alpha, P_0, P_1) = \inf_{\Psi_\alpha} \beta(\psi). \quad (1.4)$$

Тест $\psi^{(t)}$ или ψ_α называется *минимаксным*, если удовлетворяет условиям (1.3) или (1.4).

Величины $\gamma(t)$ и $\beta(\alpha)$ характеризуют величину суммарной ошибки или ошибки второго рода, которые возможны в задаче.

1.5.3. Асимптотическая минимаксность

Минимаксные свойства семейства тестов ψ_ε характеризуются асимптотикой максимума ошибок I и II рода

$$\alpha_\varepsilon(\psi_\varepsilon) = \sup_{\theta \in \Theta_{\varepsilon,0}} \alpha_\varepsilon(\psi_\varepsilon, \theta), \quad \beta_\varepsilon(\psi_\varepsilon) = \sup_{\theta \in \Theta_{\varepsilon,1}} \beta_\varepsilon(\psi_\varepsilon, \theta)$$

или их суммой $\gamma_{\varepsilon,t}(\psi_\varepsilon) = t\alpha_\varepsilon(\psi_\varepsilon) + \beta_\varepsilon(\psi_\varepsilon)$. Обозначим

$$\gamma_\varepsilon(t) = \inf \gamma_{\varepsilon,t}(\psi_\varepsilon), \quad 0 \leq \gamma_\varepsilon(t) \leq \min(1, t),$$

где нижняя грань берется по всем тестам $\psi_\varepsilon \in \Psi_\varepsilon$. Совокупность тестов $\psi_\varepsilon^{(t)}$ называется асимптотически минимаксной, если $\gamma_{\varepsilon,t}(\psi_\varepsilon^{(t)}) = \gamma_\varepsilon(t) + o(1)$, при $\varepsilon \rightarrow \varepsilon_0$. Исходя из подхода Неймана-Пирсона, рассмотрим совокупности тестов ψ_ε , удовлетворяющие неравенству $\alpha_\varepsilon(\psi_\varepsilon) \leq \alpha + o(1)$. Совокупность тестов $\psi_{\varepsilon,\alpha}$ называется асимптотически минимаксной, если

$$\alpha_\varepsilon(\psi_{\varepsilon,\alpha}) \leq \alpha + o(1), \quad \beta_\varepsilon(\psi_{\varepsilon,\alpha}) \leq \beta + o(1); \quad \varepsilon \rightarrow \varepsilon_0.$$

Мы рассматриваем задачу изучения асимптотики $\gamma_\varepsilon(t)$ или $\beta_\varepsilon(\alpha)$ и построения асимптотически минимаксных семейств тестов $\psi_\varepsilon^{(t)}$ и $\psi_{\varepsilon,\alpha}$.

2. Обнаружение сигнала в многоканальной системе

2.1. Постановка задачи

В статье [1] рассмотрена задача обнаружения сигнала в многоканальной системе, каждый из каналов которой может содержать или не содержать сигнал $f(t)$, $t \in (0,1)$,

наблюдаемый в гауссовском белом шуме. Рассмотрен случай, когда число каналов n велико, и изучается асимптотика при $n \rightarrow \infty$.

Точная постановка задачи состоит в следующем. Рассмотрена модель гауссовского белого шума, т.е. наблюдались n независимых случайных процессов $X_i(t)$, $t \in (0, 1)$, каждый из которых соответствует i -му каналу системы и удовлетворяет стохастическому дифференциальному уравнению

$$dX_i(t) = \xi_i f(t) + \varepsilon dW_i(t), \quad t \in (0, 1), \quad i = 1, \dots, n. \quad (2.1)$$

Здесь $W_i(t)$ – независимые винеровские процессы, $\varepsilon = \varepsilon_n > 0$ – известный уровень шума, $\xi_i \in \{0, 1\}$ – параметры, соответствующие наличию сигнала $f \in L_2(0, 1)$ в i -ом канале ($\xi_i = 1$) или отсутствию сигнала ($\xi_i = 0$), где $L_2(0, 1)$ – пространство измеримых функций на $(0, 1)$ и суммируемых со степенью 2. Рассмотрен случай сигнала известной формы $f(t) = c\phi(t)$, $\phi \in L_2(0, 1)$, $\|\phi\| = 1$; здесь ϕ – известная функция, определяющая форму сигнала, а величина $c = c_n = \|f\| > 0$ определяет уровень сигнала.

За $\bar{f} = (\xi_1 f, \dots, \xi_n f) \in L_2^n(0, 1)$ обозначен набор функций, характеризующих наличие или отсутствие сигнала f в каналах. Величина $K(\bar{\xi}) = \sum_{i=1}^n \xi_i$ есть число каналов, содержащих сигналы; $0 \leq K(\bar{\xi}) \leq n$.

Пусть $F_n = \{\bar{f}\} \subset L_2^n(0, 1)$ – заданное множество наборов сигналов в каналах. Задачу обнаружения сигнала можно рассматривать как задачу проверки гипотезы $H_0: \bar{f} = 0$ против альтернативы $H_{1, F_n}: \bar{f} \in F_n$.

Изучена задача с точки зрения минимаксного подхода. Именно, пусть $\psi_n = \psi_n(X_1, \dots, X_n)$ – тест, т.е. измеримая функция от наблюдений со значениями в интервале $[0, 1]$. Пусть $\alpha_n(\psi_n) = E_{n,0}(\psi_n)$ есть вероятность ошибок I рода и $\beta_n(\psi_n, \bar{f}) = E_{n, \bar{f}}(1 - \psi_n)$ есть вероятность ошибок II рода теста n (здесь и ниже $E_{n, \bar{f}}$ есть математическое ожидание по мере $P_{n, \bar{f}}$, соответствующей наблюдениям (2.1); мера $P_{n,0}$ соответствует отсутствию сигналов в системе). Обозначена за $\gamma_n(\psi_n, F_n)$ сумма вероятности ошибок I рода и максимальной вероятности ошибок II рода

$$\gamma_n(\psi_n, F_n) = \alpha_n(\psi_n) + \max_{\bar{f} \in F_n} \beta_n(\psi_n, \bar{f}).$$

Пусть $\gamma_n = \gamma_n(F_n) = \inf \gamma_n(\psi_n, F_n)$. Величина $\gamma_n(F_n)$ называется минимаксной ошибкой обнаружения для альтернативы. Легко видеть, что $0 \leq \gamma_n \leq 1$.

Будем говорить, что тест ψ_n является минимаксным для альтернативы H_{1, F_n} , если $\gamma_n(\psi_n, F_n) = \gamma_n(F_n)$.

Рассматриваемая задача также представлена в несколько другой форме. Рассмотрим статистику

$$x_i = \varepsilon^{-1} \int_0^1 \phi(t) dX_i(t), \quad i = 1, \dots, n. \quad (2.2)$$

Легко видеть, что это достаточные статистики в рассматриваемой задаче, и они независимы как при гипотезе, так и при альтернативе, соответствующей любому набору $\bar{\xi}_n = (\xi_1, \dots, \xi_n)$. При гипотезе H_0 статистики x_i имеют стандартное гауссовское распределение $N(0, 1)$. Пусть $b = c/\varepsilon$ есть отношение сигнал-шум. Тогда при альтер-

нативе, соответствующей набору $\bar{\xi}_n = (\xi_1, \dots, \xi_n)$, статистики x_i имеют гауссовское распределение $N(v_i, 1)$, $v_i = b\xi_i$. При переходе к статистикам (2.2), возникает вектор наблюдений вида $x = v + \eta$, $v \in R^n$, где $\eta = (\eta_1, \dots, \eta_n)$ – стандартный n -мерный нормальный вектор, а $v = (b\xi_1, \dots, b\xi_n)$ – n -мерный вектор, соответствующий наличию или отсутствию сигнала уровня c в каналах. Получается эквивалентная задача проверки гипотез о среднем n -мерного нормального распределения $H_0 : v = 0$, $H_{1,V_n} : v \in V_n$, где $V_n \subset R^n$ – множество, соответствующее множеству $F_n \subset L_2^2(0, 1)$. Величины $\alpha_n(\psi_n)$, $\beta_n(\psi_n, v)$, $\gamma_n(\psi_n) = \gamma(\psi_n, V_n)$, $\gamma_n = \gamma(V_n)$ в этой задаче определяются аналогично.

Пусть число каналов k , $1 \leq k \leq n$, содержащих сигналы, и отношение сигнал-шум $b < 0$ известны. Пусть

$$V_n(b, k) = \left\{ v = (\xi_1 b, \dots, \xi_n b), \quad \xi_i \in \{0, 1\}, \quad \sum_{i=1}^n \xi_i = k \right\}, \quad (2.3)$$

и рассматривается эквивалентная задача проверки гипотез о среднем n -мерного нормального распределения $H_0 : v = 0$ против альтернативы $H_{1,V_n}(b, k) : v \in V_n(b, k)$. Указывается вид минимаксного теста в этой задаче. Именно, пусть π -равномерное априорное распределение на дискретном множестве $V_n(b, k)$ имеет вид $\pi(v) = M^{-1} \quad \forall v \in V_n(b, k)$; $M = (V_n(b, k)) = C_k^n$.

Рассматривается байесовская задача проверки гипотезы $H_0 : P = P_{n,0}$ против альтернативы $H_1 : P = P_\pi$ о распределении P наблюдений x ; здесь $P_\pi = \int_{R^n} P_{n,v} \pi(dv)$ – смесь по априорному распределению. Пусть ψ_π – байесовский тест, минимизирующий

$\gamma(\psi, P_\pi) = E_{n,0} \psi + E_{P_\pi} (1 - \psi)$. По лемме Неймана-Пирсона он имеет вид $\psi_\pi = 1_{\{L(x) > 1\}}$,

где $L(x) = \frac{dP_\pi}{dP_{n,0}}(x) = \int \exp(-|v|^2/2 + (x, v)) \pi(dv) = (C_k^n)^{-1} e^{-kb^2/2} \sum_{A_{n,k}} \exp(b \sum_{i \in A_{n,k}} x_i)$, а суммирование ведется по всевозможным подмножествам $A_{n,k} \subset \{1, \dots, n\}$, содержащим k элементов. Из соображений инвариантности следует, что есть минимаксный тест для альтернативы $V_n(b, k)$, т.е. $\psi(V_n(b, k)) = \gamma(\psi_n, V_n(b, k))$. Это означает, что есть наиболее благоприятное априорное распределение. Однако структура статистики $L(x)$ достаточно сложна при $k < n$, и трудно исследовать вероятности ошибок минимаксного теста. Кроме того, статистика $L(x)$ существенно зависит от параметров k, b .

В этой связи изучается асимптотический вариант задачи при $n \rightarrow \infty$. Считается, что последовательность тестов ψ_n является асимптотически минимаксной для альтернативы H_{1,V_n} , если $\gamma_n(\psi_n, V_n) = \gamma_n(V_n) + o(1)$, и состоятельной для альтернативы H_{1,V_n} , если $\gamma_n(\psi_n, V_n) \rightarrow 0$. Здесь и далее пределы рассмотрены при $n \rightarrow \infty$.

2.2. Основные результаты

Пусть $k = k_n, b = b_n$ – заданные последовательности, $b_n > 0, 1 \leq k_n \leq n$. Вводится величина $\delta_n = (\log k_n) / (\log n) \in [0, 1]$, т.е. $k_n = n^{\delta_n}$. Полагается

$$u_n^2 = u_n(k_n, b_n) = n^{-1} k_n^2 (e^{b_n^2} - 1), \quad \tilde{u}_n = \tilde{u}_n(k_n, b_n) = n^{-1/2} k_n b_n, \quad T = \sqrt{2 \log n}.$$

$$b_n^*(k_n) = \sqrt{\log(n k_n^{-2})} = \sqrt{(1 - 2\delta_n) \log n}. \quad (2.4)$$

Если $\delta_n \rightarrow \delta \in (0, 1/4)$, то точная асимптотика минимаксных вероятностей ошибок в задаче описывается в терминах специальных безгранично делимых распределений без гауссовской компоненты. Однако можно получить другую точную границу между условиями различимости и неразличимости при $\liminf \delta_n \leq 1/4$.

Пусть при $\delta \in [0, 1/4]$

$$b_n^*(k_n) = (1 - \delta_n^{1/2}) \sqrt{2 \log n}. \quad (2.5)$$

Теорема [1, стр. 92]

Пусть $V_n = V_n(b_n, k_n)$ – множества вида (2.3) и $\delta_n \rightarrow \delta \in (0, 1/4)$. Тогда справедливы утверждения:

1. Если $\limsup b_n / b_n^*(k_n) < 1$, то $\gamma_n \rightarrow 1$.

2. Если $\liminf b_n / b_n^*(k_n) > 1$, то $\gamma_n(\psi_n^{thr}) \rightarrow 0$, где тесты ψ_n^{thr} , основаны на статистиках

$$t_{n,b_n} = A_n \sum_{i=1}^n v(x_i, b_n), \quad A_n = (n(e^{b_n^2} - 1))^{-1/2}, \quad v(t, b) = (e^{-b^2/2+tb} - 1). \quad (2.6)$$

Более того, это справедливо при $\delta_n \rightarrow \delta$ для любого $\delta \in [0, 1/2)$.

Тесты ψ_n^{thr} , являющиеся состоятельными при условиях $\liminf b_n / b_n^*(k_n) > 1$ и $\liminf b_n / b_n^*(k_n) < 1/2$, где определены равенствами (2.4), (2.5), не зависят от параметров b_n и k_n .

3. Моделирование и обнаружение сигналов в многоканальной системе в среде MATLAB [5]

3.1. Исходные данные

Количество каналов системы: $n = 47$. Количество каналов с сигналами: $k = 1; 2; 3$;
4. Отношение сигнал-шум: $b = \sqrt{c \log n}$, где $c = 1, 3, 5$.

3.2. Моделирование статистик $L(x)$

Сигнал многоканальной системы имеет вид: $x = v + \eta$, $v \in R^n$.

Для проверки гипотезы с альтернативой вида $H_0 : v = 0$, $H_{1,V_n} : v \in V_n$ необходимо построить на базе леммы Неймана-Пирсона байесовский критерий при наиболее неблагоприятном априорном распределении n , он базируется на статистике $L(x)$:

$$L(x) = \frac{dP_\pi}{dP_{n,0}}(x) = \int \exp(-|v|^2 / 2 + (x, v)) \pi(dv) = (C_k^n)^{-1} e^{-kb^2/2} \sum_{A_{n,k}} \exp(b \sum_{i \in A_{n,k}} x_i).$$

3.3. Моделирование теста по максимуму

Моделирование статистик $L(x)$ в среде MATLAB с использованием описанного выше метода занимает довольно много объективного компьютерного времени, поэтому байесовский критерий неудобен в применении. В вычислении использовался персональный компьютер на базе процессора Pentium-III, тактовая частота которого составляет 800 МГц. На указанной машине формирование только одного набора из одной тысячи пятисот статистик $L(x)$ для $n = 47$ и $k = 4$ заняло 25 часов компьютерного времени. Всего же таких наборов нужно было смоделировать не менее восьми. Также было проведено тестирование программы и на более мощном компьютере, результаты теста показали, что и в случае увеличения оперативной памяти компьютера и тактовой частоты процессора формирование массива из тысячи $L(x)$ займет не менее 10–12 часов.

Существуют очень простые в применении отдельно для большого числа включенных каналов k асимптотически минимаксные тесты (критерии) (линейные тесты), и отдельно для малых k – оптимальные по порядку различимости (тесты по максимуму), но и те, и другие гарантированно имеют «хорошие» свойства только при больших n . Число каналов 50 – сравнительно малое, поэтому проводится сравнение работы этих приближенных критериев (так как $k = 1 \dots 4$, т.е. малое, то проводился тест по максимуму) с точным (байесовским). Для проведения теста по максимуму:

$$D(\bar{X}) = \max(x_1, \dots, x_{47}), \quad \varphi(X, T) = \begin{cases} 1, & D(\bar{X}) > T, \\ 0, & D(\bar{X}) \leq T, \end{cases} \quad T = \sqrt{c \log n}.$$

$$\text{Ошибка I рода: } \alpha(\varphi(X, T)) = 1 - (\Phi(T))^n.$$

$$\text{Ошибка II рода: } \beta(\varphi(X, T)) = (\Phi(T))^{n-k} (\Phi(T-b))^k.$$

Для каждого k ($k = 1, 2, 3, 4$) получено 3 набора по 500 значений статистик $L(x)$ и $L1(x)$.

3.4. Построение графиков

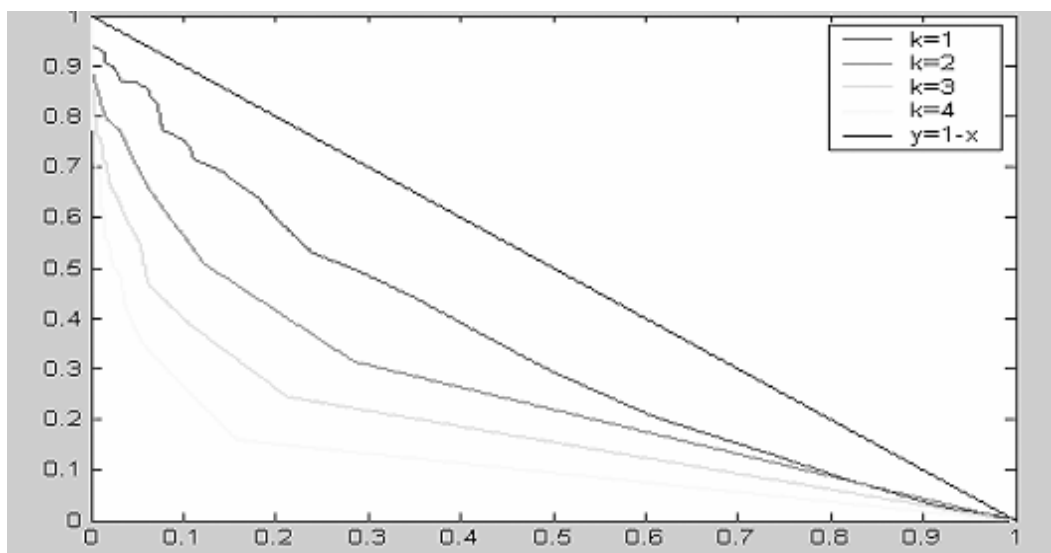


Рис. 3.1. Кривые обнаружения для байесовского теста

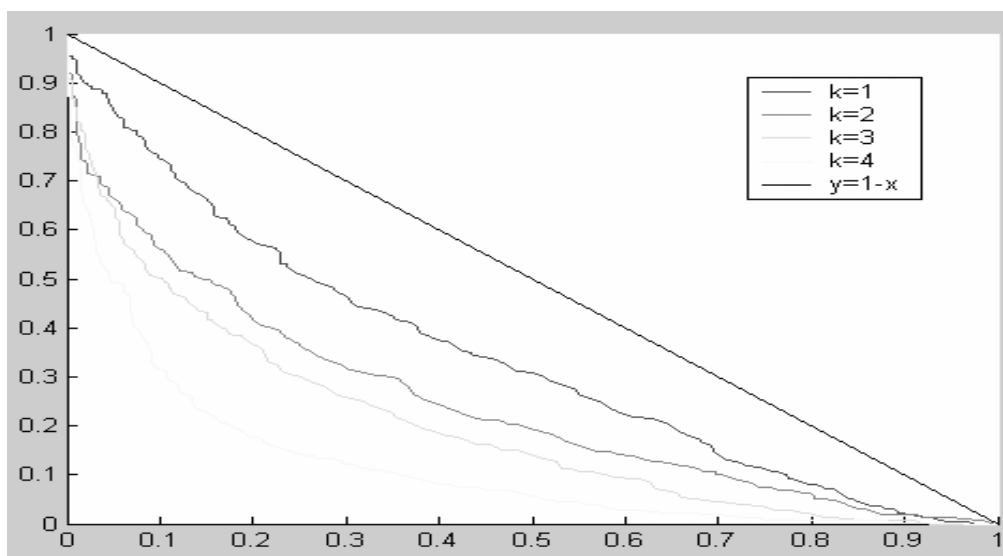


Рис. 3.2. Кривые обнаружения для теста по максимуму

Для анализа полученных результатов, основываясь на полученных статистиках $L(x)$ (только шум) и $L1(x)$ (наличие сигнала), строим кривые обнаружения для байесовского теста, основанного на лемме Неймана-Пирсона. Аналогично строим кривые обнаружения для теста по максимуму.

Из рисунков видно, что при увеличении количества включенных каналов k и/или уровня сигнал-шум b различимость становится лучше и для теста по максимуму, и для байесовского теста, кривые обнаружения располагаются ближе к осям координат.

3.5. Сравнение работы для теста по максимуму и байесовского теста.

Как уже говорилось выше, байесовский критерий, базирующийся на статистике $L(x)$, из-за сложности моделирования статистики неудобен в применении, а простой в применении тест по максимуму гарантированно дает «хороший» результат при больших n . Нужно провести сравнение, которое идет на базе кривых обнаружения теста по максимуму и байесовского теста, а также, по возможности, проследить динамику k .

Ниже приведены графики сравнения кривых обнаружения для байесовского теста и теста по максимуму (см. рис. 3.5, 3.6)

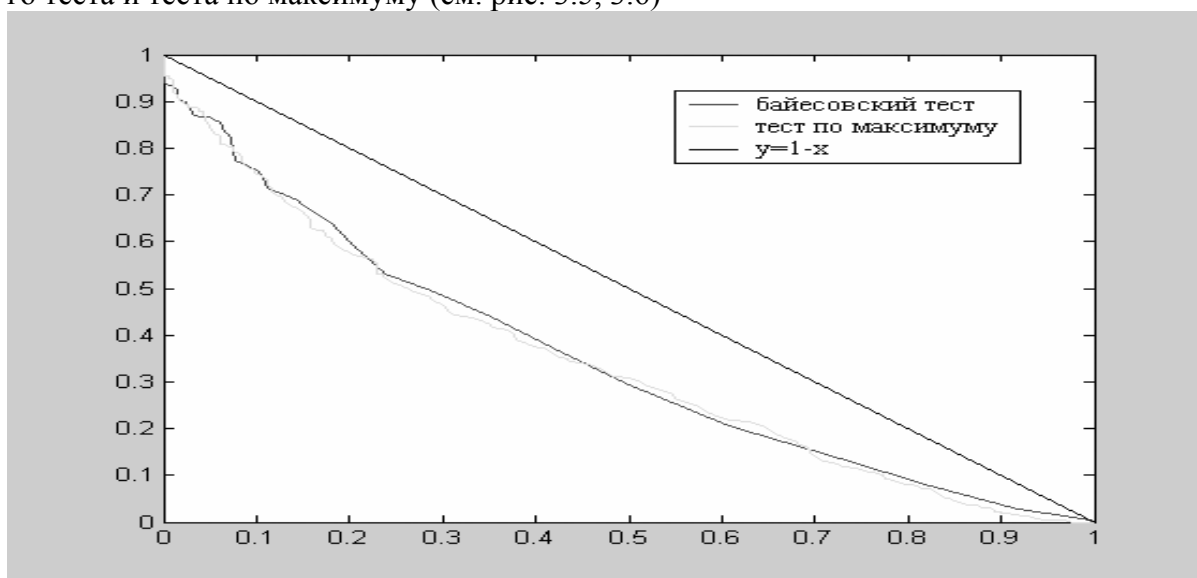


Рис. 3.3. Сравнение кривых обнаружения, при $k = 1$

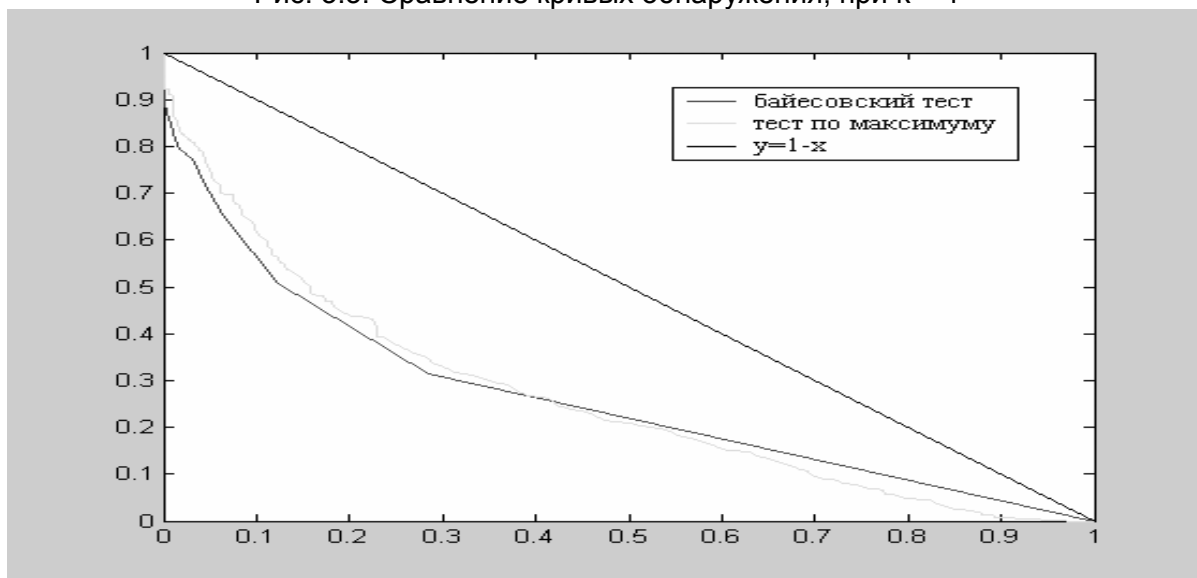


Рис. 3.4. Сравнение кривых обнаружения, при $k = 2$

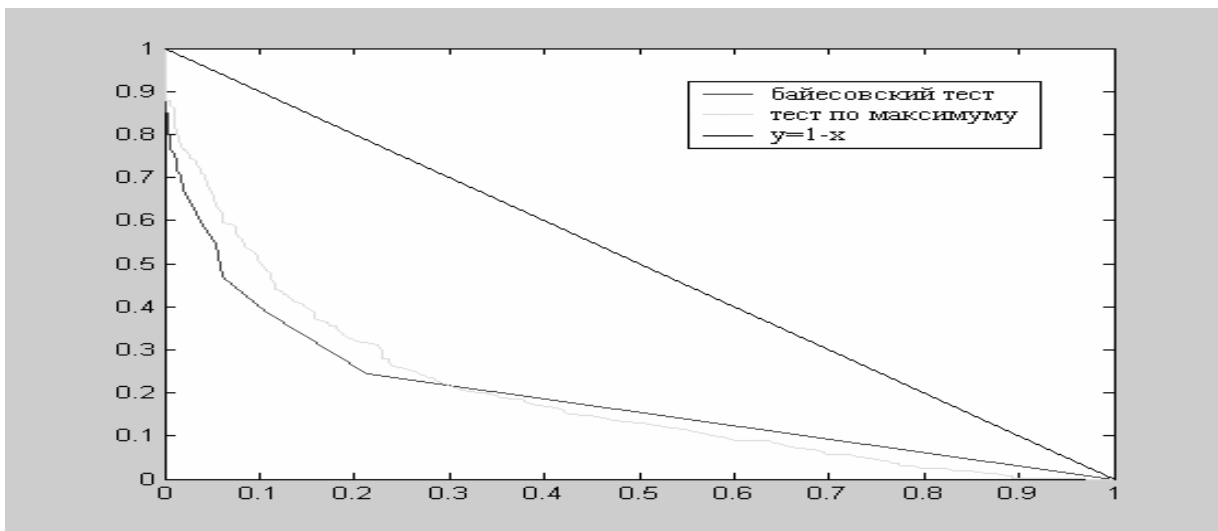


Рис. 3.5. Сравнение кривых обнаружения, при $k = 3$

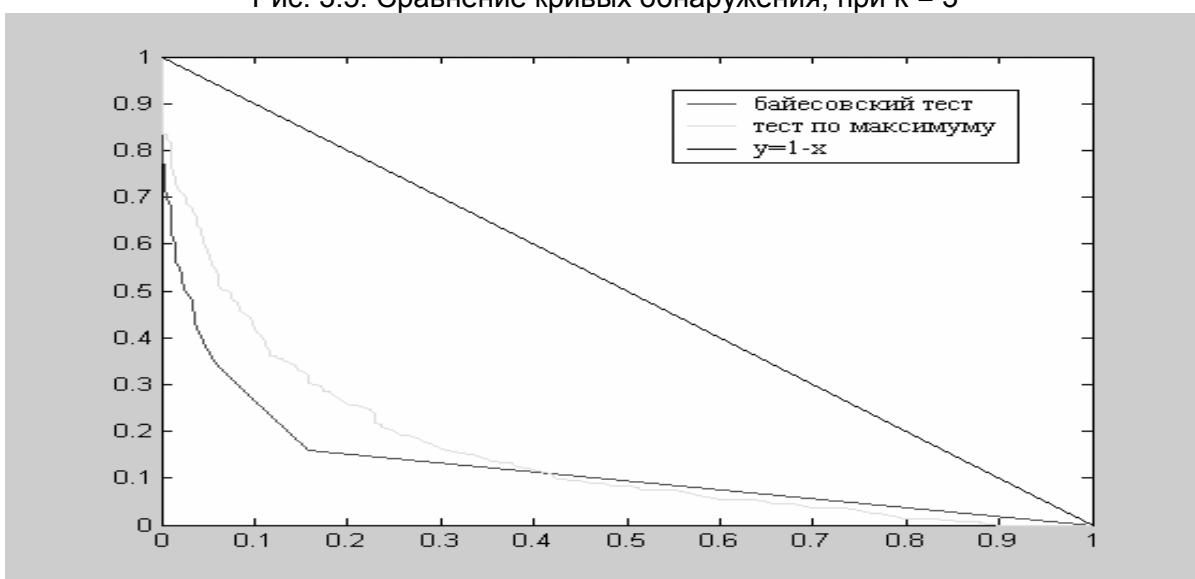


Рис. 3.6. Сравнение кривых обнаружения, при $k = 4$

На рис. 3.5 можно видеть, что кривые обнаружения практически сливаются. При увеличении количества включенных каналов расстояние между кривыми немного увеличивается, но в целом можно признать, что тест по максимуму довольно «хорошо» работает и со сравнительно малым числом каналов. Из рис. 3.6 видно, что различия между кривыми обнаружения незначительны, что свидетельствует о состоятельности теста по максимуму как критерия проверки гипотез.

Заключение

В ходе работы была смоделирована многоканальная система с наличием (или отсутствием) сигнала определенного вида в каждом канале. На базе леммы Неймана-Пирсона в этой системе со сравнительно малым числом каналов ($n = 47$) был построен байесовский критерий при наиболее неблагоприятном априорном распределении. В среде MATLAB также был смоделирован тест по максимуму, который очень прост в применении, но гарантированно имеет хорошие свойства только при больших n . Эксперименты проводились для разного числа включенных каналов и для разного уровня сигнала. По результатам экспериментов выполнено сравнение точных байесовских

тестов и тестов по максимуму (оптимальных по порядку различимости) на базе полученных кривых обнаружения.

В результате анализа экспериментов были сделаны выводы о состоятельности теста по максимуму как критерия проверки гипотез. Сравнение байесовского теста с тестом по максимуму показало, что последний можно применять без существенной потери качества и для многоканальных систем со сравнительно малым числом каналов.

Литература

1. Ингстер И.Ю., Суслина И.А. Об обнаружении сигнала известной формы в многоканальной системе. // Записки научных семинаров ПОМИРАН. 2002. Т. 294. С.88–112.
2. Бодрова Н.А., Родина Т.В., Суслина И.А. Элементы теории вероятностей и математической статистики (Под редакцией В.П. Смирнова). СПб: СПбГУ ИТМО, 2001.
3. Бородин А.Н. Элементарный курс теории вероятностей и математической статистики. СПб: Лань, 1998.
4. Ингстер Ю.И. Асимптотические методы в статистике. СПб: ПГУПС, 2000.
5. Потемкин В.Г. Система MATLAB. М.: ДИАЛОГ-МИФИ, 1997.

ИССЛЕДОВАНИЕ РЕЖИМОВ ЗАХВАТА И АВТОСОПРОВОЖДЕНИЯ ЦЕЛИ В СИСТЕМАХ ПРОСТРАНСТВЕННОГО СЛЕЖЕНИЯ

М.М. Мотылькова

Научный руководитель – д.т.н., профессор В.В. Григорьев

В статье исследуется система пространственного слежения в наиболее распространенных режимах ее работы – захват и автосопровождение цели для различных вариантов построения регуляторов для исполнительных механизмов. Делается вывод о достижении заданных показателей качества

Введение

Автоматический обзор пространства с целью получения информации о находящихся в нем объектах представляет существенный интерес для ряда областей промышленности. В настоящее время возник ряд задач, связанных с исследованием и просмотром больших территорий и пространств. Локационные системы пространственного слежения позволяют измерять с большой точностью координаты и скорость объекта, получать исчерпывающую информацию о его форме, определять характеристики его поверхности.

К системам пространственного слежения относятся следящие системы, предназначенные для автоматического сопровождения и выработки пространственных координат подвижных объектов, перемещающихся в пространстве по неизвестной траектории. Основная задача этих систем – воспроизведение данных о входном изображении, создаваемом излучением от объекта.

В системах автоматического сопровождения входными координатами радиолокационных следящих систем являются угловые положения объекта сопровождения и дальность. Сигналы, определяющие текущие координаты объекта, получаются с помощью пеленгационного устройства и поступают на вход следящих систем. Выходными координатами следящих систем являются текущие координаты объекта сопровождения, передаваемые приборами управления. Выходные координаты отличаются от входных ошибкой системы сопровождения. Автоматическое измерение текущих координат объекта производится путем непрерывного совмещения равносигнального направления с направлением на объект, т.е. устранением угла ошибки, возникающего при перемещении объекта сопровождения и носителя в пространстве [1].

В данной работе исследуются режимы захвата и автосопровождения в следящей локационной станции с коническим сканированием с П-регулятором и И-регулятором, а также с новым типом регулятора, построенным по ключевой схеме.

Функциональный состав системы пространственного слежения

В следящей системе радиолокационной станции антенны, с одной стороны, представляют собой объекты управления, с другой, являются элементами сравнения или измерительными элементами, где из g (текущих угловых координат объекта) производится вычитание y (угловых положений равносигнального направления в пространстве). Ошибки e действуют на пеленгационное устройство, на выходе которого возникают напряжения сигналов ошибки $u_{ош}$. Эти напряжения преобразуются общим трактом в управляющие напряжения, которые подводятся к исполнительным механизмам ИМ, которые развивают скорость вращения $\Omega_{об}$. Исполнительные двигатели через редукторы поворачивают антенны и тем самым изменяют угловое положение равносигнального направления в пространстве, стремясь совместить его с направлением на объект со-

проведения. Главные обратные связи физически замыкаются через пространство и структурно имеют коэффициент передачи «-1».

Функциональная схема следящей системы представлена на рис. 1.

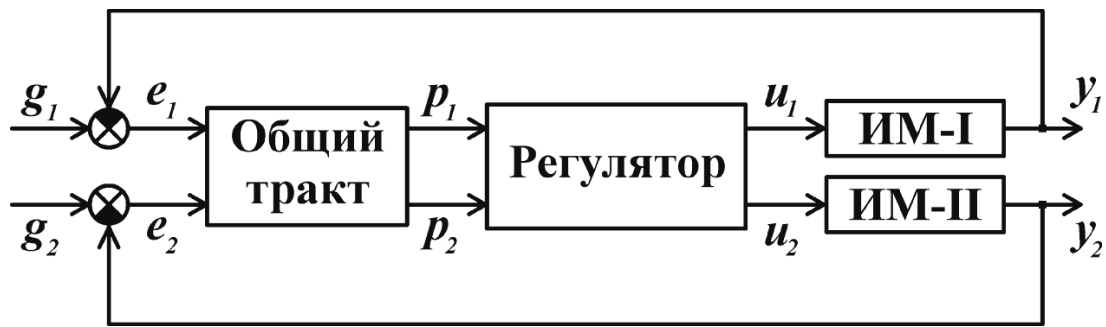


Рис. 1. Функциональная схема следящей системы радиолокационной станции

Следящие системы радиолокационных станций автоматического сопровождения и управления по направлению представляют собой замкнутые автоматические системы, принцип действия которых основан на сравнении действительного значения угловой координаты объекта в пространстве с положением равносигнального направления. С помощью выявляемого при сравнении рассогласования, или ошибки, осуществляется управление антенной таким образом, чтобы равносигнальное направление непрерывно совмещалось с направлением на объект. При этом по угловому положению антенны определяются текущие угловые координаты, а по скорости вращения антенны – угловая скорость объекта относительно станции.

Пеленгационное устройство удерживает в поле зрения подвижный объект, отрабатывая с помощью приводов углов места (УМ) и азимута (УА) рассогласование между направлением на объект и осью направленности излучения (осью визирования). Относительное рассогласование между направлением на объект и осью визирования преобразуется пеленгатором и общим трактом в электрические сигналы, которые поступают на регулятор, управляющий приводами УМ и УА. Выходные сигналы общего тракта, как правило, зависят нелинейно от ошибок по УМ и УА.

В системах пространственного слежения переход от режима поиска к режиму автосопровождения объекта наблюдения называют захватом цели. Под режимом захвата цели будем понимать состояние системы, начиная с момента попадания объекта наблюдения в поле зрения системы до момента, начиная с которого система пространственного слежения сопровождает цель с требуемой для обеспечения функционирования этой системы точностью. Анализ захвата состоит в определении условий захвата – области начальных отклонений, из которой все траектории стягиваются к устойчивому положению равновесия, т.е. осуществляется переход к режиму автосопровождения. Необходимо определить области захвата цели для системы пространственного слежения.

Отслеживание траектории движения цели системой называется автосопровождением. Необходимо определить максимальные скорости автосопровождения, т.е. области определения устойчивости системы в режиме слежения [1].

Моделирование системы пространственного слежения

С помощью пакета моделирования MATLAB «Simulink» была построена схема моделирования следящей радиолокационной станции с общим трактом, по физическим закономерностям своей работы наиболее приближенная к реальным. Подробное описание построения общего тракта можно найти в [2] и [3]. Затем была построена общая схема моделирования, которая приведена на рис. 2. С помощью этой схемы было про-

ведено исследование режима захвата и автосопровождения в следящей радиолокационной станции с коническим сканированием с пропорциональным, интегральным регуляторами и регулятором, построенным по ключевой схеме. При этом исследовались несколько случаев: при различных частотах сканирования, при различном количестве посылок на один оборот сканирования.

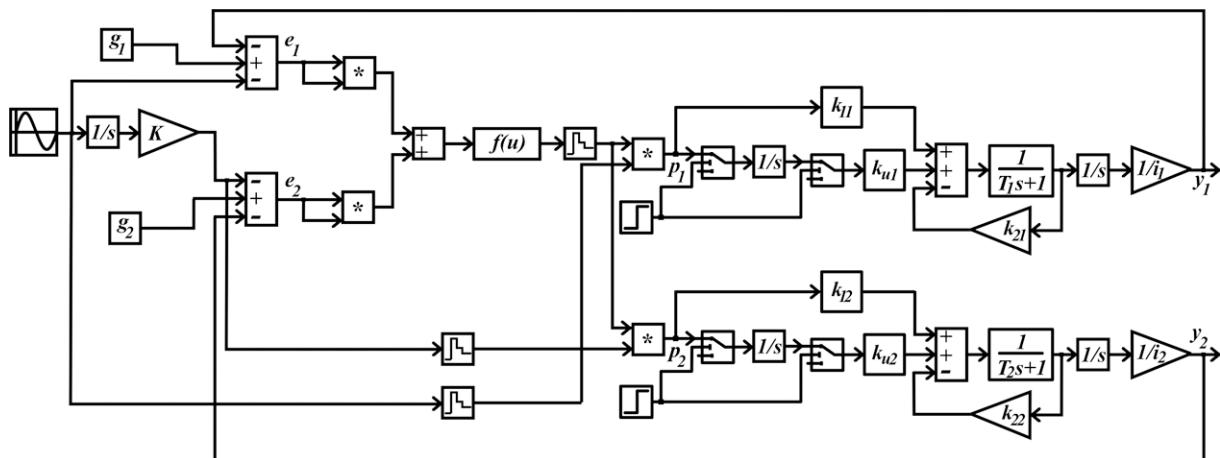


Рис. 2. Общая схема моделирования следящей радиолокационной системы

Вследствие большого количества результатов моделирования приведем лишь часть результатов в виде графиков ошибок (для частоты сканирования 50 Гц и 4 посылок на оборот сканирования). Остальные результаты моделирования сведены в таблицы (табл. 1 и табл. 2 – данные для режима захвата цели с П- и И-регуляторами, соответственно, табл. 3 и табл. 4 – данные для режима автосопровождения цели с П- и И-регуляторами, соответственно).

Частота сканирования, Гц	Количество посылок на оборот сканирования	Время переходного процесса, с	Перерегулирование, %	Амплитуда установившейся ошибки, град	Значение, при котором происходит срыв захвата, град
20	4	0.82	0	0.0164	2.46
50		0.77	0	0.0059	2.45
100		0.785	0	0.00255	2.46
20	16	0.77	0	0.0185	2.26
50		0.8	0	0.0068	2.32
100	12	0.8	0	0.0028	2.33
20	48	0.76	0	0.018	2.25
50		0.78	0	0.00685	2.3
100		0.83	0	0.0028	2.34

Таблица 1. Результаты моделирования для режима захвата цели с пропорциональным регулятором

Частота сканирования, Гц	Количество посылок на оборот сканирования	Время переходного процесса, с	Перерегулирование, %	Амплитуда установившейся ошибки, град	Значение, при котором происходит срыв захвата, град
20	4	0.87	19.83	0.044	2.42
50		0.79	26.36	0.0072	2.58
100		0.9	26.31	0.0017	2.68
20	16	0.82	27.11	0.046	2.25
50		0.79	25.71	0.0086	2.45
100	12	0.83	24.22	0.0022	2.56
20	48	0.86	27.68	0.0465	2.24
50		0.84	25.31	0.0086	2.45
100		0.76	24.31	0.0022	2.55

Таблица 2. Результаты моделирования для режима захвата цели с интегральным регулятором

Частота сканирования, Гц	Количество посылок на оборот сканирования	Значение, при котором происходит срыв слежения, град/с
20	4	25.06
50		4.6
100		4.7
20	16	5.2
50		4.9
100	12	5
20	48	5.2
50		5.1
100		5.1

Таблица 3. Результаты моделирования для режима автосопровождения цели с пропорциональным регулятором

Частота сканирования, Гц	Количество посылок на оборот сканирования	Значение, при котором происходит срыв слежения, град/с
20	4	37.87
50		37.56
100		37.45
20	16	38.01
50		37.08
100	12	36.76
20	48	38.27
50		37.18
100		36.76

Таблица 4. Результаты моделирования для режима автосопровождения цели с интегральным регулятором

На рис. 3–8 приведены результаты моделирования (графики ошибок по каждому из каналов) для режима захвата и автосопровождения цели с параметрами, указанными выше.

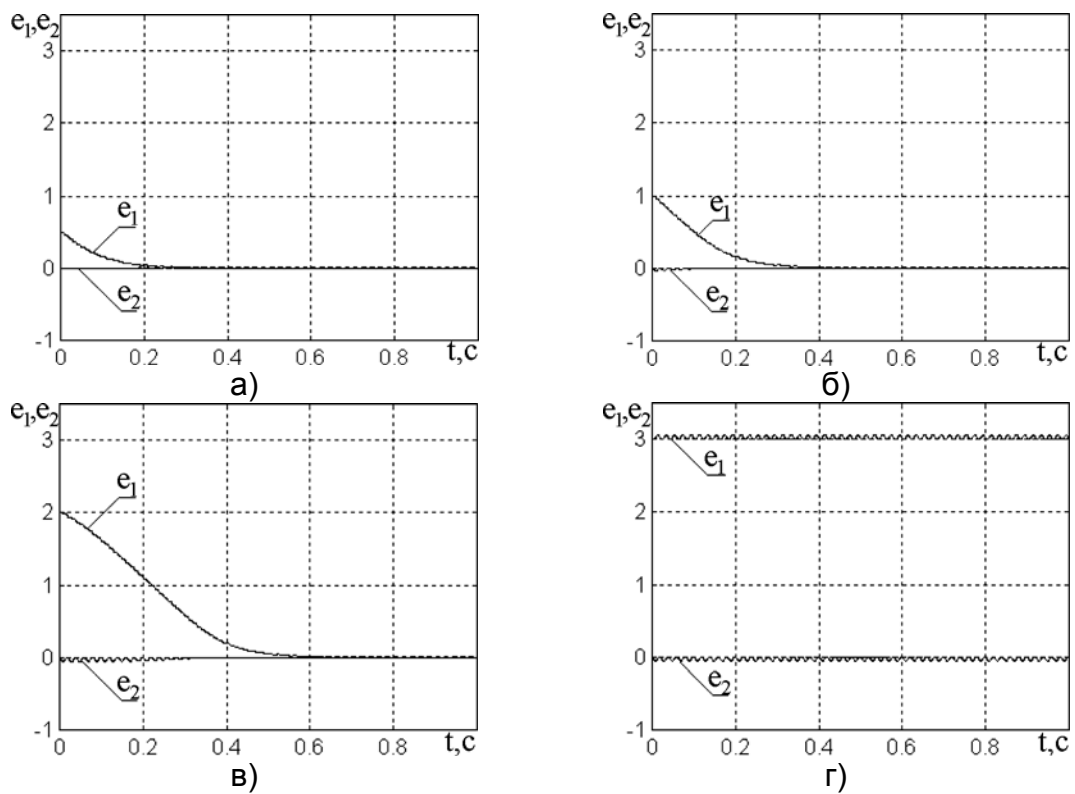


Рис. 3. Графики ошибок для режима захвата цели для системы с пропорциональным регулятором: а) $e_1=0.5$, $e_2=0$; б) $e_1=1$, $e_2=0$; в) $e_1=2$, $e_2=0$; г) $e_1=3$, $e_2=0$

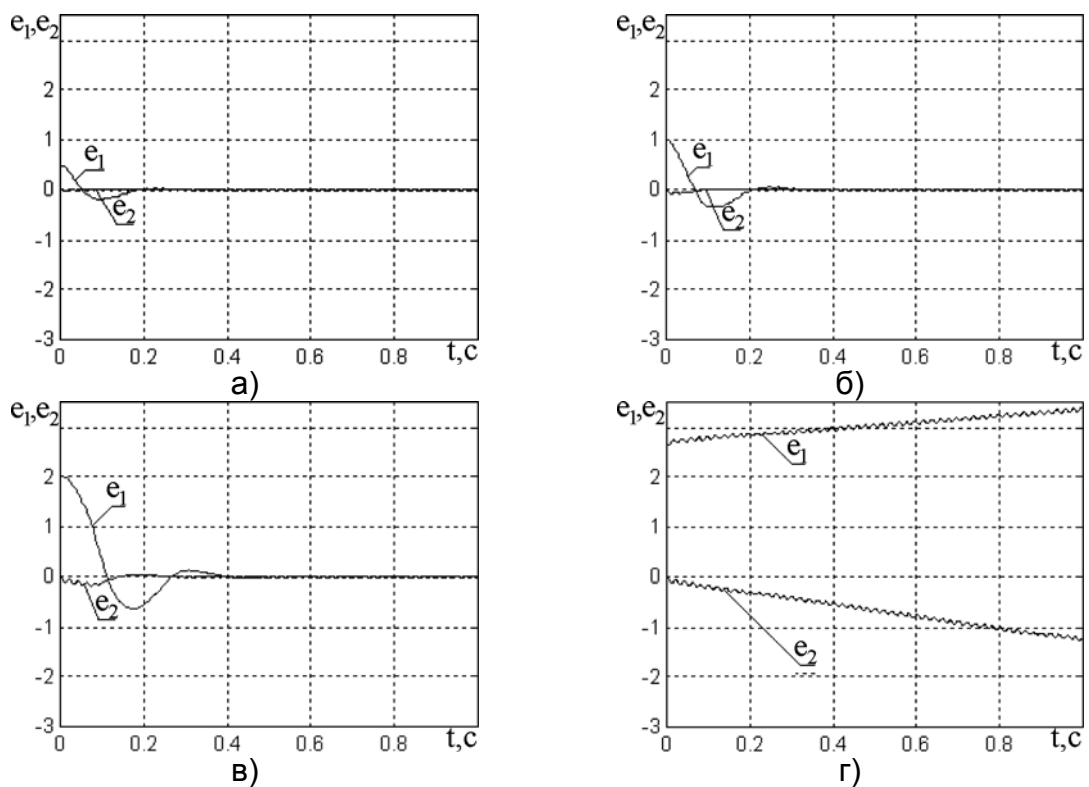


Рис. 4. Графики ошибок для режима захвата цели для системы с интегральным регулятором: а) $e_1=0.5$, $e_2=0$; б) $e_1=1$, $e_2=0$; в) $e_1=2$, $e_2=0$; г) $e_1=2.65$, $e_2=0$

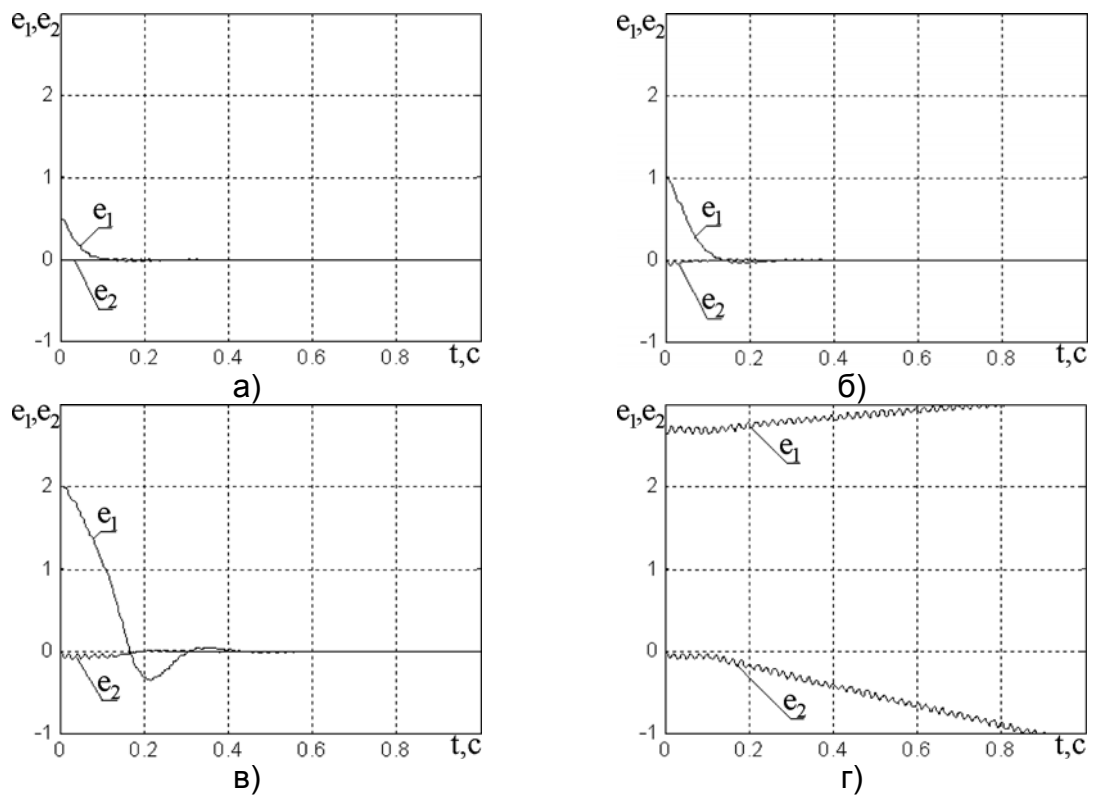


Рис. 5. Графики ошибок для режима захвата цели для системы с ключом:
 а) $e_1=0.5, e_2=0$; б) $e_1=1, e_2=0$; в) $e_1=2, e_2=0$; г) $e_1=2.65, e_2=0$

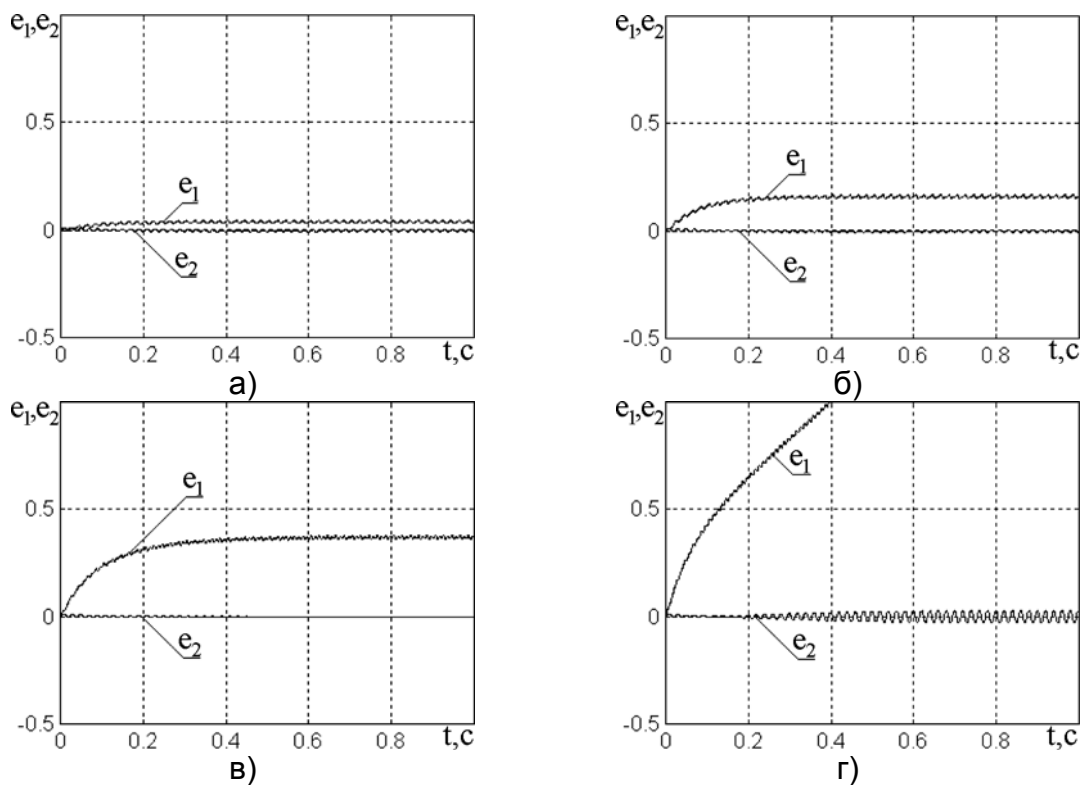


Рис. 6. Графики ошибок для режима автосопровождения цели для системы с пропорциональным регулятором:
 а) $e_1=0.5t, e_2=0$; б) $e_1=2t, e_2=0$; в) $e_1=4t, e_2=0$; г) $e_1=7t, e_2=0$

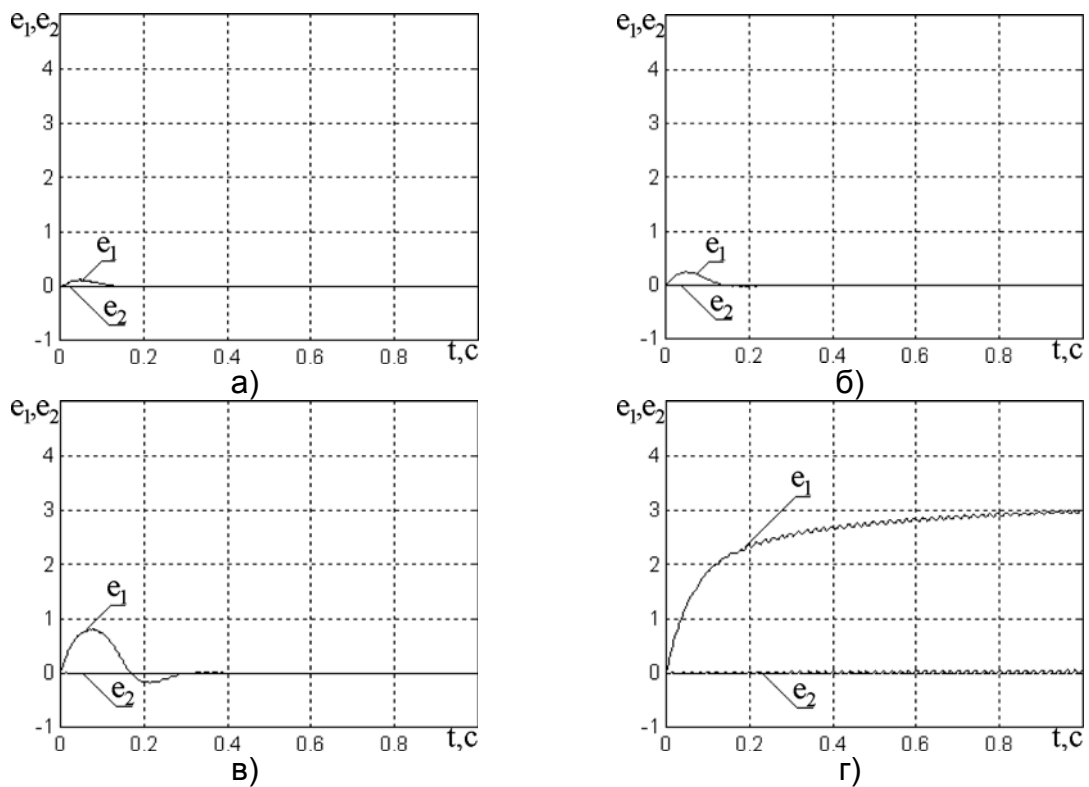


Рис. 7. Графики ошибок для режима автосопровождения цели для системы с интегральным регулятором: а) $e_1=5t$, $e_2=0$; б) $e_1=10t$, $e_2=0$; в) $e_1=25t$, $e_2=0$; г) $e_1=37.57t$, $e_2=0$

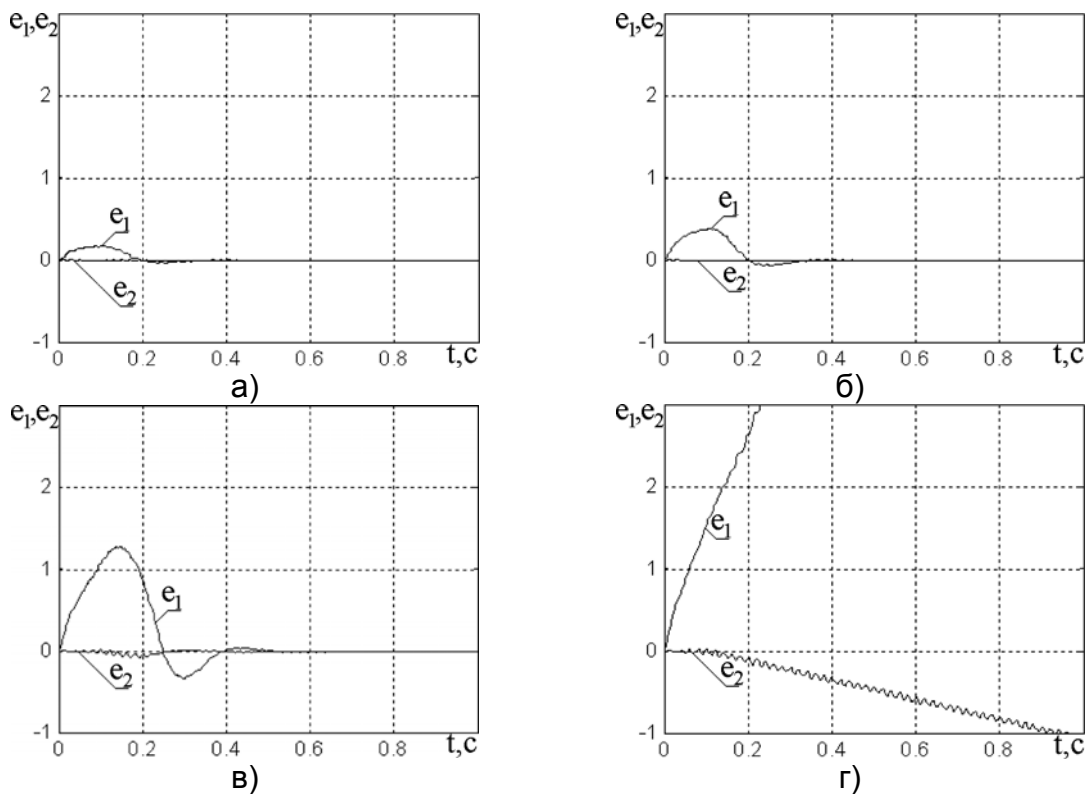


Рис. 8. Графики ошибок для режима автосопровождения цели для системы с ключом: а) $e_1=5t$, $e_2=0$; б) $e_1=10t$, $e_2=0$; в) $e_1=20t$, $e_2=0$; г) $e_1=25t$, $e_2=0$

Заключение

Синтезированная в ходе работы система обеспечивает заданные показатели качества. По результатам моделирования можно сделать вывод о том, что в режиме захвата целесообразнее использовать схему с пропорциональным регулятором, так как обеспечиваются высокие динамические и точностные показатели, а в режиме автосопровождения – с интегральным, так как в этом случае пропорциональный регулятор не обеспечивает высокой точности. Ключевая схема обеспечивает надежный захват, а в режиме автосопровождения – высокую точность слежения.

Литература

1. Артамонов В.М. Следящие системы радиолокационных станций автоматического сопровождения и управления. Л.: Судостроение, 1969. 488 с.
2. Motylkova M., Kozis D. Allocation systems with scanning. // Preprints of BOAC'2004. Saint-Petersburg, 2004. P.221–224.
3. Ткаченко В.Р., Демин А.В. Математическая модель пеленгационной системы с редкими посылками. / Межинститутский сборник. Л.: ЛИТМО, 1989. С.88–93.
4. Бесекерский В.А., Попов Е.П. Теория систем автоматического регулирования. М.: Наука, 1975. 768 с.
5. Бобцов А.А., Мирошник И.В. Линейные системы автоматического управления. СПб: СПбГИТМО (ТУ), 2001. 245 с.

ЭКСПЕРИМЕНТАЛЬНОЕ ОПРЕДЕЛЕНИЕ ИНФОРМАТИВНОСТИ АНАЛОГОВЫХ КАНАЛОВ МИКРОКОНВЕРТЕРОВ*

В.И. Бойков, И.С. Волков, А.Н. Коровьяков

Научный руководитель – д.т.н., профессор В.В. Григорьев

Получены оценки погрешностей аналого-цифровых преобразователей микроконвертеров, порождаемых процессами преобразования информации как непосредственно в аналого-цифровых преобразователях, так и под действием внешних случайных возмущений. Определена информативность аналоговых каналов микроконвертеров в зависимости от технической реализации схемы включения микроконвертеров.

Введение

В настоящее время в цифровых измерительных устройствах находят широкое применение микроконвертеры. Проблема выбора подходящего микроконвертера, стоящая перед разработчиком измерительного устройства, часто упирается в выбор наименее шумного аналого-цифрового преобразователя (АЦП) микроконвертера (МК).

В данном исследовании проводится сравнительная характеристика информативности аналоговых каналов двух микроконвертеров ф. Analog Devices ADuC812 и ADuC841. Произведено сравнение характеристик аналого-цифровых преобразований при использовании нескольких вариантов схемотехнических исполнений аналоговой части. Получены числовые характеристики случайных процессов преобразования и определена оценка информативности микроконвертеров, что позволяет в процессе разработки измерительных устройств наиболее обосновано выбрать подходящий микроконвертер и схемотехническое исполнение.

Описание эксперимента

Для оценки информативности аналоговых каналов в ходе эксперимента на входы 12-и разрядных АЦП МК фирмы Analog Device ADuC812 и ADuC841 подается постоянный во времени и по уровню напряжения входной сигнал. Получаемый в результате работы аналого-цифровых преобразователей объем данных, представляющий собой некоторое случайное распределение, оценивается количественно (при помощи показателей качества) и качественно (графически).

Для количественного описания статистических свойств исследуемых массивов данных, а, следовательно, и МК, было решено использовать среднеквадратичное отклонение (СКО), определяемое, согласно [1], как

$$\sigma_x = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}, \quad (1)$$

где x_i – текущее значение измеряемой случайной величины, \bar{x} – среднее значение измеряемой случайной величины; n – количество измерений.

Для оценки информативности АЦП МК используется эффективное число бит (ENOB), определяемое согласно [2] (с поправкой на условия данного исследования):

$$ENOB = B - \log_2(\sigma_x), \quad (2)$$

где B – номинальное число бит.

Для наглядности и качественного оценивания различных МК и схемотехнических решений в пакетах MS Office Excel и StatGraphics Plus построены следующие графики:

- а) распределения отсчетов в порядке измерения,
- б) эмпирической плотности распределения (гистограммы),

* Работа выполнена при финансовой поддержке РФФИ

в) частных автокорреляционных распределений.

Для проведения экспериментов было выбрано несколько электронных измерительных устройств, в дальнейшем именуемых экспериментальные макеты (см. табл. 1).

Особенности исполнения Макет	Тип МК	Операционный усилитель	Наличие гальванической развязки ЛС
Макет №1 (EVAL-AduC812QS QuickStart ^{nv} Rev:B01)	ADUC812BS 9938 F108672.1	OP491 G 004 C240029	Нет
Макет №2 (PGV2106 4 разъема)	ADUC812BS 0119 F147229.1	LM324D 2AJ038	Нет
Макет №3 (PGV2106 8 разъемов)	ADUC812BS 0119 F147229.1	LM324D 2AJ038	Нет
Макет №4 (I T-403 v 2.1)	ADUC812BS 0119 F147229.1	AD8552 A 513 C9898	Есть
Макет №5 (PGV2106 4 разъема)	ADUC841 BS62-3 F21 0510 628637.1	LM324D 2AJ038	Нет

Таблица 1. Исследуемые экспериментальные макеты

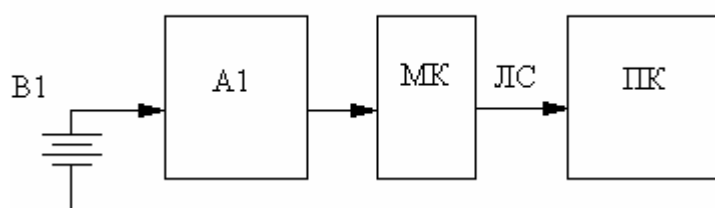


Рис. 1. Структурная схема экспериментального макета: В1 – эталонный источник сигнала (батарея номиналом 1,5 В), А1 – аналоговый усилитель, МК – микроконвертер, ЛС – линия связи RS-232, ПК – персональный компьютер

Макет №1 поставляется производителем в качестве тренировочного макета, его условно можно разделить на две части – заводскую и пользовательскую. В заводской части размещен непосредственно микроконвертер и схема «обвязки» его элементов, необходимых для работы. В пользовательской части размещены элементы, необходимые непосредственно для проведения эксперимента – RC-фильтр, переменный резистор и пр. Все остальные макеты выполнены на печатных платах меньшего размера та-

ким образом, что их нельзя условно разделить на заводскую и пользовательскую части. Все макеты отличаются между собой типом используемых микроконвертеров, операционными усилителями, наличием гальванической изоляции линии связи RS-232 (ЛС), см. табл. 1, при сохранении общей структуры схемы, представленной на рис. 1.

Порядок проведения эксперимента

Опрос данных результата аналого-цифрового преобразования происходит с помощью разработанного ПО: ведомой микропрограммы и ведущей программы, запускаемой с ПК, и цифровой линии связи RS-232, работающей по протоколу ModBus RTU, (см. рис. 2).

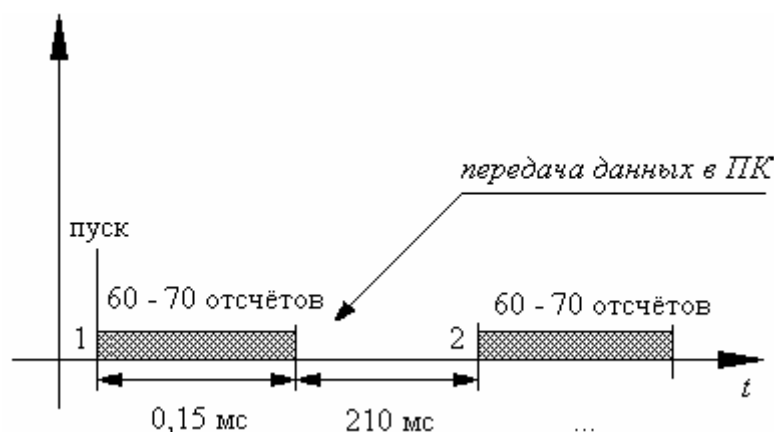


Рис. 2. Порядок проведения эксперимента

Пользователем задается номер опрашиваемого входа АЦП. При записи «1» в определенный регистр МАК происходит пуск аналого-цифровых (АЦ) преобразований, а именно, последовательно опрашивается установленный входной порт АЦП на номинальной частоте работы МК, при этом происходит серия из 60–70 АЦ преобразований (в зависимости от типа МК) величины напряжения на заданном входе АЦП в цифровой код, хранящийся в памяти микроконвертера.

По окончании серии преобразований статического напряжения (возможны отклонения значений, вызванные включением/отключением устройства) массив накопленных значений передается по запросу с ПК из МК в память персонального компьютера и сохраняется в файл с заданным пользователем именем. После этого происходит следующая серия последовательных АЦ преобразований-измерений входного напряжения и дописывается в тот же файл. Таким образом, снимается заданный пользователем массив измерений, кратный количеству измерений в каждой серии (60–70 отсчетов). По этому массиву измерений строятся перечисленные выше графики и рассчитываются указанные количественные оценки.

Обработка экспериментальных данные

В табл. 2 перечислены проведенные эксперименты и их количественные оценки, рассчитанные по формулам (1), (2). Из таблицы видно, что величина ENOB максимальна при использовании МК ADuC841 и ADuC812, размещенного на печатной плате (Макет №1) фирмой-производителем. Это убедительно свидетельствует о качестве исполнения остальных печатных плат. Согласно [3], при проведении гистограммного теста АЦП, стандартизированного IEEE, для получения качественных количественных оценок интегральной и дифференциальной нелинейности (по условиям проведения эксперимента их получить в данной работе не представляется возможным), а также показа-

теля ENOB (эффективного числа бит) необходимо провести порядка полумиллиона измерений (стандарт проведения гистограммного теста), чтобы достигнуть доверительной вероятности 95 %. В нашем случае обработка такого массива информации оказалась бы затруднительной ввиду отсутствия специальных программных решений. Следовательно, выявление характерных графических качественных оценок является наглядным и простым решением, в том числе и для данной работы. В то же время оценка СКО также является информативной для изучения информационной, измерительной способности АЦП МК.

Порядок проведения эксперимента	Номер макета	Количество отсчетов в массиве данных	СКО	ENOB
1	1	4900	4,29	9,90
2	5	60	5,69	9,49
3	5	1790	6,02	9,41
4	5	60	6,61	9,28
5	4	600	21,60	7,57
6	2	600	37,80	6,76
7	3	600	14,80	8,11
8	5	2400	4,90	9,71

Таблица 2. Данные экспериментов

На рис. 3–6 приводятся графики распределения отсчетов в порядке проведения экспериментов.

цифровой код отсчёта

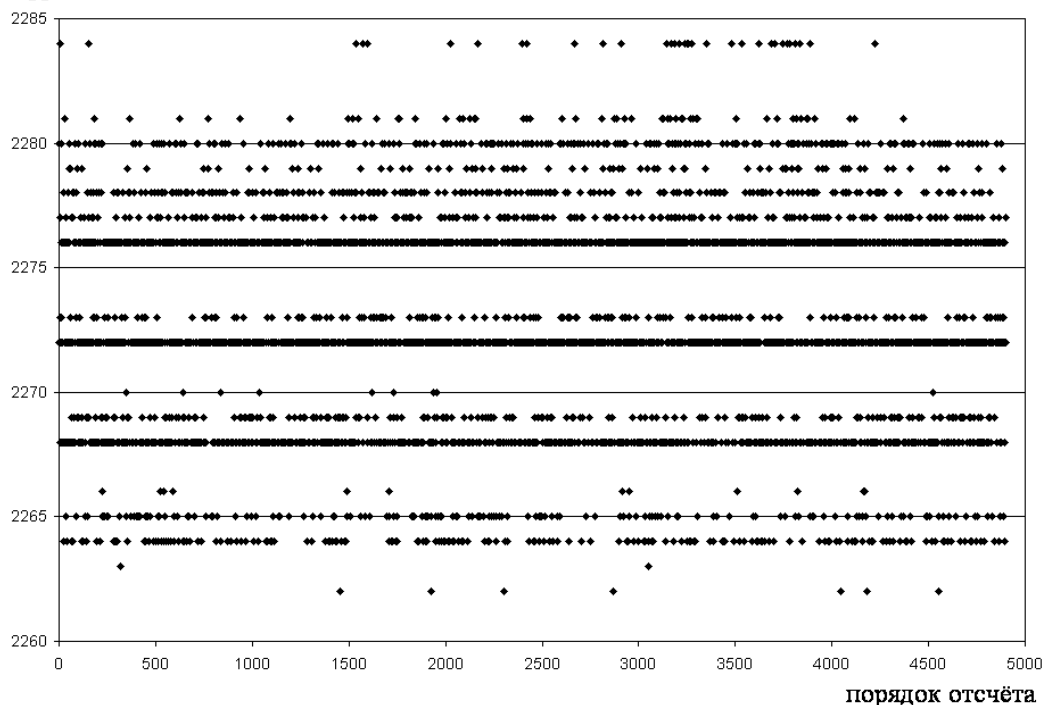


Рис. 3. Распределение отсчетов в 1-ом эксперименте

Согласно рис. 3, в МК серии AduC812 заметны «сильные провалы» (между тремя «сплошными» линиями), по всей видимости, это происходит из-за некорректной работы младших разрядов АЦП.

цифровой код отсчёта

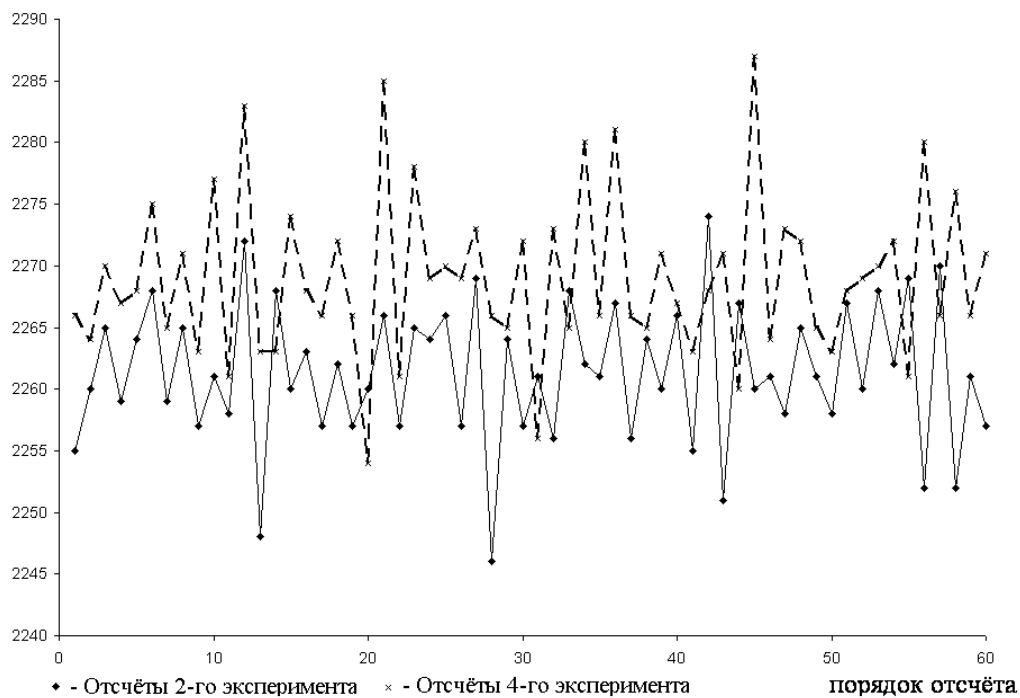


Рис. 4. Распределение отсчетов во 2-ом и 4-ом экспериментах

цифровой код отсчёта

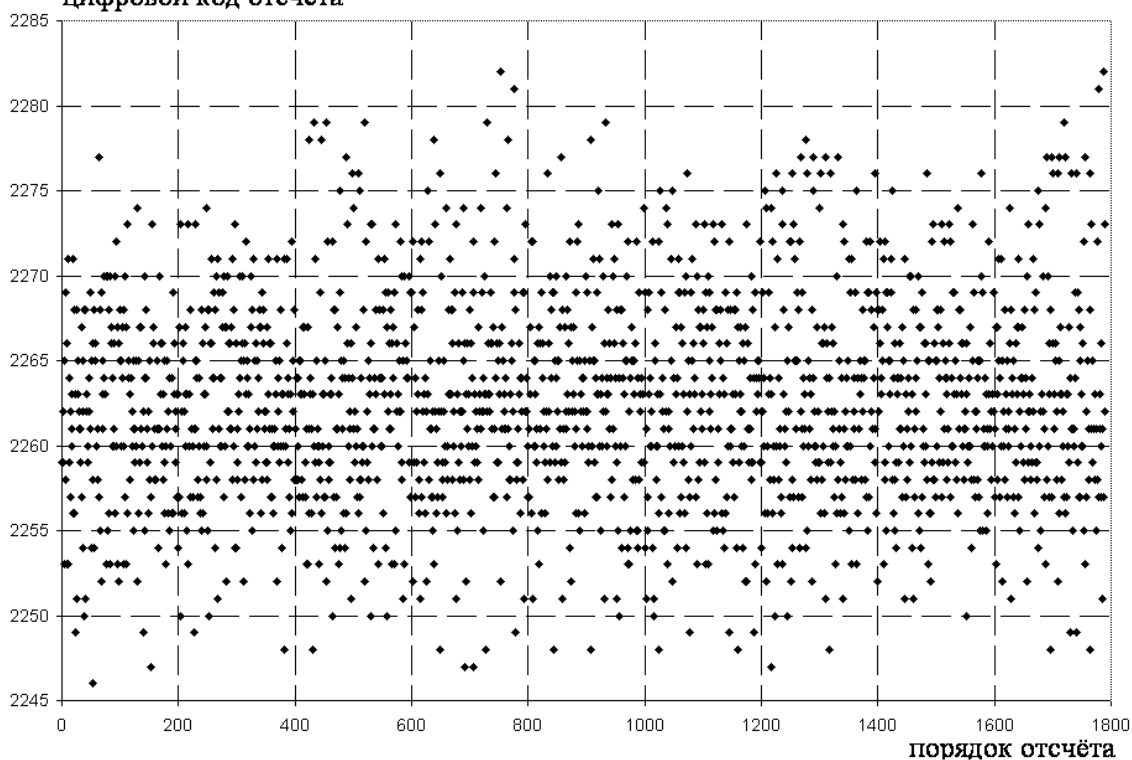


Рис. 5. Распределение отсчетов в 3-ем эксперименте

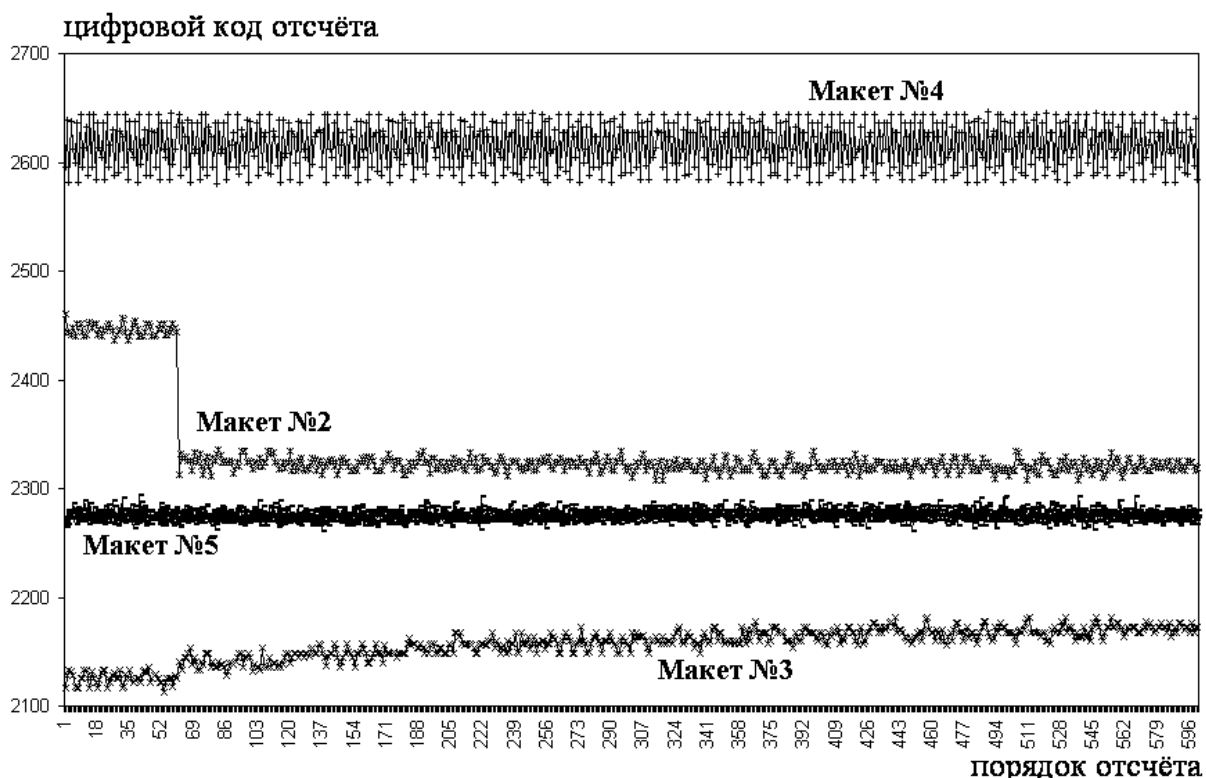


Рис. 6. Распределение отсчетов со 2-го по 5-й эксперименты

Различия в поведении АЦП особенно заметны на рис. 6: видна большая «зашумленность» данных макета № 4, дрейф измерений макета № 3 и скачкообразный переход в значениях макета № 2, что свидетельствует о влиянии конкретного схемотехнического решения на процесс измерения. Стоит отметить, что измерения макетов №№ 2–5 были произведены с минимальным необходимым для переключения схем интервалом.

На рис. 7–9 приводятся гистограммы, построенные на основе экспериментальных данных.

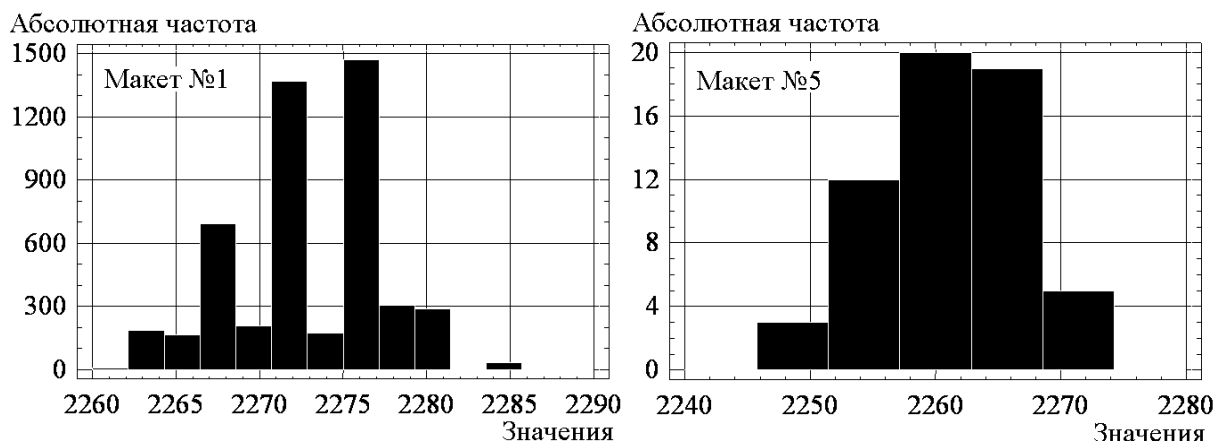


Рис. 7. Гистограммы макетов №1 и №5

На рис. 7 заметны «провалы» в распределении выборок данных ADuC812 (Макет №1) в сравнении с выборками данных ADuC841 (Макет №5).

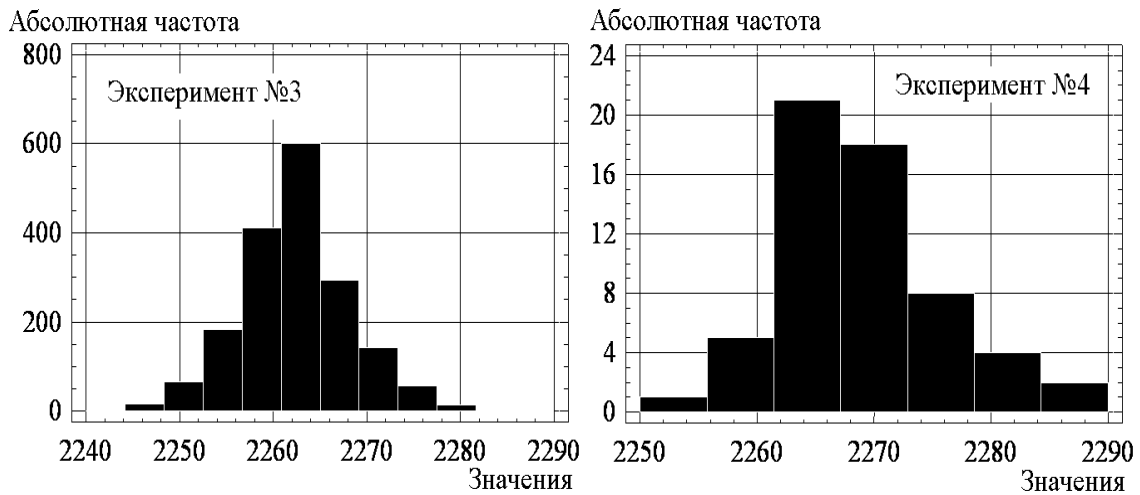


Рис. 8. Гистограммы макета № 5, эксперименты № 3 и № 4

Из рис. 8 ясно, что одна мода проявляется независимо от количества измерений.

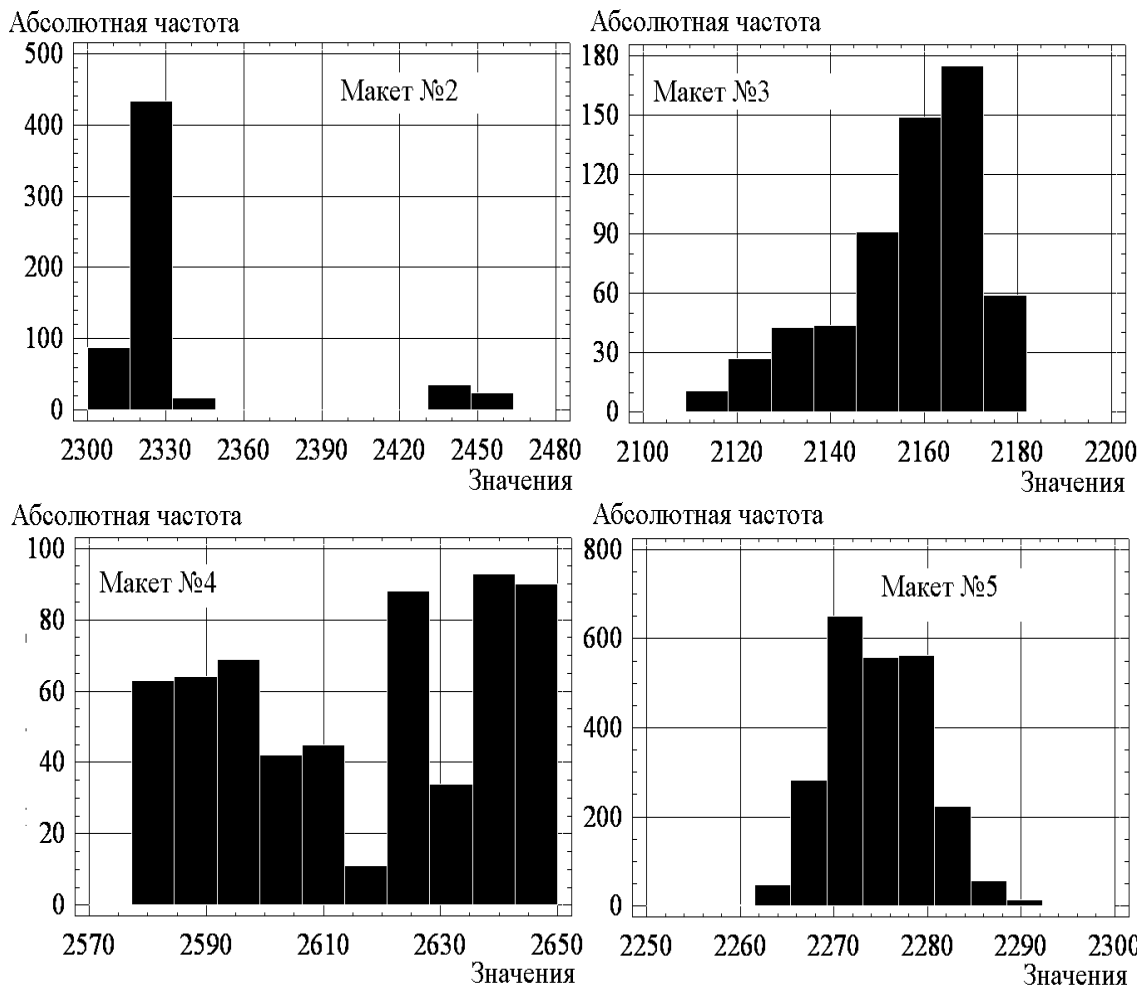


Рис. 9. Гистограммы макетов №№2–5, эксперименты №№ 6, 7, 5, 8, соответственно

Из рис. 9 видно, что разброс выборок данных макета № 2 произошел вследствие скачка значения входного напряжения (см. рис. 6), а разброс выборок данных макета № 3 произошел вследствие «дрейфа» измерений во время снятия данных (см. рис. 6). На гистограмме макета № 4 основная масса значений расположена ближе к границам, что объясняется большим разбросом снятых измерений, произошедших, несмотря на

гальваническую изоляцию линии связи. Лучшие показатели у макета № 5, однако стоит отметить, что на нем отснято в 4 раза больше отсчетов.

На рис. 10 показаны наиболее характерные графики частной автокорреляции для экспериментов с наибольшим числом измерений. Наиболее характерными графиком обладают макеты, показавшие «худшие» измерения.

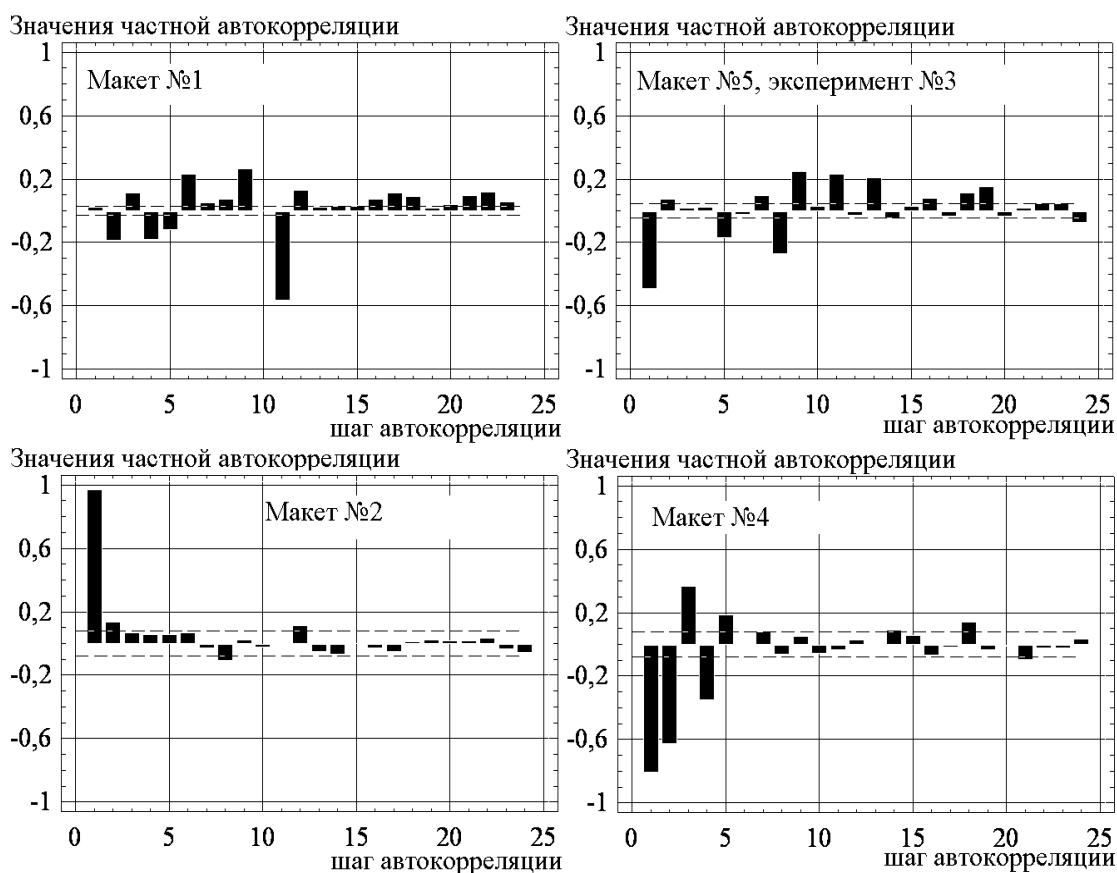


Рис. 10. Графики частной автокорреляции для макетов №№ 1,5,2,4

Заключение

В ходе экспериментов были исследованы микроконвертеры, отличающиеся различными схемотехническими решениями и типом АЦП. В результате проведенных измерений входного сигнала постоянного напряжения была собрана статистическая информация, получены количественные характеристики (СКО и ENOB), построены графики. Это позволило наглядно оценить измерительную и информативную способность МК.

Литература

1. Катранов А.Г., Самсонова А.В. Компьютерная обработка данных экспериментальных исследований. [Текст]: учеб. пособие/ СПб.: Св. План, 2004. 124 с., ил.
2. István Kollár1, Jerome J. Blair. Improved Determination of the Best Fitting Sine Wave in ADC Testing // Manuscript IM-4462, revised and resubmitted to IEEE Trans. Instrumentation and Measurement, June 30, 2005.
3. Histogram Testing Determines DNL and INL Errors. // APPLICATION NOTE 2085: May 31, 2003 – [Ссылка в сети Интернет] http://www.maxim-ic.com/appnotes.cfm/appnote_number/2085

ОЦЕНКА ПОГРЕШНОСТИ ОПРЕДЕЛЕНИЯ ПАРАМЕТРОВ ИНТЕРФЕРЕНЦИОННОГО ПОЛЯ

В.И. Бойков, Д.А. Камнев, А.А. Орешков

Научный руководитель – д.т.н., профессор В.В. Григорьев

Использование интерференционных измерителей микроперемещений позволяет достичь высокой разрешающей способности – до 0,01 длины волны используемого излучения. Целью работы является оценка погрешности определения параметров интерференционной картины. В работе рассмотрен интерференционный измеритель микроперемещений, построенный по схеме интерферометра Майкельсона, излучение которого имеет сферический волновой фронт. Получены аналитические зависимости, позволяющие по известному уровню шумов оценить величину погрешностей вычисления параметров интерференционной картины. Проведенное исследование позволяет выбрать разрешающую способность фотоэлектронной приемной системы измерителя для обеспечения заданной погрешности измерений величины микроперемещений.

Введение

Развитие технических устройств и механизмов в настоящее время проходит в направлении все большего уменьшения размеров самих устройств и диапазонов их действия. Немалую роль в этом играют представляющие огромный интерес и стремительно развивающиеся нанотехнологии. Одно из важных направлений в этой области – обеспечение перемещений и исследование динамики объектов. Широкое применение для проведения микроперемещений получили управляемые пьезоприводы. В зависимости от конструктивного исполнения они могут совершать как линейные, так и угловые перемещения.

Помимо выполнения непосредственно самих микроперемещений, требуется также решить задачи их измерения и обеспечения контроля точности и четкости. Измерение малых перемещений представляет собой достаточно сложную техническую задачу. Одним из решений этой задачи является использование интерференционного оптико-электронного измерителя микроперемещений. Его отличает высокая точность измерений (задаваемая длиной волны рабочего излучения) и высокое быстродействие.

На кафедре систем управления и информатики СПбГУ ИТМО был создан исследовательский стенд интерференционного измерителя микроперемещений. На рис. 1 приведена функциональная схема стенда.



Рис. 1. Функциональная модель

Персональный компьютер осуществляет управление приводом микроперемещений, который, в свою очередь, воздействует на оптическую систему интерференциометра. При этом изменяется вид и параметры интерференционной картины. Съём оптической информации осуществляется камерой, подключенной (например, посредством USB-интерфейса) к ЭВМ, который производит анализ информации и вычисляет величины перемещений. Оптическая часть стенда построена на основе классической схемы ин-

терферометра Майкельсона. Она состоит из монохроматического лазера с расширителем пучка, светоделительной пластинки, неподвижного зеркала, жестко закрепленного подвижного зеркала на перемещающемся объекте измерений и экрана, на котором отображается интерферограмма (рис. 2).

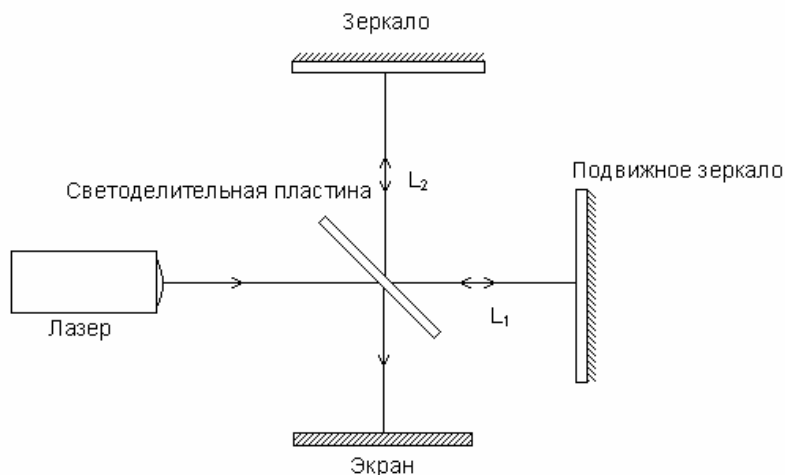


Рис. 2. Оптическая часть измерителя

При изначальном равенстве «плеч» интерферометра L_1 и L_2 наблюдается интерференция, и в центре картины присутствует светлое пятно. При перемещении объекта измерения перемещается и подвижное зеркало, что влечет за собой изменение длин хода разделенных оптических лучей. Поэтому к экрану приходят два когерентных луча, сдвинутые относительно друг друга по фазе. На интерферограмме в результате появляются интерференционные кольца, которые через длину волны излучения характеризуют перемещение подвижного объекта.

Таким образом, интерференционная картина представляет собой набор чередующихся светлых и темных колец, соответствующих максимумам и минимумам волнового фронта (на рис. 3 приведен пример интерференционной картины). Обычно для анализа интерферограм, используемых при контроле целостности объектов, применяют громоздкие математические алгоритмы [2]. Их использование приводит к большим временным затратам, к тому же в ходе вычислений могут возникнуть погрешности. В нашем случае нет нужды пользоваться сложными методами, что позволяет увеличить быстрдействие и упростить алгоритмы обработки оптической информации.

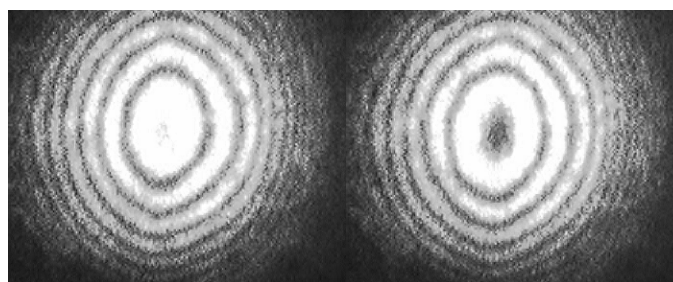


Рис. 3. Интерференционные кольца

При линейном перемещении подвижного зеркала вдоль оси светового пучка кольца на интерференционной картине начинают сходиться к центру картины, при этом происходит соответствующее смещение максимумов и минимумов интерферограммы. Это смещение отражает движение зеркала – зная смещение колец, можно измерить величину микроперемещений. В идеальном случае интерферограмма представляет собой набор концентрических колец, но в реальной установке появляются помехи от неравномерного движения зеркала, погрешностей при снятии оптической информации, ис-

кажения прямолинейности оптических лучей из-за несовершенства оптики, изменения температуры и т.д.

Конфигурация и расположение интерференционных колец определяются тремя величинами – координатами центра окружностей x_u, y_u и радиусом R . Для работы используются три ближайших к центру интерферограммы окружности, так как далее существенное влияние имеют помехи. Определение параметров интерференционных колец позволяет определить положение и перемещение объекта измерений:

$$\tilde{d} = f(d + \xi) = f((x_u, y_u, R_1, R_2, R_3) + \xi). \quad (1)$$

В формуле (1) величина \tilde{d} – оценка перемещения, ξ – помехи в оптическом канале измерителя, $f(\)$ – преобразовательная характеристика измерителя.

Снятие интерференционной картины происходит с использованием видеокамеры. Видеокамера имеет разрешение 640 на 480 пикселей. Из рис. 3 видно, что на реальной интерферограмме кольца являются не четкой окружностью, а некоей областью. При пиксельном сканировании выделяются области колец и центр интерферограммы по пороговой яркости пикселя. Далее путем усреднения выбирается центральная часть интерференционного кольца. Таким образом, создается массив точек N_f , представляющих собой окружность. При этом точки располагаются на некотором расстоянии от идеальной окружности. Это отклонение вызвано ограничением разрешающей способности матрицы видеокамеры, а также помехами при снятии оптической информации.

Окружность для ЭВМ представляется в виде полигонального контура. Вершины контура выбираются на основании информативности контура [3]. Вводится мера информативности контура μ , которая связана с необходимой точностью измерителя. На основании этой меры из массива точек N_f выбираются точки, удовлетворяющие этой мере. При этом решается как задача отсеивания малоинформативных точек, так и точек с очень высоким весом как явно «шумовых». После отсеивания точек остается массив из N_n точек или полигональных вершин, которые представляют собой полигональное отображение окружности. Вокруг этого многоугольника описывается окружность. При этом положения вершин отклоняются от линии окружности на некую величину ε , которая в идеальном случае равна 0. Все вершины удовлетворяют уравнению окружности с учетом этого отклонения (2): оно проявляется в смещении центра окружности и радиуса от идеальных.

$$(x - x_u)^2 + (y - y_u)^2 = R^2 + \varepsilon. \quad (2)$$

Положение центра окружности характеризует угловые отклонения положения плоскости подвижного зеркала от падающего луча. Радиус колец характеризует линейное положение отражающего зеркала на оси хода луча. Координаты положения центра можно найти путем усреднения положения центра относительно всех полигональных вершин.

Имея N_n точек окружности, мы получаем n уравнений (2) окружности, которому удовлетворяют эти точки. При этом нормируем значение R^2 путем введения коэффициента m :

$$\begin{aligned} (x_1 - x_u)^2 + (y_1 - y_u)^2 &= mR^2 + \varepsilon_1, \\ (x_2 - x_u)^2 + (y_2 - y_u)^2 &= mR^2 + \varepsilon_2, \\ \dots \\ (x_n - x_u)^2 + (y_n - y_u)^2 &= mR^2 + \varepsilon_n. \end{aligned} \quad (3)$$

Для нахождения оптимальной окружности, описывающей полученный многоугольник, т.е. для нахождения величин x_u, y_u, R , необходимо найти минимум суммы всех отклонений: $I = \sum_{i=1}^n \varepsilon_i^2 \rightarrow \min$. При достижении минимума производная $\frac{\partial I}{\partial m}$ равна 0:

$$\frac{\partial I}{\partial m} = \frac{\partial \left[\sum_{i=1}^n [(x_i - x_u)^2 + (y_i - y_u)^2 - m]^2 \right]}{\partial m} = 0. \quad (4)$$

Из выражения (4) получаем формулу для вычисления m :

$$m = \frac{\sum_{i=1}^n (x_i - x_u)^2 + (y_i - y_u)^2}{n}. \quad (5)$$

Внесем в формулу (5) пометку определения координат точки (x_i, y_i) :

$$\begin{aligned} (\xi + x_i + x_u) &= (\xi + \tilde{x}), \\ (\xi + y_i + y_u) &= (\xi + \tilde{y}). \end{aligned} \quad (6)$$

Получим:

$$\tilde{m} = \frac{\sum_{i=1}^n (\xi + \tilde{x})^2 + (\xi + \tilde{y})^2}{n} = \frac{\sum_{i=1}^n \xi^2}{n} + \frac{\sum_{i=1}^n 2\xi \tilde{x}}{n} + \frac{\sum_{i=1}^n \tilde{x}^2}{n} + \frac{\sum_{i=1}^n \xi^2}{n} + \frac{\sum_{i=1}^n 2\xi \tilde{y}}{n} + \frac{\sum_{i=1}^n \tilde{y}^2}{n}. \quad (7)$$

В формуле (7) члены $\frac{\sum_{i=1}^n 2\xi \tilde{x}}{n}$ и $\frac{\sum_{i=1}^n 2\xi \tilde{y}}{n}$ стремятся к 0 в силу того, что среднее значение пометки равно 0. Преобразуя формулу (7), получим

$$\tilde{m} = m_n + \tilde{\sigma}_n^2, \quad (8)$$

где $m_n = \frac{\sum_{i=1}^n \tilde{x}^2}{n} + \frac{\sum_{i=1}^n \tilde{y}^2}{n}$, $\tilde{\sigma}_n^2 = \frac{\sum_{i=1}^n \xi^2}{n} + \frac{\sum_{i=1}^n \xi^2}{n}$.

Формула (8) характеризует погрешность определения параметров интерферограммы. Значение погрешности определения зависит от числа точек, по которым строится интерференционное кольцо.

Заключение

В ходе работы рассмотрен интерференционный измеритель микроперемещений, построенный на основе интерферометра Майкельсона. Был рассмотрен процесс снятия оптической информации и ее обработки. Также была оценена погрешность определения параметров интерференционной картины.

Литература

1. Коломийцев Ю.В. Интерферометры. Л.: Машиностроение, 1976.
2. Гуров И.П., Джабиев А.Н. Интерферометрические системы дистанционного контроля объектов. СПб: СПбГИТМО (ТУ), 2000. 190 с.
3. Лепский А.Е., Броневич А.Г., Бачило С.А. Выделение контрольных точек на основе меры информативности контура. / 4-ая Международная Конференция DSPA-2002.

РЕАЛИЗАЦИЯ НАБЛЮДАТЕЛЯ СОСТОЯНИЙ АСИНХРОННОГО ДВИГАТЕЛЯ С КОРОТКОЗАМКНУТЫМ РОТОРОМ В БЕЗДАТЧИКОВОЙ СИСТЕМЕ ВЕКТОРНОГО УПРАВЛЕНИЯ

А.С. Исаков

Научный руководитель – д.т.н., профессор А.В. Ушаков

Рассматривается процесс построения бездатчиковой системы управления. Проводится анализ модельного представления процессов, протекающих в асинхронном двигателе с короткозамкнутым ротором, особое внимание уделено построению наблюдателя, позволяющего отказаться от измерения механических величин. Приведены результаты моделирования сходимости наблюдателя и работы системы управления в целом.

Введение

Для получения высокого качества управления электроприводом в статических и динамических режимах в широком диапазоне регулирования скорости необходимо иметь возможность быстрого непосредственного управления моментом электродвигателя. Момент любого электродвигателя в каждый момент времени определяется величиной (амплитудой) и фазой двух моментобразующих составляющих – тока и магнитного потока. Так как токи и потокосцепления статора и ротора вращаются с разными угловыми скоростями, то они имеют разные изменяющиеся во времени фазовые параметры и не подлежат непосредственному измерению и управлению. Доступной управляемой переменной в асинхронном двигателе является лишь ток статора, который имеет составляющие, образующие магнитный поток и момент. Фазовая ориентация двух этих составляющих может быть осуществлена только внешним управляющим устройством. Иными словами, в асинхронном двигателе необходимо обеспечить управление, как амплитудой, так и фазой тока статора [1].

Для управления фазой тока, а значит, и фазой магнитного поля статора относительно вращающегося ротора необходимо знать точное положение ротора в любой момент времени. Эта задача может быть решена с использованием датчика положения ротора или наблюдателя для оценивания положения ротора. Использование датчика положения в таких системах экономически и конструктивно невыгодно. В системах, построенных с использованием наблюдателя, измерению подлежат лишь токи и напряжения статора асинхронного двигателя. Наблюдатель представляет собой систему дифференциальных уравнений, решаемых в реальном времени. Неизвестными функциями времени в этой системе являются векторы потокосцеплений ротора и скорость ротора. Для устойчивой работы наблюдателя применяется непрерывная коррекция решений системы на основе разностей измеренных и вычисленных значений токов статора. Правильность и эффективность работы наблюдателя зависит от выбора его внутренней структуры и детальности описания модели объекта управления.

Модель асинхронного двигателя с короткозамкнутым ротором

При построении систем векторного регулирования асинхронным электроприводом на основе преобразователей частоты с широтно-импульсной модуляцией (ШИМ) и с микропроцессорным (МК) управлением наиболее целесообразно представление модели двигателя в системе координат, ориентированной по вектору потокосцепления ротора двигателя, поскольку в ней амплитуда и частота выходного напряжения преобразователя частоты, приложенного к статорным обмоткам, присутствуют в явном виде [2].

С учетом вектора случайных воздействий $w(k)$ и вектора случайных помех $q(k)$, сопровождающих измерения, модель будет иметь следующий вид [3]:

$$\begin{cases} x(k+1) = f[x(k), u(k), k] + w(k) \\ y(k) = C \cdot x(k) + q(k) \end{cases}, \quad (1)$$

где

$$\begin{aligned} u &= (u_{S\alpha} \quad u_{S\beta})^T \\ x &= (i_{S\alpha} \quad i_{S\beta} \quad \psi_{R\alpha} \quad \psi_{R\beta} \quad v)^T, \\ y &= (i_{S\alpha} \quad i_{S\beta})^T \\ \begin{cases} f_1 = \left(1 - \frac{T}{T_{S2}}\right) i_{S\alpha}(k) + \frac{x_m}{x_{sig} x_R} \frac{T}{T_R} \psi_{R\alpha}(k) + \frac{x_m \omega_b T}{x_{sig} x_R} v(k) \psi_{R\beta}(k) + \frac{\omega_b T}{x_{sig}} u_{S\alpha}(k) \\ f_2 = \left(1 - \frac{T}{T_{S2}}\right) i_{S\beta}(k) - \frac{x_m \omega_b T}{x_{sig} x_R} v(k) \psi_{R\alpha}(k) + \frac{x_m}{x_{sig} x_R} \frac{T}{T_R} \psi_{R\beta}(k) + \frac{\omega_b T}{x_{sig}} u_{S\beta}(k) \\ f_3 = \frac{x_m T}{T_R} i_{S\alpha}(k) + \left(1 - \frac{T}{T_R}\right) \psi_{R\alpha}(k) - \omega_b T v(k) \psi_{R\beta}(k) \\ f_4 = \frac{x_m T}{T_R} i_{S\beta}(k) + \omega_b T v(k) \psi_{R\alpha}(k) + \left(1 - \frac{T}{T_R}\right) \psi_{R\beta}(k) \\ f_5 = v(k) \end{cases}, \\ C &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix}, \\ T_{S2} &= \left(\frac{r_S \omega_b}{x_{sig}} + \frac{x_m^2}{x_{sig} x_R T_R} \right)^{-1}, \quad T_R = \frac{x_R}{r_R \omega_b}, \quad x_{sig} = x_S - \frac{x_m^2}{x_R} \end{aligned}$$

$u_{S\alpha}, u_{S\beta}, i_{S\alpha}, i_{S\beta}$ – относительные напряжения и токи статора в неподвижной системе координат, $\psi_{R\alpha}, \psi_{R\beta}$ – относительное потокосцепление ротора в неподвижной системе координат, v – относительная скорость вращения ротора, ω_b – относительная скорость вращения системы координат, T – период квантования, x_S, x_R – относительные индуктивности статора и ротора, x_m – относительная взаимная индуктивность между статором и ротором, r_S, r_R – относительные сопротивления статора и ротора.

Предполагается, что система при $w(t) \equiv 0$ и $q(t) \equiv 0$ наблюдаема. Будем считать $w(t)$ и $q(t)$ гауссовскими случайными процессами типа белого шума с нулевыми математическими ожиданиями

$$M\{w(t)\} = 0, \quad M\{q(t)\} = 0$$

и ковариационными матрицами

$$\text{cov}(w) = M\{ww^T\} = Q$$

$$\text{cov}(q) = M\{qq^T\} = R$$

Наблюдатель

Для наблюдателя (Н) вектора состояния асинхронного короткозамкнутого двигателя (АД) входными величинами являются напряжение, подводимое к двигателю, выходными – фазные токи, а поток ротора и скорость – восстановленными переменными состояния. Оценка потока ротора позволяет вычислить ориентацию вектора потокосцепления без использования датчика положения ротора. На рис. 1 приведена функциональная

схема системы управления без датчика положения ротора, где были приняты следующие обозначения: АИ – автономный инвертор, ПК – преобразователь координат, U_a, U_b, U_c и i_a, i_b, i_c – напряжения и токи статора в неподвижной трехфазной системе координат, U_α, U_β и $i_{S\alpha}, i_{S\beta}$ – напряжения и токи статора в неподвижной двухфазной системе координат, v_{ref}, v – задающее значение скорости и скорость вращения ротора, Θ – угол поворота вращающейся системы координат относительно неподвижной.

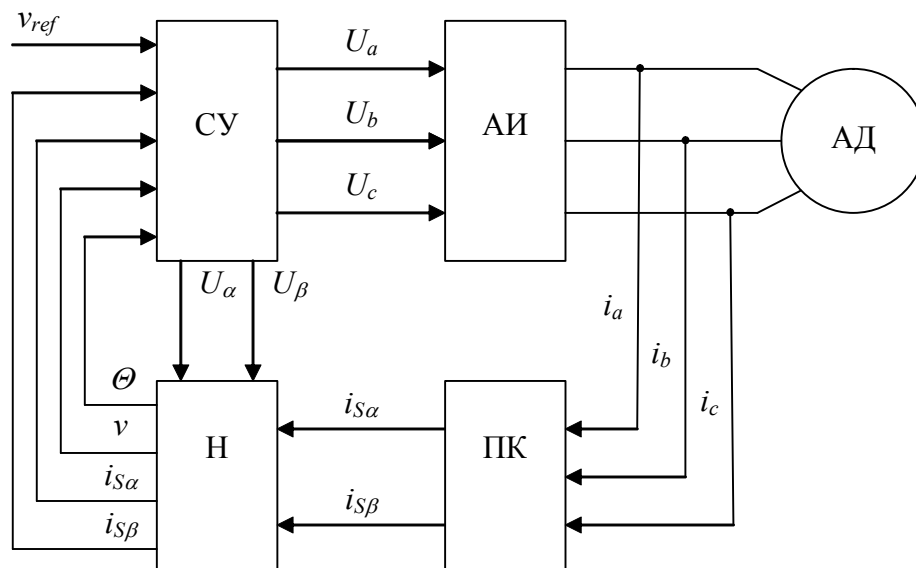


Рис. 1. Функциональная схема бездатчиковой системы управления

Точность такого наблюдателя достаточно сильно зависит от точности параметров модели и точности измеренных величин. Для асинхронных двигателей существует некоторая неопределенность параметров двигателя (температурное изменение) и погрешность в измерениях фазных напряжений и токов. Это обстоятельство приводит к отклонению оценочных переменных вектора состояния от реальных переменных и к некачественному управлению двигателем. Для преодоления данных проблем воспользуемся фильтром Калмана для нелинейных систем. В отличие от других наблюдателей [4], его структура обладает наименьшей чувствительностью к случайному изменению параметров двигателя и к воздействию шумов на измеряемые величины [5].

В литературе [6, 7], для оценки потокосцепления и скорости ротора, широкое распространение получили системы типа MRAS (адаптивное управление с эталонной моделью), где используются два векторных уравнения потокосцепления ротора, одно из которых является базовой, а второе – настраиваемой моделями. Такое построение системы позволяет оценивать только одно состояние системы, и наблюдатель чувствителен к воздействию шумов, хотя его структура значительно проще фильтра Калмана, а его реализация требует меньше вычислительных ресурсов. Это построение подходит для недорогих решений с невысокими требованиями к показателям качества управления.

Фильтр Калмана (ФК) представляет собой стохастический наблюдатель для нелинейных систем. Он предназначен для оценки вектора состояния системы при воздействии на систему и измерения случайных возмущений. Как и обычный наблюдатель, ФК содержит предоценщик (построенный на основе модели системы) и корректор (на основе ошибки наблюдения). В основе корректирующей части лежат коэффициенты обратных связей по ошибке наблюдения. Для ФК эти коэффициенты являются переменными и рассчитываются каждый период выборки, поскольку зависят от шума системы. Структура ФК приведена на рис. 2.

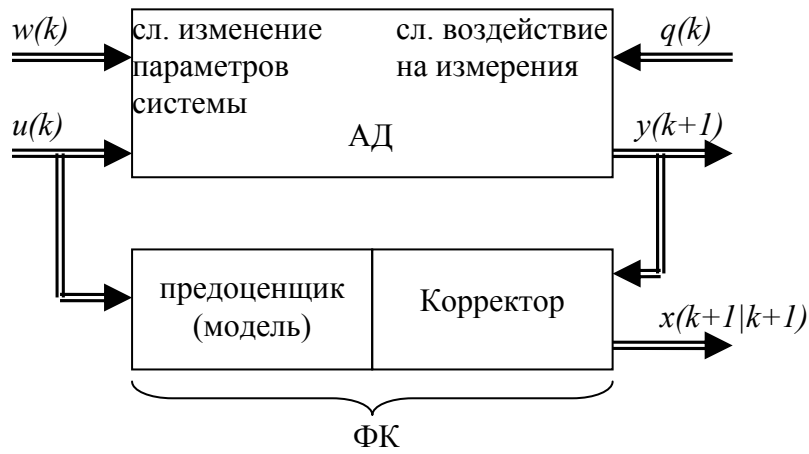


Рис. 2. Фильтр Калмана

Рассмотрим алгоритм работы ФК [3]. Первым шагом алгоритма является предоценка следующих значений вектора состояний $x(k+1|k)$. Она находится из соотношения $x(k+1|k) = f[x(k|k), u(k), k]$. (2)

Для получения реального вектора x необходимо учитывать ошибки вычисления и измерения. Вектор состояния в момент времени $k+1$ может быть вычислен как сумма предоцененного и скорректированного значения:

$$x(k+1|k+1) = x(k+1|k) + K(k+1) \cdot \{y(k+1) - C \cdot x(k+1|k)\}, \quad (3)$$

где $K(k+1)$ – матрица коэффициентов Калмана, которая вычисляется по формуле

$$K(k+1) = P(k+1|k)C^T \{CP(k+1|k)C^T + R\}^{-1}. \quad (4)$$

Здесь R – матрица ковариации ошибок измерения и $P(k+1|k)$ – матрица ковариации ошибок оценивания. Последняя матрица должна рассчитываться каждый период выборки по следующей формуле:

$$P(k+1|k) = A(k)P(k|k)A^T(k) + Q, \quad (5)$$

где Q – матрица ковариации ошибок параметров системы, $P(k|k)$ – матрица ковариации ошибок фильтрации и $A(k)$ – матрица линеаризованной системы. Матрица $A(k)$ вычисляется по следующей формуле:

$$A(k) = \left. \frac{df}{dx} \right|_{x=x(k|k)} = \begin{pmatrix} 1 - \frac{T}{T_{S2}} & 0 & \frac{x_m T}{x_{sig} x_R T_R} & \frac{x_m \omega_b T}{x_{sig} x_R} \nu(k|k) & \frac{x_m \omega_b T}{x_{sig} x_R} \psi_{R\beta}(k|k) \\ 0 & 1 - \frac{T}{T_{S2}} & -\frac{x_m \omega_b T}{x_{sig} x_R} \nu(k|k) & \frac{x_m T}{x_{sig} x_R T_R} & -\frac{x_m \omega_b T}{x_{sig} x_R} \psi_{R\alpha}(k|k) \\ \frac{x_m T}{T_R} & 0 & 1 - \frac{T}{T_R} & -\omega_b T \nu(k|k) & -\omega_b T \psi_{R\beta}(k|k) \\ 0 & \frac{x_m T}{T_R} & \omega_b T \nu(k|k) & 1 - \frac{T}{T_R} & \omega_b T \psi_{R\alpha}(k|k) \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Матрица представляет собой линеаризованную систему относительно вектора состояния и рассчитывается каждый период выборки. Алгоритм оценки также требует обновления матрицы ковариации ошибок фильтрации по следующей формуле:

$$P(k+1|k+1) = \{I - K(k+1)C\}P(k+1|k). \quad (6)$$

На рис. 3 представлена диаграмма, суммирующая шаги работы алгоритма ФК.

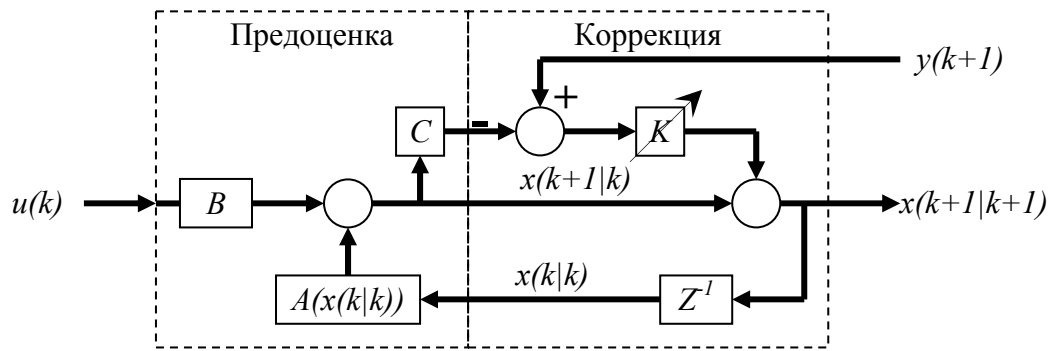


Рис. 3. Диаграмма работы ФК

Бездатчиковая система векторного управления

Векторное управление асинхронным электроприводом строится в системе координат, ориентированной по вектору потокосцепления ротора, по принципу подчиненного регулирования с независимым регулированием потокосцепления ротора и скорости вращения ротора асинхронного двигателя [8].

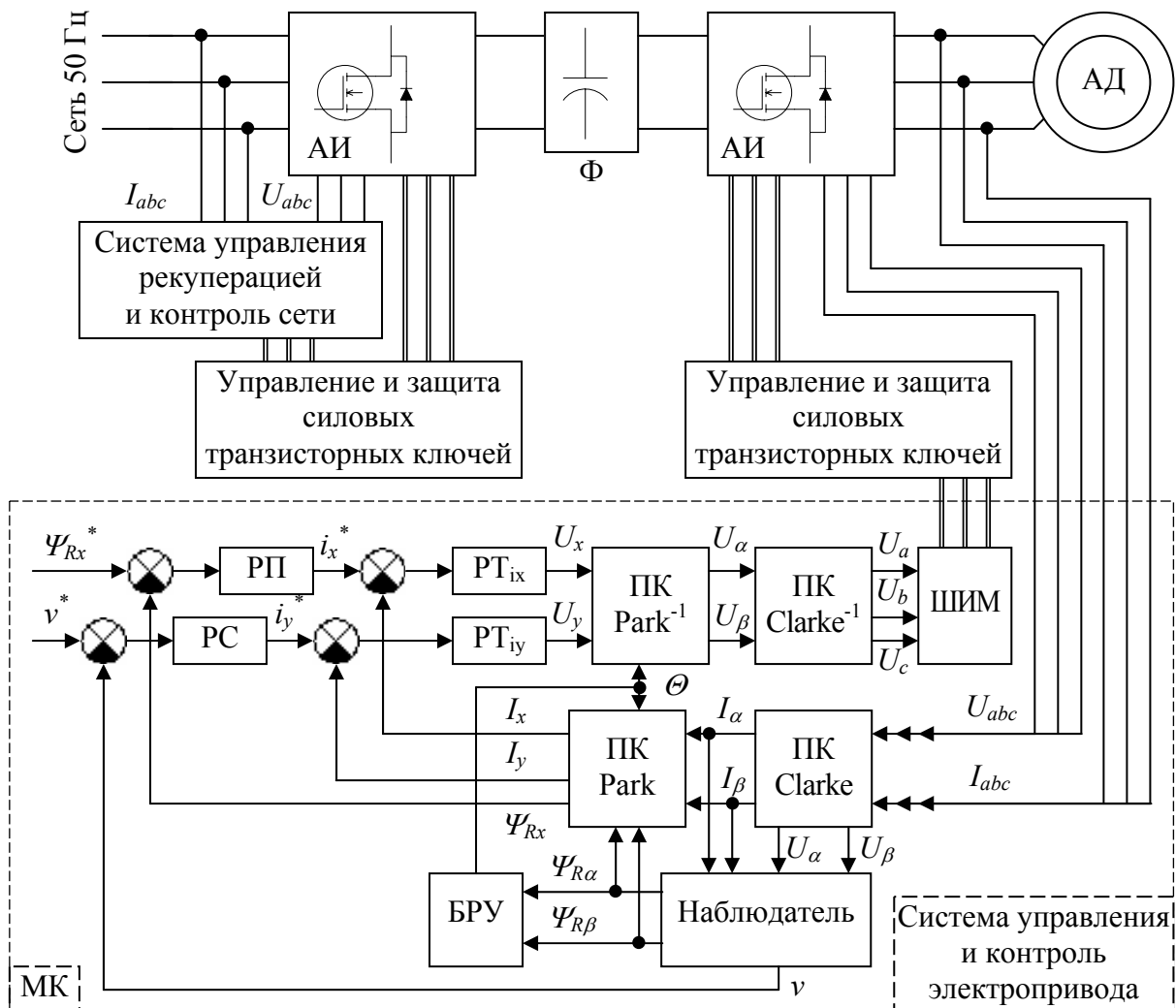


Рис. 4. Функциональная схема системы векторного управления АД

На рис. 4 показана функциональная схема системы векторного управления асинхронным электроприводом, на которой введены следующие обозначения: Φ – фильтр

звена постоянного тока, БРУ – блок расчета угла поворота системы координат, РП и РС – регуляторы потокосцепления и скорости, РТ_{ix} и РТ_{iy} – регуляторы тока активной и реактивной составляющих, U_x, U_y и i_x, i_y – напряжения и токи статора во вращающейся системе координат, $\Psi_{R\alpha}, \Psi_{R\beta}, \Psi_{Rx}$ – потокосцепления ротора в неподвижной и вращающейся системах координат.

Результаты моделирования

На рис. 5–8 представлены результаты моделирования бездатчиковой системы векторного управления асинхронным двигателем с короткозамкнутым ротором в пакете моделирования MatLab/Simulink. На рис. 5 представлен график отработки задающего воздействия по скорости в относительных величинах, на рис. 6–8 – графики относительной ошибки сходимости вектора состояний наблюдателя. Помимо исследований свойств наблюдателя, был проведен анализ динамики системы управления в целом, анализ системы во всем диапазоне скоростей, анализ системы при влиянии возмущающих воздействий в канале измерения и при изменении параметров двигателя.

Исследования показали, что внешние возмущения в канале измерения тока и изменения параметров двигателя (таких как сопротивление ротора) вследствие воздействия окружающей среды (нагрева) не влияют на работу наблюдателя. Это позволяет сделать вывод о правильности выбора структуры наблюдателя.

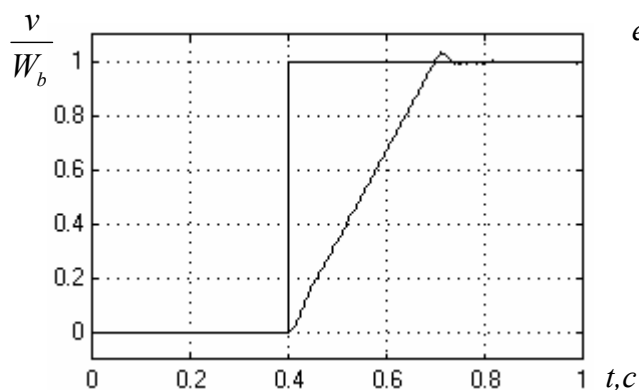


Рис. 5. График скорости

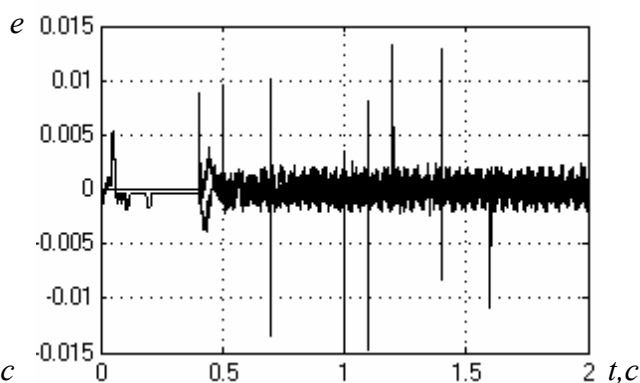


Рис. 6. График ошибок слежения по токам

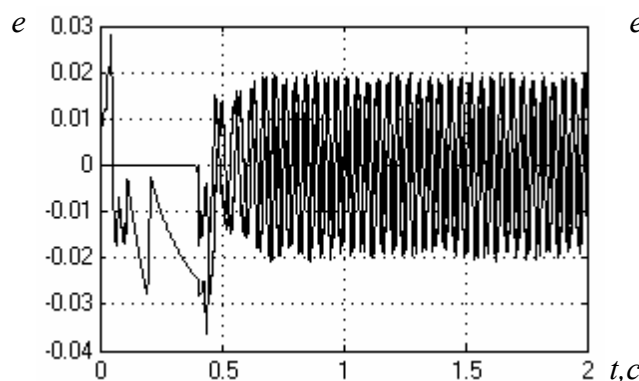


Рис. 7. График ошибок слежения по потокосцеплению ротора

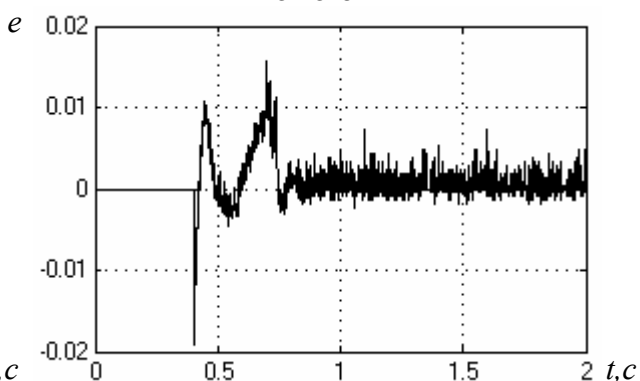


Рис. 8. График ошибки слежения по скорости

Заключение

В статье рассмотрен процесс построения бездатчиковой системы управления асинхронным двигателем с короткозамкнутым ротором, особое внимание уделено по-

строению наблюдателя, позволяющего отказаться от измерения механических величин. Структура наблюдателя выбрана таким образом, чтобы она обладала наименьшей чувствительностью к случайному изменению параметров двигателя и к воздействию шумов на измеряемые величины. В конце статьи приведены результаты моделирования сходимости наблюдателя и работы системы управления в целом.

Литература

1. Башарин А.В., Новиков В.А., Соколовский Г.Г. Управление электроприводами. Л.: Энергоиздат. Ленингр. отделение, 1982. 392 с., ил.
2. Герман-Галкин С.Г. Компьютерное моделирование полупроводниковых систем в MatLab 6.0: Учебное пособие. СПб.: КОРОНА принт, 2001. 320 с., ил.
3. Sensorless Field Oriented Speed Control of Three Phase AC Induction Motor using TMS320F240. <http://www.ti.com>, 1998. 74 с., ил.
4. D. Atkinson, P. Acarnley and J.W. Finch. Observers for induction motor state and parameter estimation. // IEEE Trans. Industry Applications. 1991. V. 27. № 6. P.1119–1127.
5. Балакришнан А. Теория фильтрации Калмана: Пер. с англ. М.: Мир, 1988. 168 с., ил.
6. F.Z. Peng and T. Fukao. Robust speed identification for speed sensorless vector control of induction motors. // IEEE Trans. Industry Applications. 1994. V. 30. № 5. P.1234–1239.
7. C. Schauder. Adaptive speed identification for vector control of induction motor without rotational transducers. // IEEE Trans. Industry Applications. 1992. V. 28. № 5. P.1054–1061.
8. Усольцев А.А. Векторное управление асинхронными двигателями. СПб: СПбГИТМО (ТУ), 2002. 42 с., ил.

УПРАВЛЕНИЕ В УСЛОВИЯХ ЗАПАЗДЫВАНИЯ

А.А. Пыркин

Научный руководитель – д.т.н., профессор А.А. Бобцов

Работа посвящена проблеме управления в условиях запаздывания и параметрической неопределенности. Рассматривается линейный стационарный объект управления с запаздыванием в каналах управления и измерения. Представлен адаптивный идентификационный подход, позволяющий получать упреждающие оценки выходной переменной и вектора переменных состояния, а также оценки всех параметров объекта.

1. Введение

В последнее время наблюдается повышенный интерес к системам управления с запаздыванием. Это объясняется тем, что в большинстве производственных процессов имеются запаздывания, которыми нельзя пренебречь, ибо их влияние на динамику процесса весьма велико.

Стремительный рост требований к темпам развития производства и качеству продукции способствовал бурному развитию автоматизации производственных процессов. Повышение темпов привело к тому, что управляемые величины стали изменяться во времени, причем с большими скоростями. В этих условиях стало играть важную роль время от момента измерения значения регулируемой величины до момента принятия решения и выдачи сигнала управления. Запаздывание реакции управляющей системы на возникшее нарушение процесса приводит, как правило, к возникновению автоколебаний в замкнутой системе, а нередко и к потере устойчивости системы.

Синтез систем управления для объектов с запаздыванием требует учета влияния величины запаздывания на устойчивость и качество переходных процессов в замкнутой системе [1]. Поэтому возникает заманчивая идея компенсировать запаздывание, получив устройство упреждающей оценки. В случае, когда параметры объекта известны, эта задача решается с использованием регуляторов Смита [2] и Ресвика для непрерывных систем и регулятора Цыпкина для дискретных.

В данной работе рассматривается задача синтеза адаптивного предиктора для нахождения упреждающей оценки регулируемой переменной и вектора переменных состояния в условиях параметрической неопределенности.

2. Основная часть

2.1. Проблематика

Эффект запаздывания при наличии обратной связи особенно ярко проявляется при автоматическом управлении удаленными объектами (рис. 1).



Рис. 1. Управление удаленными объектами

Транспортное запаздывание может возникать в силу конструктивных особенностей системы управления. Например, при автоматическом управлении впрыском топлива в инжекторном двигателе анализатор выхлопного газа достаточно проблематично поместить непосредственно в камеру сгорания. На рис. 2 схематично показано устройство системы впрыска.

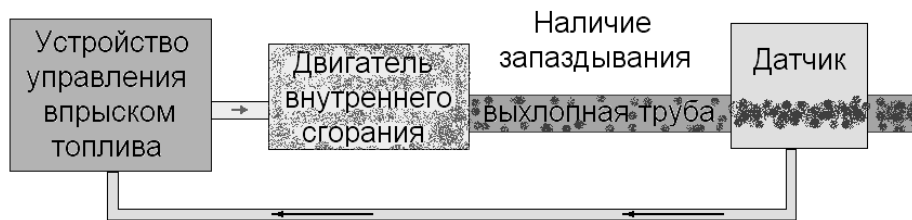


Рис. 2. Наличие транспортного запаздывания в инжекторном двигателе

Запаздывание может быть вызвано сложными вычислительными устройствами (рис. 3), контроллерами, требующими время на расчет управляющего воздействия. Реализуя сложные алгоритмы управления, необходимо учитывать временные затраты процессора.

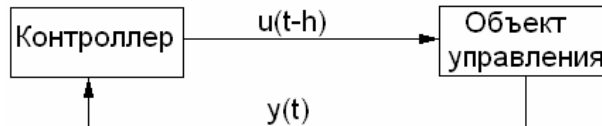


Рис. 3. Запаздывание, вносимое при расчете

При синтезе законов управления сложными химическими реакторами запаздывание имеет место в силу особенностей протекания химических реакций. Также запаздывание можно встретить, работая с экологическими системами, транспортными системами и многими другими.

2.2. Классический подход

Рассмотрим простейший объект управления, описываемый апериодическим звеном первого порядка. Закон управления определяется ПИД-регулятором. Показатели качества определяются коэффициентами регулятора.

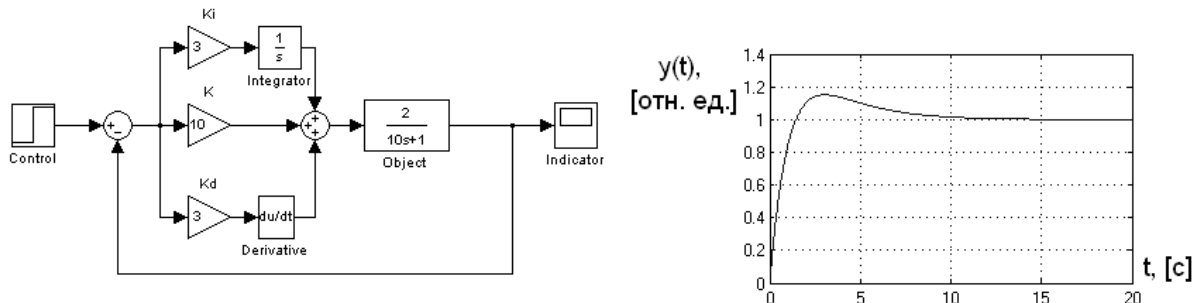


Рис. 4. Система управления с ПИД-регулятором

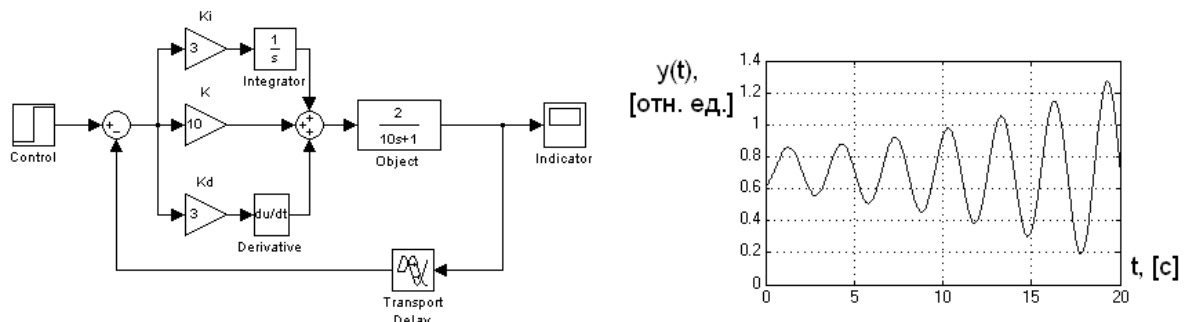


Рис. 5. Потеря устойчивости системы управления из-за запаздывания в измерении

Если канал измерения характеризуется запаздыванием, то становится затруднительным обеспечение требуемых показателей качества с помощью того же регулятора (рис. 5). При некотором критическом значении запаздывания система управления потеряет устойчивость [3]. Это объясняется тем, что звено запаздывания вносит отрицательный сдвиг фа-

зы в систему. Таким образом, необходимо искать новые подходы, позволяющие работать в условиях запаздывания как в канале управления, так и в канале измерения.

2.3. Общая постановка задачи

Рассматривается объект управления вида

$$\begin{cases} \dot{x}(t) = f(x(t), u(t-h_1)), \\ y(t) = r(x(t)), \end{cases}$$

где $x \in R^n$ – вектор состояния системы, $y \in R^1$ – выходная переменная, $u \in R^1$ – управляющее воздействие, $f(\cdot)$ – вектор-функция, $r(\cdot)$ – скалярная функция, h_1 – запаздывание в канале управления. Измеряемым является сигнал $y(t-h_2)$, либо весь вектор переменных состояния $x(t-h_2)$, где h_2 – запаздывание в канале измерения.

При предполагаемом законе управления

$$u(t+h_1) = v(\hat{y}(t+h_1), g(t+h_1), p),$$

где $v(\cdot)$ – скалярная функция, $\hat{y}(t+h_1)$ – оценка выходной переменной $y(t+h_1)$, $g(t+h_1)$ – задающее воздействие, $p = \frac{d}{dt}$ – оператор дифференцирования, требуется

построить предиктор, обеспечивающий предельное равенство $\lim_{t \rightarrow \infty} (y(t+h_1) - \hat{y}(t+h_1)) = 0$ или неравенство $\lim_{t \rightarrow \infty} (y(t+h_1) - \hat{y}(t+h_1)) \leq \varepsilon$, где ε – неко-

торое маленькое число. Может рассматриваться дополнительная цель, описываемая равенством $\lim_{t \rightarrow \infty} (x(t+h_1) - \hat{x}(t+h_1)) = 0$ или неравенством $\lim_{t \rightarrow \infty} (x(t+h_1) - \hat{x}(t+h_1)) \leq \varepsilon_{n \times 1}$, где

$\varepsilon_{n \times 1}$ – некоторый вектор маленьких чисел.

На рис. 6 показана структурная схема замкнутой системы управления с предиктором.

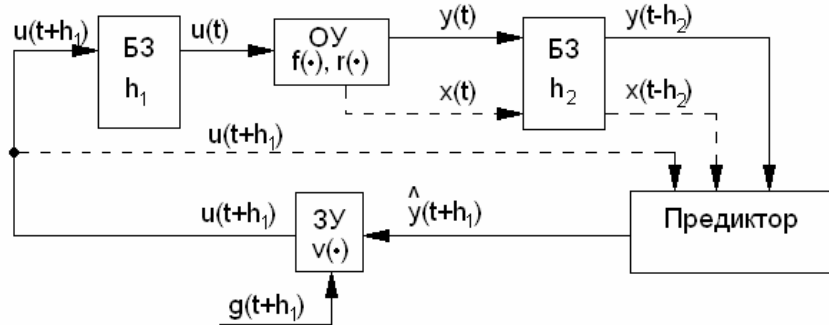


Рис. 6. Замкнутая система с предиктором

2.4. Адаптивный идентификационный подход

Рассматривается линейный стационарный объект управления, модель вход-состояние-выход которого представлена в каноническом наблюдаемом базисе:

$$\dot{x}(t) = A \cdot x(t) + b \cdot u(t-h_1), \quad y(t) = c^T x(t), \quad (1)$$

$$\text{где } A_{n \times n} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -a_1 \\ 1 & 0 & 0 & \cdots & 0 & -a_2 \\ 0 & 1 & 0 & \cdots & 0 & -a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -a_n \end{bmatrix} = \begin{bmatrix} 0_{1 \times n} & \\ I_{(n-1) \times (n-1)} & -a_{n \times 1} \end{bmatrix}, \quad b_{n \times 1} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \\ b_n \end{bmatrix},$$

$$c_{n \times 1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad x \in R^n \text{ – вектор переменных состояния, } y \in R^1 \text{ – выходная переменная,}$$

$u \in R^1$ – управляющее воздействие, h_1 – постоянное запаздывание в канале управления. Измеряемым является сигнал $y(t-h_2)$ и вектор переменных состояния $x(t-h_2)$, где h_2 – запаздывание в канале измерения. Значения h_1 и h_2 полагаются известными. Параметры $a_{n \times 1}$ и $b_{n \times 1}$ определены неточно или неизвестны. Закон управления имеет вид

$$u(t+h_1) = W(p) \cdot \hat{y}(t+h_1) + g(t+h_1),$$

где $W(p)$ – некоторая интегрально-дифференциальная передаточная функция, $g(t+h_1)$ – задающее воздействие.

Требуется построить предиктор, обеспечивающий цель

$$\lim_{t \rightarrow \infty} (y(t+h_1) - \hat{y}(t+h_1)) = 0, \quad \lim_{t \rightarrow \infty} (x(t+h_1) - \hat{x}(t+h_1)) = 0, \quad (2)$$

где $\hat{y}(t+h_1)$ – оценка выходной переменной $y(t+h_1)$, $\hat{x}(t+h_1)$ – вектор оценок переменных состояния $x(t+h_1)$. Предиктор описывается системой уравнений:

$$\begin{cases} \dot{\hat{x}}(t+h_1) = \hat{A}(t-h_2) \cdot \hat{x}(t+h_1) + \hat{b}(t-h_2) \cdot u(t+h_1), & \hat{y}(t+h_1) = c^T \hat{x}(t+h_1), \\ \dot{\hat{a}}(t-h_2) = -k_a \cdot (x(t-h_2) - \hat{x}(t-h_2)) \cdot \hat{y}(t-h_2), \\ \dot{\hat{b}}(t-h_2) = -k_b \cdot (x(t-h_2) - \hat{x}(t-h_2)) \cdot u(t-h_2), \\ \dot{\hat{x}}(t-h_2) = \hat{A}(t-h_2) \cdot \hat{x}(t-h_2) + \hat{b}(t-h_2) \cdot u(t-h_2), & \hat{y}(t-h_1) = c^T \hat{x}(t-h_1). \end{cases}, \quad (3)$$

$$\text{где } \hat{A}_{n \times n} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & -\hat{a}_1 \\ 1 & 0 & 0 & \cdots & 0 & -\hat{a}_2 \\ 0 & 1 & 0 & \cdots & 0 & -\hat{a}_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -\hat{a}_n \end{bmatrix} = \begin{bmatrix} 0_{1 \times n} & \\ & -\hat{a}_{n \times 1} \end{bmatrix}, \quad \hat{b}_{n \times 1} = \begin{bmatrix} \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ \vdots \\ \hat{b}_n \end{bmatrix}, \quad c_{n \times 1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}, \quad \text{где}$$

$\hat{y}(t+h_1)$ – оценка выходной переменной $y(t+h_1)$, $\hat{x}(t+h_1)$ – вектор оценок переменных состояния $x(t+h_1)$, $\hat{y}(t-h_2)$ – оценка выходной переменной $y(t-h_2)$, $\hat{x}(t-h_2)$ – вектор оценок переменных состояния $x(t-h_2)$, $\hat{a}(t-h_2)$ – вектор оценок параметров a , $\hat{b}(t-h_2)$ – вектор оценок параметров b , $k_a > 0$ и $k_b > 0$ – некоторые положительные числа. Алгоритм настройки параметров аналогичен базовому [4]. Предиктор (3) обеспечивает цель (2). Если сигнал $u(t)$ – «частотно богатый», то предиктор обеспечит дополнительную цель

$$\lim_{t \rightarrow \infty} (a - \hat{a}(t-h_2)) = 0, \quad \lim_{t \rightarrow \infty} (b - \hat{b}(t-h_2)) = 0.$$

Пример

Пусть параметры объекта управления (1) имеют вид

$$A_{2 \times 2} = \begin{bmatrix} 0 & -6 \\ 1 & -3 \end{bmatrix}, \quad b = \begin{bmatrix} 1 \\ 2 \end{bmatrix}, \quad c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad h_1 = 3, \quad h_2 = 2.$$

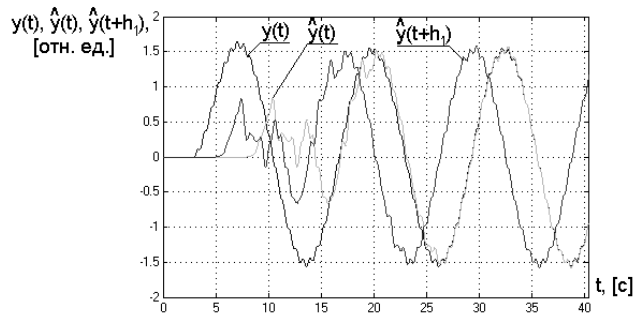


Рис. 7. Временные диаграммы сигналов $y(t)$, $\hat{y}(t)$, $\hat{y}(t+h_1)$

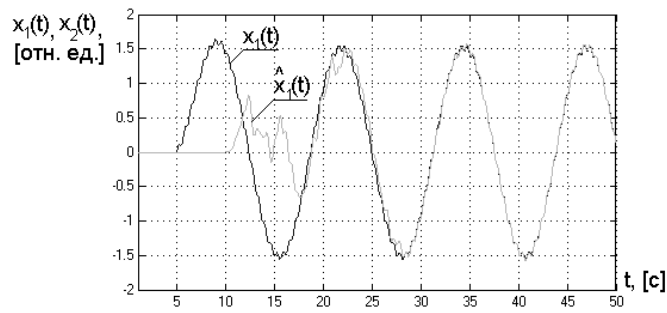


Рис. 8. Временные диаграммы сигналов $x_1(t)$, $\hat{x}_1(t)$

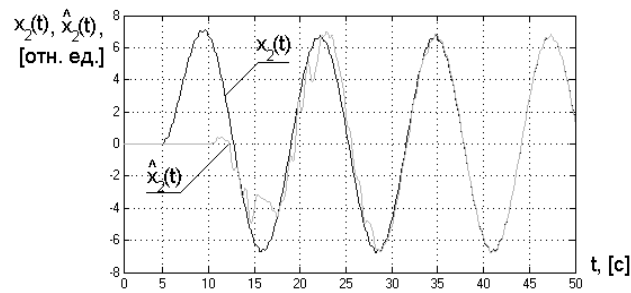


Рис. 9. Временные диаграммы сигналов $x_2(t)$, $\hat{x}_2(t)$

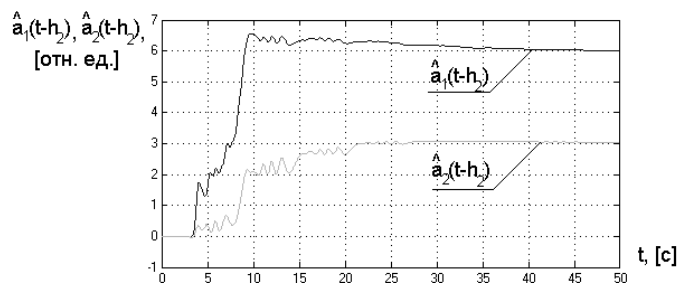


Рис. 10. Временные диаграммы оценок параметров $\hat{a}_1(t-h_2)$, $\hat{a}_2(t-h_2)$

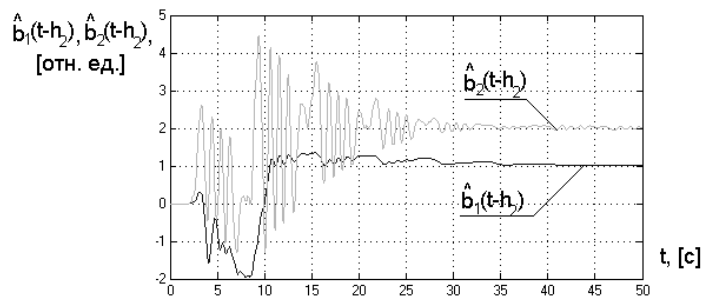


Рис. 11. Временные диаграммы оценок параметров $\hat{b}_1(t-h_2)$, $\hat{b}_2(t-h_2)$

На рис. 7–11 представлены результаты моделирования предиктора (3), из которых видно, что цель (2) достигнута.

2.5. Другие методы

Если вектор переменных состояния не измеряется, то схема предиктора, а именно алгоритм адаптации, усложняется. Тем не менее, существует адаптивный подход, основанный на параметризации объекта управления и ряда линейных фильтров, обеспечивающих адекватную настройку параметров и получение упреждающей оценки выходной переменной.

Если сигнал является периодическим, то упреждающую оценку можно легко получить, пропустив сигнал через звено запаздывания, величина которого вычисляется в соответствии с периодом сигнала. Этот подход можно использовать при работе с квазигармоническими сигналами. Внося в каждую гармонику соответствующее запаздывание, нетрудно получить требуемую упреждающую оценку.

Предиктор можно построить по неадаптивной схеме. Легко показать, что звено чистого запаздывания имеет передаточную функцию вида $W(p) = e^{-hp}$, где h – величина запаздывания. Следовательно, звено с передаточной функцией $\bar{W}(p) = e^{hp}$, по сути, является звеном «чистого предсказания». Реализовать такое звено весьма затруднительно, но, разложив экспоненту в ряд Маклорена, можно получить с некоторой точностью линейный фильтр. Нетрудно видеть, что появится новая проблема с реализацией звена дифференцирования.

Существует метод замены координат [5]. В этом подходе в новой системе координат формально пропадает звено запаздывания.

3. Заключение

В работе исследовалась проблема управления в условиях запаздывания. Были продемонстрированы примеры технических систем с запаздыванием. Были показаны недостатки классических методов. Решалась задача получения упреждающей оценки измеряемых сигналов.

Насматривается линейный стационарный объект управления, имеющий запаздывание в каналах управления и измерения, решается задача синтеза адаптивного предиктора для нахождения упреждающей оценки регулируемой переменной и вектора переменных состояния в условиях параметрической неопределенности. Показан пример использования данного подхода, проиллюстрированы результаты моделирования, подтверждающие работоспособность алгоритма.

Литература

1. Гурецкий Х. Анализ и синтез систем управления с запаздыванием. М.: Машиностроение, 1973. 328 с.
2. http://www.erudition.ru/referat/ref/id.35532_1.html
3. Бесекерский В.А., Попов Е.П. Теория систем автоматического регулирования. М.: Наука, 1975. 768 с.
4. Мирошник И.В., Никифоров В.О., Фрадков А.Л. Нелинейное и адаптивное управление сложными динамическими системами. СПб: Наука, 2000. 549 с.
5. Arstein Z. Linear systems with delayed controls. // IEEE Transactions on Automatic control. 1982. V. AC–27. №3. P. 869–879.

ИСПОЛЬЗОВАНИЕ ПОЛИНОМОВ КОТЕЛЬНИКОВА В ЗАДАЧЕ АППРОКСИМАЦИИ МОДЕЛИ ТЕРМОДИНАМИЧЕСКОГО ОБЪЕКТА

А.А. Блинников

Научный руководитель – к.т.н., доцент В.И. Бойков

Произведен анализ вычислительной ошибки, порождаемой принятой схемой моделирования процессов, протекающих в термодинамическом объекте. Предложен способ выбора шага пространственной сетки для конечномерной аппроксимации исходного уравнения в частных производных, основанный на использовании теоремы Котельникова.

Для ряда задач, сутью которых является моделирование термодинамических процессов в объемных объектах, имеется возможность существенного упрощения вычислительной схемы. За счет анизотропии процессов можно выделить плоскости решения уравнения теплопроводности, сведя трехмерную задачу к совокупности двумерных. Частным случаем такой задачи является моделирование термодинамического процесса в машине непрерывного литья заготовок (МНЛЗ). Схематическое изображение машины непрерывного литья заготовок приведено на рис. 1.

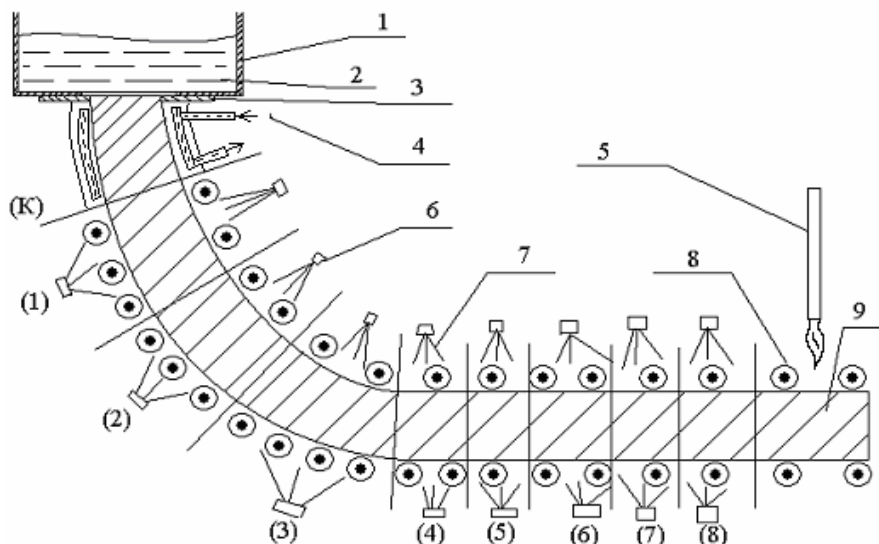


Рис. 1. Схема МНЛЗ: 1 – пром. ковш; 2 – жидкая сталь; 3 – пластины формователи; 4 – нагнетатель воды в кристаллизаторе (К); 5 – система нарезки заготовок; 6 – форсунка подачи воды в зонах вод.-возд. охлаждения; 7 – охладитель (вода); 8 – прокатные валки; 9 – сляб

Моделирование процессов охлаждения заготовки позволяет по известному температурному полю восстановить ряд параметров, таких как:

- толщина корки – слоя твердого металла, окружающего жидкую фазу;
- величина термических напряжений в корке;
- протяженность жидкой фазы вдоль заготовки;
- термическая деформация сортовой заготовки.

Эти параметры необходимы для качественного контроля производства заготовок со стороны технологов и операторов МНЛЗ.

Существует много моделей [1, 3], решающих эту задачу и основанных на дискретизации уравнения теплопроводности. Однако выбор параметров дискретизации формализован весьма слабо. В статье рассмотрен один из способов выбора шага равномерной пространственной сетки, основанный на использовании теоремы Котельникова.

Постановка задачи

Если пренебречь влиянием валков и изгибом сляба, то объект, подобный показанному на рис. 1, представляет собой параллелепипед с диффузно-конвективным типом теплообмена, который описывается трехмерным уравнением в частных производных с граничными условиями первого, второго и третьего рода [2, 3]. Поскольку объект, описанный в терминах уравнений в частных производных, является бесконечномерным, для практического применения модели необходимо произвести аппроксимацию исходной модели конечномерной в одном из принятых видов, а именно:

- метод конечных элементов (МКЭ);
- метод конечных разностей (МКР);
- модальный метод (ММ).

Наиболее простым и, в то же время, применимым на практике из перечисленных способов аппроксимации является МКР, однако выбор шага пространственной сетки в этом методе формализован весьма слабо. Традиционно в качестве критерия выбора шага пространственной сетки рассматривают:

- соблюдение условий устойчивости принятой разностной схемы [3];
- пространственную близость узлов сетки к границе области задачи, с целью корректного учета граничных условий.

Ввиду простоты пространственной конфигурации области задачи второй критерий может дать лишь расположение граничных узлов, в то время как первый работает только для условно устойчивых схем, да и для них дает лишь размер шага, лежащий на точке бифуркации свойств объекта. Целью данной статьи является получение хорошо формализованного алгоритма выбора минимального количества узлов пространственной сетки, обеспечивающего заданную погрешность решения уравнения теплопроводности по всему пространству задачи.

Математические предпосылки

Означенный класс объектов может быть описан в виде (1) [2, 3].

$$\left\{ \begin{array}{l} \frac{dQ(x, y, z, t)}{dt} = a^2 \left(\frac{d^2 Q(x, y, z, t)}{dx^2} + \frac{d^2 Q(x, y, z, t)}{dy^2} + \right. \\ \left. + \frac{d^2 Q(x, y, z, t)}{dz^2} \right) - V \frac{dQ(x, y, z, t)}{dz} \\ \frac{dQ(x, y, z, t)}{dx} \Big|_{(x=x_0)} = \frac{h_{x0}}{k} (Q(x_0, y, z, t) - T_g(t)); \\ \frac{dQ(x, y, z, t)}{dx} \Big|_{(x=x_1)} = -\frac{h_{x1}}{k} (Q(x_1, y, z, t) - T_g(t)); \\ \frac{dQ(x, y, z, t)}{dy} \Big|_{(y=y_0)} = \frac{h_{y0}}{k} (Q(x, y_0, z, t) - T_g(t)); \\ \frac{dQ(x, y, z, t)}{dy} \Big|_{(y=y_1)} = -\frac{h_{y1}}{k} (Q(x, y_1, z, t) - T_g(t)); \\ \frac{dQ(x, y, z, t)}{dz} \Big|_{(z=z_0)} = T_{Ж}, \frac{dQ(x, y, z, t)}{dz} \Big|_{(z=z_1)} = 0 \end{array} \right. , \quad (1)$$

где $a = \sqrt{\frac{k}{c\rho}}$, причем k , Вт/(м*К) – коэффициент теплопроводности; c , Дж/(кг*К) – удельная теплоемкость; ρ , г/м³ – плотность. h_{x0} , h_{x1} , h_{y0} , h_{y1} , Вт/(м²*К) – коэффициенты теплопередачи вдоль соответствующей оси на границах области \bar{D} (0 – соответствует 0 координаты, 1 – максимальному значению координаты (высота или ширина заготовки)). Коэффициенты отображают эффективность теплообмена и определяются экспериментально [4]. \bar{V} , м/с – скорость продольного движения сляба (вдоль оси Z); $Q(x, y, z, t)$, C – динамическое температурное поле; $T_g(t)$, C – температура воды (охлаждителя); $T_{ж}$, C – температура жидкой стали;

Оси координат располагаются следующим образом: центр системы координат – нижняя угловая точка в левом, на рис. 1, торце; ось Z направлена вдоль продвижения сляба; ось X – направлена в поперечном сечении сляба в горизонтальной плоскости, а Y – в вертикальной.

Учитывая то, что теплообмен в направлении Z носит преимущественно конвективный характер, а вдоль осей X, Y – диффузный, а также симметричность охлаждения относительно плоскости YZ, модель (1) часто упрощают до (2). Выражение (2) описывает поведение i -го сечения, движущегося вместе с заготовкой со скоростью V вдоль оси Z. В начальный момент времени i -ое сечение равномерно прогрето.

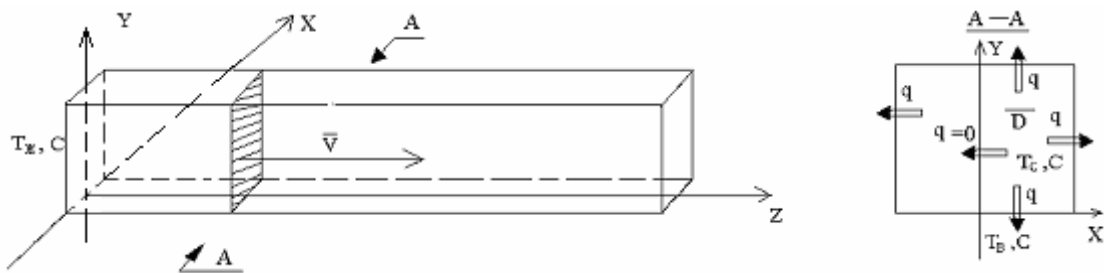


Рис. 2. Редукция модели сляба в 2 измерения

$$\left\{ \begin{array}{l} \frac{dQ_i(x, y, t_i)}{dt} = a^2 \left(\frac{d^2 Q_i(x, y, t_i)}{dx^2} + \frac{d^2 Q_i(x, y, t_i)}{dy^2} \right) \\ \frac{dQ_i(x, y, t_i)}{dx} \Big|_{(x=x_0)} = 0; \\ \frac{dQ_i(x, y, t_i)}{dx} \Big|_{(x=x_1)} = -\frac{h_{x1}}{k} (Q_i(x_1, y, t_i) - T_g(k)); \\ \frac{dQ_i(x, y, t_i)}{dy} \Big|_{(y=y_0)} = \frac{h_{y0}}{k} (Q_i(x, y_0, t_i) - T_g(k)); \\ \frac{dQ_i(x, y, t_i)}{dy} \Big|_{(y=y_1)} = -\frac{h_{y1}}{k} (Q_i(x, y_1, t_i) - T_g(k)); \\ Q_i(x, y, 0) = 0 \end{array} \right. , \quad (2)$$

где $Q_i(x, y, t_i) = Q(x, y, z = V * t_i, t_i)$.

Модель (2), также как и (1), является бесконечномерной. Для моделирования на ЭВМ уравнение (2) обычно представляется в дискретном виде (в нашем случае – МКР), для чего необходимо найти параметры дискретизации. Решение поставленной задачи будем искать на основе теоремы Котельникова.

Исследование динамики ошибки интерполяции значения температурного поля между узлами сетки

В соответствии с теоремой Котельникова [5], некоторая функция $S(t)$, удовлетворяющая условиям Дирихле (ограничена, кусочно-непрерывная, имеет конечное число экстремумов), с ограниченным по ширине спектром полностью определяется своими мгновенными значениями, отсчитанными через интервалы переменной t (3)

$$\Delta t = \frac{\pi}{\omega_c} = \frac{1}{2f_c}, \quad (3)$$

где f_c – граничная частота спектра функции $S(t)$. Согласно теореме Котельникова, такая функция $S(t)$, ограниченная по частоте, а значит, неограниченная по шкале аргумента, может быть восстановлена по отсчетам t_k . Для восстановления функции используется полином Котельникова (4)

$$S(t) = \sum_{-\infty}^{+\infty} S(t_k) \frac{\sin(\omega_c(t-t_k))}{\omega_c(t-t_k)}. \quad (4)$$

Реальные сигналы имеют неограниченный спектр, поэтому для них теорема Котельникова, строго говоря, не применима. С другой стороны, амплитуда гармоник $S(t)$ обычно резко затухает с ростом частоты, и на практике существует некоторая частота, которую можно принять за f_c . Тогда полином (4) по значениям $S(t)$ через шаг (3) дает аппроксимацию $S^*(t)$. Средняя квадратичная ошибка $\sigma_{отн}$ (СКО) такой аппроксимации определяется как

$$\sigma_{отн}^2 = \frac{\int_0^T (S(t) - S^*(t))^2 dt}{\int_0^T S^2(t) dt}. \quad (5)$$

Как показано в [5], интервал дискретности, дающий относительное СКО $\sigma_{отн}$, может быть найден следующим образом:

- (1) рассчитывается мощность среза как $P_{срез} = P_{сиг} * (1 - \sigma_{отн}^2/3)$, где $P_{сиг}$ – мощность всего сигнала;
- (2) последовательно увеличивая количество учтенных гармоник Ng , определяется значение, когда их суммарная мощность превысит $P_{срез}$;

- (3) рассчитывается интервал дискретности как $\Delta t = \left\lceil \frac{T}{\frac{T\omega_{срез}}{\pi}} \right\rceil$, где T – область задачи, а

$\omega_{срез}$ – частота среза, т.е. частота, для которой совокупность гармоник с частотами $\leq \omega_{срез}$ имеет мощность $\geq P_{срез}$.

Для применения теоремы Котельникова было построено представление модели (2) в виде гармонического ряда. С этой целью была построена ММ в виде (6)–(8), синтез которой хорошо формализован [3, 6]. С учетом конфигурации охлаждения в МНЛЗ параметры ММ рассчитаны по (9) – (17).

$$Q_{nm}(\mu_n, \eta_m, t) = \int_{x_0}^{x_1} \int_{y_0}^{y_1} Q(x, y, t) \varphi_n(\mu_n, x) \varphi_m(\eta_m, y) r(x) r(y) dy dx, \quad (6)$$

где $r(x) = r(y)$ – весовая функция преобразования.

$$Q(x, y, t) = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} Q_{nm}(\mu_n, \eta_m, t) \varphi_n(\mu_n, x) \varphi_m(\eta_m, y) \quad (7)$$

$$\begin{cases} Q_{nm}(\mu_n, \eta_m, t) = A Q_{nm}(\mu_n, \eta_m, t) + B T_e \\ y(k) = C Q_{nm}(\mu_n, \eta_m, t) \end{cases}, \quad (8)$$

где

$$A = \begin{bmatrix} -\mu_1^2 - \eta_1^2 & 0 & 0 & 0 \\ 0 & -\mu_1^2 - \eta_2^2 & 0 & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & -\mu_n^2 - \eta_m^2 \end{bmatrix};$$

$$B = \begin{bmatrix} Rux(\mu_1, \eta_1, t) + Ruy(\mu_1, \eta_1, t) \\ Rux(\mu_1, \eta_2, t) + Ruy(\mu_1, \eta_2, t) \\ \dots \\ Rux(\mu_n, \eta_m, t) + Ruy(\mu_n, \eta_m, t) \end{bmatrix};$$

$$C = [\varphi_1(\mu_1, x_\partial) \varphi_1(\eta_1, y_\partial) \quad \varphi_1(\mu_1, x_\partial) \varphi_2(\eta_2, y_\partial) \quad \dots \quad \varphi_n(\mu_n, x_\partial) \varphi_m(\eta_m, y_\partial)],$$

причем $Rux(\mu_n, \eta_m, t)$ и $Ruy(\mu_n, \eta_m, t)$ – пространственное изображение граничного входа по одной оси вдоль другой.

$$r(x) = r(y) = \frac{1}{\alpha^2}; \quad (9)$$

$$\varphi_n(\mu_n, x) = \frac{1}{E_n} \varphi_n^*(\mu_n, x); \quad (10)$$

$$\varphi_n^*(\mu_n, x) = \cos\left(\frac{\mu_n}{\alpha} x\right); \quad \varphi_m^*(\eta_m, y) = k \cos\left(\frac{\eta_m}{\alpha} y\right) + hy_0 \frac{\alpha}{\eta_m} \sin\left(\frac{\eta_m}{\alpha} y\right); \quad (11)$$

$$E_n^2 = \frac{1}{2} \left(\frac{1}{\alpha \mu_n} \sin\left(\frac{\mu_n}{\alpha} x_1\right) \cos\left(\frac{\mu_n}{\alpha} x_1\right) + \frac{x_1}{\alpha^2} \right); \quad (12)$$

$$E_m^2 = \left(k^2 - \left(\frac{\alpha hy_0}{\eta_m} \right)^2 \right) \left(\frac{1}{2 \alpha \mu_n} \sin\left(\frac{\eta_m}{\alpha} y_1\right) \cos\left(\frac{\eta_m}{\alpha} y_1\right) \right) + \left(k^2 + \left(\frac{\alpha hy_0}{\eta_m} \right)^2 \right) \left(\frac{y_1}{2 \alpha^2} \right) + \frac{k hy_0}{\eta_m^2} \sin^2\left(\frac{\eta_m}{\alpha} y_1\right); \quad (13)$$

$$Q_{nm}(\mu_n, \eta_m, 0) = \frac{T_e}{E_n E_m \alpha^2 \mu_n \eta_m} * \quad (14)$$

$$* \left(\sin\left(\frac{\mu_n}{\alpha} x_1\right) \left(k \sin\left(\frac{\eta_m}{\alpha} y_1\right) + \frac{\alpha hy_0}{\eta_m} (1 - \cos\left(\frac{\eta_m}{\alpha} y_1\right)) \right) \right)$$

$$tg\left(\frac{\mu_n}{\alpha} x_1\right) = \frac{\mu_n h x_1}{k \alpha}; \quad tg\left(\frac{\eta_m}{\alpha} y_1\right) = \frac{k (hy_0 + h y_1) \frac{\eta_m}{\alpha}}{k^2 \frac{\eta_m^2}{\alpha^2} - hy_0 h y_1}; \quad (15)$$

$$Rux(\mu_n, \eta_m, t) = \frac{h x_1}{k E_n} \cos\left(\frac{\mu_n}{\alpha} x_1\right) * \quad (16)$$

$$* \left(\frac{k}{\alpha \eta_m E_m} \sin\left(\frac{\eta_m}{\alpha} y_1\right) + \frac{hy_0}{\eta_m^2 E_m} (1 - \cos\left(\frac{\eta_m}{\alpha} y_1\right)) \right)$$

$$R_{uy}(\mu_n, \eta_m, t) = \frac{1}{2\mu_n E_n} \sin\left(\frac{\mu_n}{\alpha} x_1\right) * \quad (17)$$

$$* \left(\frac{hy_1}{kE_m} \left(k \cos\left(\frac{\eta_m}{\alpha} y_1\right) + \frac{\alpha hy_0}{\eta_m} \sin\left(\frac{\eta_m}{\alpha} y_1\right) \right) - \frac{hy_0}{E_m} \right)$$

В начальный момент времени спектральный состав сигнала определяется начальными условиями (14). Для объекта (2) вид частотного распределения на начальный момент времени (равномерный прогрев) приведен на рис. 3.

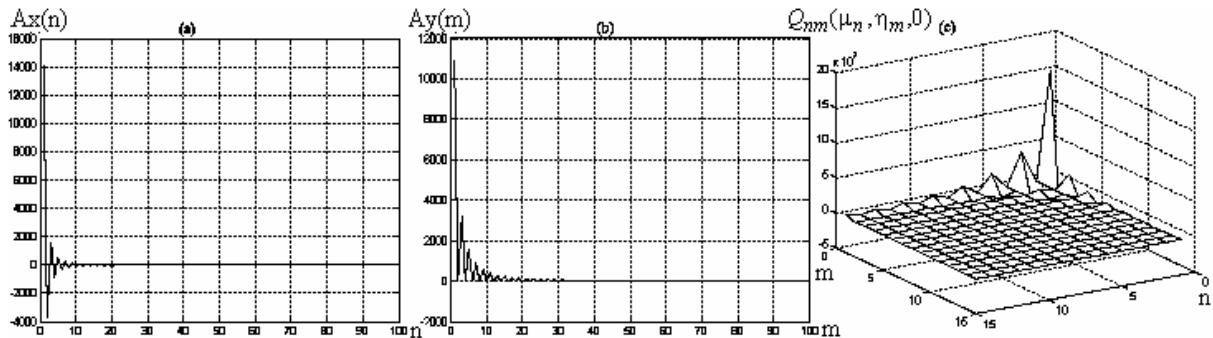


Рис. 3. Распределение коэффициентов гармонического разложения задачи (2) в подпространствах X(a) и Y(b) и в плоскости XY(c) в начальный момент времени

Моделирование представления (2) в виде (8), с параметрами (9) – (17) дало частотное распределение температурного поля на момент времени $T=400$ сек (рис. 4).

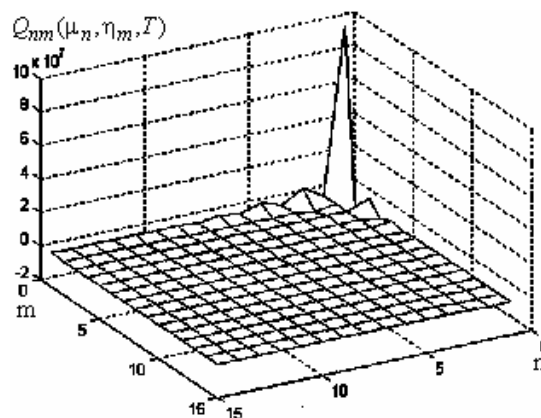


Рис. 4. Распределение коэффициентов гармонического разложения задачи (2) в плоскости XY(c) в момент времени $T=400$ с

Как видно, спектр частотного представления модели (2) со временем сужается, а значит, $\sigma_{отн}^2$ в начальный момент времени, когда заготовка нагрета равномерно, является наибольшим. Исходя из вышесказанного, был сделан вывод: шаг сетки, рассчитанный по начальным условиям (14) и требуемой $\sigma_{отн}^2$, дает аппроксимацию поля с точностью не худшей, чем $\sigma_{отн}^2$, на всем протяжении расчета сечения.

Применение полиномов Котельникова к двумерной задаче

Поскольку теорема Котельникова традиционно применяется к одномерным сигналам, при использовании ее в двумерной задаче необходимо было произвести ряд модификаций. Во-первых, выражение дающее $P_{срез}$ было приведено к форме (18), учитывающей наличие двух измерений. Во-вторых, (4) преобразовалось в (19).

$$P_{\text{срез}} = P_{\text{сиг}} * (1 - \sigma_{\text{отн}}^2/9), \quad (18)$$

$$S(t) = \sum_{m=1}^M \sum_{n=1}^N S(x_m, y_n) \frac{\sin(\omega_c(x-x_m))}{\omega_c(x-x_m)} \frac{\sin(\omega_c(y-y_n))}{\omega_c(y-y_n)}, \quad (19)$$

где M, N – количество узлов сетки вдоль осей.

В-третьих, процесс нахождения количества учтенных гармоник, обеспечивающего заданное $P_{\text{срез}}$, видоизменился. Ввиду того, что имеется два пространственных измерения, количество учтенных гармоник вдоль каждой оси Mg и Ng не может быть определено однозначно. Числа Mg и Ng для некоторого $\sigma_{\text{отн}}$ лежат на кривой, получаемой последовательным увеличением числа гармоник по одной координате при фиксированном для данной точки кривой по другой. Эта кривая в нашей задаче для $\sigma_{\text{отн}} = 1\%$ выглядит, как это показано на рис. 5.

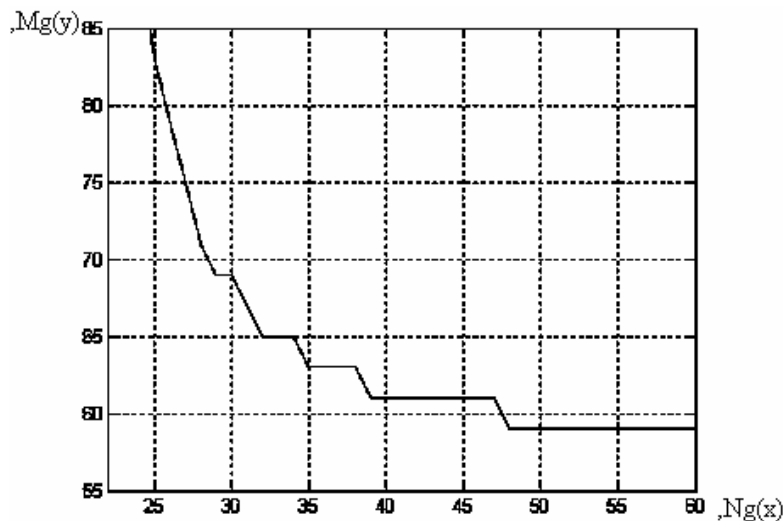


Рис. 5. Кривая равных мощностей сигнала для Ng учтенных гармоник по x и Mg – по y для $\sigma_{\text{отн}} = 1\%$

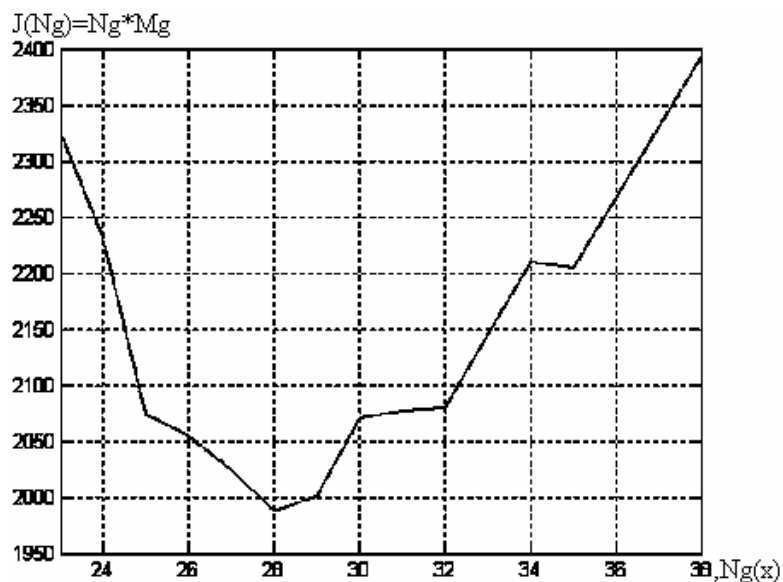


Рис. 6. График $J(Ng)$

Логично предположить, что выбор точки на кривой, приведенной на рис. 5, должен обеспечивать наименьшее количество узлов сетки. С этой целью был построен

график функции (20), приведенный на рис. 6, по сути, дающий общее количество гармоник.

$$J(Ng) = Ng * Mg, P(Ng, Mg) = \text{const.} \quad (20)$$

Из рис. 6 видно, что наименьшее количество гармоник, а значит, и узлов сетки, удовлетворяющее требованиям, соответствует минимуму функции (20), для Ng – количества гармоник по x . По y же число Mg находится по рис. 5, исходя из найденного Ng . Так, для $\sigma_{отн} = 1\%$ $Ng=28$, а $Mg=71$.

Таким образом, алгоритм расчета шага сетки изменился и приобрел следующий вид.

- (1) Назначить относительное СКО $\sigma_{отн}$.
- (2) Построить ММ задачи.
- (3) Рассчитать $P_{срез}$ по (18).
- (4) Последовательно увеличивая количество учтенных гармоник Ng и Mg , найти кривую, переход которой дает мощность сигнала $\geq P_{срез}$.
- (5) Построить графики, аналогичные приведенным на рис. 5 и рис. 6, определив по второму из них количество гармоник Ng вдоль одной оси, а по первому – Mg вдоль другой, основываясь на Ng .

(6) Рассчитать интервалы дискретности по теореме Котельникова и ММ как $\Delta x = \frac{X}{\left\lceil \frac{X\mu_{Ng}}{\pi a} \right\rceil}$ и $\Delta y = \frac{Y}{\left\lceil \frac{Y\eta_{Mg}}{\pi a} \right\rceil}$, где X, Y – область задачи вдоль осей, μ_{Ng} и η_{Mg} – собственные числа задачи (2), определенные в ММ и порождающие гармоники с частотами большими, нежели частоты среза соответствующих осей, а a – параметр из (2).

Моделирование

В соответствии с полученным алгоритмом для $\sigma_{отн} = 1\%$ сетка имеет 29×72 узла. Для начального момента времени, когда заготовка прогрета равномерно, ее сечение имеет температурное поле, приведенное на рис. 7; как было выяснено ранее, это наиболее требовательный вариант в смысле количества узлов сетки. Поле, восстановленное по узлам сетки, имеет вид, приведенный на рис. 8.

Моделирование показало, что СКО исходного температурного поля и восстановленного составляет $\sigma_{отн} = 1,04\%$, что практически совпадает с расчетным значением. В то же время из графиков рис. 7 и рис. 8 можно увидеть, что локальная ошибка восстановления решения может быть весьма значительна. В рассмотренном случае она достигает 23 %, что является недостатком способа, однако он дает довольно точную методику расчета шага пространственной сетки по СКО.

В итоге был получен способ расчета шага пространственной сетки, использующий теорему Котельникова. Показано, что выбор шага сетки для рассмотренной задачи можно произвести по начальным условиям. Произведено моделирование задачи. Показано, что рациональный выбор шага сетки позволяет обеспечить заданную среднеквадратическую ошибку восстановления решения уравнения. Выявлено, что использование восстанавливающих полиномов Котельникова дает значительные локальные выбросы ошибки даже при существенном уменьшении шага аппроксимирующей модели.

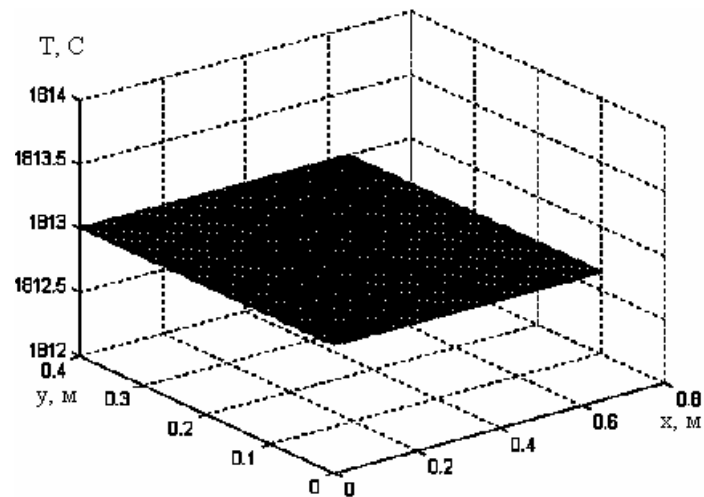


Рис. 7. Температурное поле в начальный момент времени

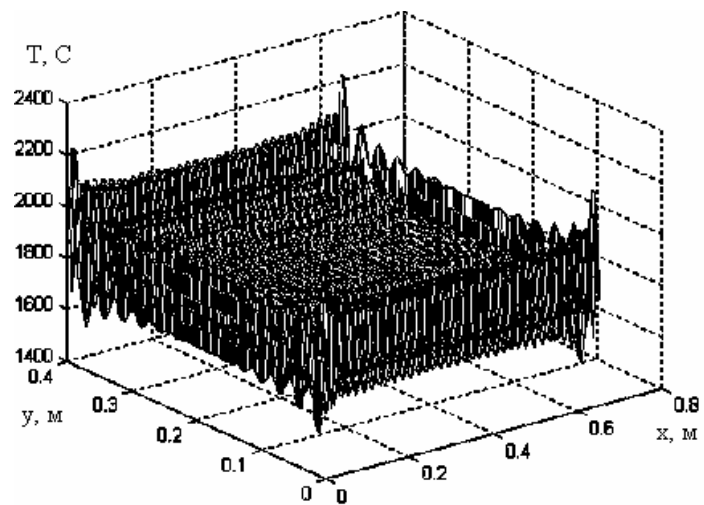


Рис. 8. Температурное поле в начальный момент времени, восстановленное по сетке 29*72, с помощью (17)

Литература

1. Левин Р.Г., Галкин М.П., Зубрев О.И., Сургаева Е.В., Глеков Л.К. Математическое моделирование процесса кристаллизации непрерывнолитой заготовки // *Металлург*, 2003. С. 42–44.
2. Фарлоу С.Г. Уравнения с частными производными для научных работников и инженеров [Текст]: М., Мир, 1985. 384 с.
3. Рапопорт Э.Я. Структурное моделирование объектов и систем управления с распределенными параметрами [Текст]: учеб. пособие/ М.: Высшая школа, 2003. 299 с.
4. Волков А.В. Промышленное освоение математической модели в составе комплексной АСУТП МНЛЗ-2 Молдавского металлургического завода. // Доклад на научно-технической конференции «Прогрессивные технологии непрерывной разливки: XXI век». 28-31 мая 2002, г. Донецк. С.В. // *Специальная техника*. 1999. №6. С. 24–29. [Ссылка в сети Интернет] http://www.oracul.org/publications/acsep/r01_02_03.pdf
5. Кавчук С.В. Руководство к лабораторной работе №2 «Временная дискретизация сигналов» по курсу Теоретические основы информационно-измерительной техники. [Текст]: учеб. пособие. Таганрог: изд. ТРТУ, 1999. 38 с.
6. Рапопорт Э.Я. Анализ и синтез систем автоматического управления с распределенными параметрами [Текст]: учеб. пособие. М.: Высшая школа, 2005. 292 с.

СИНТЕЗ АДАПТИВНОГО НАБЛЮДАТЕЛЯ ДЛЯ ХАОТИЧЕСКОЙ СИСТЕМЫ ДУФФИНГА

С.В. Арановский, Н.А. Николаев

Научный руководитель – д.т.н., профессор А.А. Бобцов

С использованием процедуры синтеза адаптивного наблюдателя для хаотической системы Дуффинга решается задача восстановления неизвестного закодированного параметра. Данная задача, в отличие от известных аналогов, решается только по измерениям выходного сигнала хаотической системы, а также в условиях полной параметрической неопределенности.

1. Введение

Проблема построения адаптивных наблюдателей для нелинейных динамических систем находится в центре внимания в течение последних лет. Одна из причин подобного интереса состоит в возможности применения адаптивных наблюдателей для кодирования и передачи информации. Это одно из новых направлений передачи данных, когда информация кодируется в параметрах динамической системы «передатчика». Сигнал выхода этой системы передается «приемнику», в задачу которого входит восстановление неизмеряемых сигналов «передатчика» и параметров модели «передатчика». Структурная схема подобной системы приведена на рис. 1, где θ – вектор параметров модели системы «передатчика», кодирующий передаваемую информацию; y – передаваемый по каналу связи выход «передатчика»; $\hat{\theta}$ – оценка вектора θ , вырабатываемая приемником.

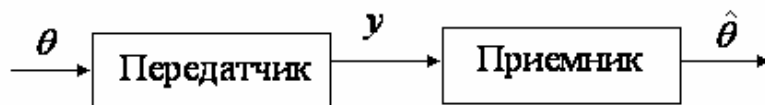


Рис. 1. Структурная схема системы передачи данных

Особенно перспективным представляется использование хаотических динамических систем в качестве моделей «передатчиков», так как сигнал выхода хаотической системы, с одной стороны, имеет широкий частотный диапазон, а с другой стороны, решения подобных систем демонстрируют слабую зависимость от начальных условий, что повышает защищенность системы от несанкционированного восстановления информационной составляющей сигнала. В этом случае в задачу «приемника» входит построение адаптивного наблюдателя хаотической системы [1].

Для синтеза адаптивных наблюдателей обычно используется несколько групп методов [2, 3], большинство из которых основываются на возможности пассивации обратной связью модели системы «передатчика» в предположении, что эта модель имеет нулевую или единичную относительную степень. Другие решения предполагают доступность для измерения всего вектора состояния системы «передатчика» [4, 5]. В статье [6] было предложено решение, позволяющее синтезировать адаптивные наблюдатели для моделей «передатчика», имеющих высокую (выше единичной) относительную степень и не допускающих пассивацию обратной связью по выходу. Эти результаты основаны на использовании новой канонической формы нелинейных адаптивных наблюдателей, предложенной в [6].

Результат работы [6] был усилен в [7–9], где рассматривалась задача построения адаптивных наблюдателей по части переменных для неавтономных нелинейных динамических систем. Использование внешнего возбуждающего сигнала является одним из подходов к созданию хаотических режимов движения в нелинейных системах. Приме-

рами таких систем являются модель Дуффинга и модель бруссельатора [9, 10], демонстрирующие хаотическое поведение только в присутствии соответствующего гармонического возмущения. Распространение классических результатов на задачу построения адаптивных наблюдателей для неавтономных систем, во-первых, позволяет существенно расширить класс допустимых моделей для системы «передатчика», а, во-вторых, увеличивает защищенность системы от несанкционированного доступа.

В данной статье рассматривается задача синтеза адаптивного наблюдателя для хаотических сигналов, генерируемых хаотической системой Дуффинга. Суть этой задачи связана с выделением из хаотического сигнала полезной информации, передаваемой по каналу связи. В отличие от известных результатов (см., например, [9]), в данной статье будет рассмотрена задача синтеза наблюдателя только по измерению выходной переменной хаотического сигнала при полной параметрической неопределенности его модели.

2. Постановка задачи

Рассмотрим хаотическую систему Дуффинга, описываемую уравнением вида (см., например, [9])

$$\ddot{y}(t) + c_1 \dot{y}(t) + c_2 y(t) - \bar{\theta} f(y) - w(t) = 0, \quad (1)$$

где c_1 , c_2 и $\bar{\theta}$ – неизвестные числа, нелинейная функция $f(y) = y^3$, $w(t) = A \sin(\omega t + \phi)$ – неизмеримый гармонический сигнал.

Ставится задача синтеза наблюдателя, обеспечивающего восстановление неизвестного параметра $\bar{\theta}$ модели (1) (см., например, [9]).

Будем допускать, что измеряется только выходная переменная $y(t)$ модели (1). Также предполагается, что параметры хаотической системы (1) c_1 , c_2 , $\bar{\theta}$, A , ω и ϕ являются неизвестными числами.

3. Синтез адаптивного наблюдателя

Для вывода основного результата перепишем модель (1) следующим образом:

$$y(t) = \frac{1}{a(p)} \left[\bar{\theta} y^3 + w(t) \right], \quad (2)$$

где полином $a(p) = p^2 + c_1 p + c_2$ и $p = d/dt$.

Переходя к изображениям Лапласа, для уравнения (2) имеем

$$Y(s) = \frac{\bar{\theta} F(s)}{s^2 + c_1 s + c_2} + \frac{W(s)}{s^2 + c_1 s + c_2} + \frac{D(s)}{s^2 + c_1 s + c_2}, \quad (3)$$

где s – комплексная переменная, $Y(s) = L\{y(t)\}$, $F(s) = L\{f(y(t))\}$, $W(s) = L\{w(t)\}$ – образы Лапласа, соответственно, функций $y(t)$, $f(y(t))$ и $w(t)$, полином $D(s)$ обозначает сумму всех членов, содержащих ненулевые начальные условия.

Преобразуем модель (3) следующим образом:

$$Y(s) = \frac{\bar{\theta} F(s)}{s^2 + c_1 s + c_2} + \frac{W(s)}{s^2 + c_1 s + c_2} + \frac{D(s)}{s^2 + c_1 s + c_2},$$

$$(s+1)^2 Y(s) = a_1(s) Y(s) + \bar{\theta} F(s) + W(s) + D(s),$$

откуда

$$Y(s) = \frac{a_1(s)}{(s+1)^2} Y(s) + \frac{\bar{\theta} F(s)}{(s+1)^2} + \frac{W(s)}{(s+1)^2} + \frac{D(s)}{(s+1)^2}, \quad (4)$$

где полиномы $a_1(s) = (s+1)^2 - a(s)$ и $a(s) = s^2 + c_1s + c_2$. Из уравнения (4) имеем

$$y(t) = \frac{a_1(p)}{(p+1)^2} y(t) + \frac{\bar{\theta}}{(p+1)^2} y^3(t) + \frac{1}{(p+1)^2} w(t) + \varepsilon_y, \quad (5)$$

где $f(y) = y^3$ и $\varepsilon_y(t) = L^{-1} \left\{ \frac{D(s)}{(s+1)^2} \right\}$ – экспоненциально затухающая функция времени, определяемая ненулевыми начальными условиями.

Пренебрегая экспоненциально затухающим слагаемым $\varepsilon_y(t) = L^{-1} \left\{ \frac{D(s)}{(s+1)^2} \right\}$,

произведем параметризацию модели (5). Для этого рассмотрим вспомогательные фильтры вида

$$\xi_1(t) = \frac{1}{(p+1)^2} y(t), \quad (6)$$

$$\xi_2(t) = \frac{1}{(p+1)^2} y^3(t). \quad (7)$$

Подставляя (6) и (7) в уравнение (5), получаем

$$y(t) = a_1(p)\xi_1(t) + \bar{\theta}\xi_2(t) + \bar{w}(t), \quad (8)$$

где функция $\bar{w}(t) = \frac{1}{(p+1)^2} w(t)$. Из уравнения (8) имеем

$$y(t) = \delta(t) + \bar{w}(t), \quad (9)$$

где функция $\delta(t) = a_1(p)\xi_1(t) + \bar{\theta}\xi_2(t)$.

Рассмотрим фильтр (6)

$$\xi_1(t) = \frac{1}{(p+1)^2} y(t) = \frac{1}{(p+1)^2} \delta(t) + \frac{1}{(p+1)^2} \bar{w}(t),$$

откуда

$$\frac{1}{(p+1)^2} \bar{w}(t) = \xi_1(t) - \frac{1}{(p+1)^2} \delta(t). \quad (10)$$

Поскольку сигнал $w(t) = A \sin(\omega t + \phi)$, то в силу гурвицевости полинома $(p+1)^2$ функцию $\bar{w}(t)$ можно представить следующим образом:

$$\bar{w}(t) = \sigma \cdot \sin(\omega t + \varphi),$$

$$p^2 \bar{w}(t) = -\sigma \omega^2 \sin(\omega t + \varphi) = \theta \bar{w}(t),$$

где $\theta = -\omega^2$.

Запишем последнее уравнение следующим образом:

$$\begin{aligned} (p+1)^2 \bar{w}(t) &= p^2 \bar{w}(t) + 2p\bar{w}(t) + \bar{w}(t) = \\ &= \theta \bar{w}(t) + 2p\bar{w}(t) + \bar{w}(t) = (2p+1)\bar{w}(t) + \theta \bar{w}(t), \end{aligned}$$

$$(p+1)^2 \bar{w}(t) = (2p+1)\bar{w}(t) + \theta \bar{w}(t).$$

Поскольку $\bar{w}(t) = y(t) - \delta(t)$, то

$$\begin{aligned} (p+1)^2 (y(t) - \delta(t)) &= (2p+1)(y(t) - \delta(t)) + \theta (y(t) - \delta(t)) = \\ &= (p+1)^2 \left[\frac{(2p+1)}{(p+1)^2} (y(t) - \delta(t)) + \frac{\theta}{(p+1)^2} (y(t) - \delta(t)) \right]. \end{aligned}$$

Из последнего уравнения получаем

$$\begin{aligned}
 \bar{w}(t) &= (y(t) - \delta(t)) = \\
 &= \left[(2p+1)\xi_1 + \theta\xi_1 + \frac{(2p+1)}{(p+1)^2}(-\delta(t)) + \frac{\theta}{(p+1)^2}(-\delta(t)) \right] = \\
 &= \left[(2p+1+\theta)\xi_1 + \frac{(2p+1)}{(p+1)^2}(-\delta(t)) + \frac{\theta}{(p+1)^2}(-\delta(t)) \right] = \\
 &= \left[(2p+1+\theta)\xi_1 + \frac{(-p^2-2p-1)}{(p+1)^2}\delta(t) + \frac{p^2-\theta}{(p+1)^2}\delta(t) \right] =, \\
 &= \left[(2p+1+\theta)\xi_1 - \delta(t) + \bar{\delta}(t) \right], \tag{11}
 \end{aligned}$$

где функция

$$\begin{aligned}
 \bar{\delta}(t) &= \frac{p^2-\theta}{(p+1)^2}\delta(t) = \frac{p^2}{(p+1)^2}\delta(t) - \frac{\theta}{(p+1)^2}\delta(t) = \\
 &= \frac{p^2 a_1(p)}{(p+1)^2}\xi_1(t) - \frac{\theta a_1(p)}{(p+1)^2}\xi_1(t) + \frac{p^2 \bar{\theta}}{(p+1)^2}\xi_2(t) - \frac{\bar{\theta}\theta}{(p+1)^2}\xi_2(t). \tag{12}
 \end{aligned}$$

Проведем преобразования модели (11):

$$\theta\dot{\xi}_1(t) + \bar{\delta}(t) = \bar{w}(t) + \delta(t) - 2\dot{\xi}_1(t) - \xi_1(t) = y(t) - 2\dot{\xi}_1(t) - \xi_1(t). \tag{13}$$

Обозначим

$$z(t) = \theta\dot{\xi}_1(t) + \bar{\delta}(t) = y(t) - 2\dot{\xi}_1(t) - \xi_1(t), \tag{14}$$

где в силу измеримости сигналов $y(t)$, $\xi_1(t)$ и $\dot{\xi}_1(t)$ функция $z(t)$ также измеряется.

Подставляя в (14) уравнение (12), получаем

$$\begin{aligned}
 z(t) &= \theta\dot{\xi}_1(t) + \frac{p^2 a_1(p)}{(p+1)^2}\xi_1(t) - \frac{\theta a_1(p)}{(p+1)^2}\xi_1(t) + \\
 &+ \frac{\bar{\theta} p^2}{(p+1)^2}\xi_2(t) - \frac{\bar{\theta}\theta}{(p+1)^2}\xi_2(t) =, \\
 &= \theta \left(1 - \frac{a_1(p)}{(p+1)^2} \right) \dot{\xi}_1(t) + \bar{\theta} \frac{p^2}{(p+1)^2} \xi_2(t) - \bar{\theta}\theta \frac{1}{(p+1)^2} \xi_2(t) + \\
 &+ \frac{p^2 a_1(p)}{(p+1)^2} \xi_1(t). \tag{15}
 \end{aligned}$$

Учитывая, что $a_1(p) = a_1 p + a_0$, имеем

$$\begin{aligned}
 \theta \left(\frac{p^2 + 2p + 1 - a_1 p - a_0}{(p+1)^2} \right) \dot{\xi}_1(t) &= \theta \frac{p^2}{(p+1)^2} \dot{\xi}_1(t) + \\
 &+ (2\theta - a_1\theta) \frac{p}{(p+1)^2} \xi_1(t) + (\theta - a_0\theta) \frac{1}{(p+1)^2} \xi_1(t), \tag{16}
 \end{aligned}$$

$$\frac{p^2(a_1 p + a_0)}{(p+1)^2} \dot{\xi}_1(t) = a_1 \frac{p^2}{(p+1)^2} \dot{\xi}_1(t) + a_0 \frac{p^2}{(p+1)^2} \dot{\xi}_1(t). \tag{17}$$

Подставляя (16), (17) в уравнение (15), получаем

$$\begin{aligned}
z(t) = & (\theta + a_0) \frac{p^2}{(p+1)^2} \xi_1(t) + \bar{\theta} \frac{p^2}{(p+1)^2} \xi_2(t) - \bar{\theta} \theta \frac{1}{(p+1)^2} \xi_2(t) + \\
& + a_1 \frac{p^2}{(p+1)^2} \dot{\xi}_1(t) + (2\theta - a_1\theta) \frac{p}{(p+1)^2} \xi_1(t) + \\
& + (\theta - a_0\theta) \frac{1}{(p+1)^2} \xi_1(t). \tag{18}
\end{aligned}$$

Из уравнения (18) имеем

$$z(t) = \psi_1(t)\theta_1 + \psi_2(t)\theta_2 + \psi_3(t)\theta_3 + \psi_4(t)\theta_4 + \psi_5(t)\theta_5 + \psi_6(t)\theta_6, \tag{19}$$

где неизвестные параметры

$$\theta_1 = \theta + a_0, \theta_2 = \bar{\theta}, \theta_3 = -\bar{\theta}\theta, \theta_4 = a_1, \theta_5 = 2\theta - a_1\theta, \theta_6 = (\theta - a_0\theta)$$

и известные функции

$$\begin{aligned}
\psi_1(t) = \frac{p^2}{(p+1)^2} \xi_1(t), \quad \psi_2(t) = \frac{p^2}{(p+1)^2} \xi_2(t), \quad \psi_3(t) = \frac{1}{(p+1)^2} \xi_2(t), \\
\psi_4(t) = \frac{p^2}{(p+1)^2} \dot{\xi}_1(t), \quad \psi_5(t) = \frac{p}{(p+1)^2} \xi_1(t), \quad \psi_6(t) = \frac{1}{(p+1)^2} \xi_1(t).
\end{aligned}$$

Для оценки неизвестных параметров модели (19) будем использовать адаптивный наблюдатель вида

$$\hat{z}(t) = \psi_1(t)\hat{\theta}_1 + \psi_2(t)\hat{\theta}_2 + \psi_3(t)\hat{\theta}_3 + \psi_4(t)\hat{\theta}_4 + \psi_5(t)\hat{\theta}_5 + \psi_6(t)\hat{\theta}_6, \tag{20}$$

$$\dot{\hat{\theta}}_i = k_i \psi_i(t) e(t) = k_i \psi_i(t) (\ddot{\xi}_1 - \hat{z}(t)), \tag{21}$$

где постоянный коэффициент $k_i > 0, i = \overline{1,6}$.

Легко показать, что адаптивный наблюдатель вида (20), (21) обеспечивает

$$\lim_{t \rightarrow \infty} |z(t) - \hat{z}(t)| = 0, \tag{22}$$

$$\lim_{t \rightarrow \infty} |\theta_i - \hat{\theta}_i(t)| = 0. \tag{23}$$

4. Результаты моделирования

Произведем моделирование предлагаемой в статье схемы адаптивного восстановления неизвестного параметра $\bar{\theta}$. Моделирование проводилось при следующих параметрах хаотической системы Дуффинга (1): $c_1 = -0,5, c_0 = 0, w(t) = 7 \sin(t)$. Алгоритм настройки параметров имеет вид

$$\begin{aligned}
\dot{\hat{\theta}}_1 = 18\psi_1(t)(\ddot{\xi}_1 - \hat{z}(t)), \quad \dot{\hat{\theta}}_2 = 14\psi_2(t)(\ddot{\xi}_1 - \hat{z}(t)), \quad \dot{\hat{\theta}}_3 = 14\psi_3(t)(\ddot{\xi}_1 - \hat{z}(t)), \\
\dot{\hat{\theta}}_4 = 14\psi_4(t)(\ddot{\xi}_1 - \hat{z}(t)), \quad \dot{\hat{\theta}}_5 = 14\psi_5(t)(\ddot{\xi}_1 - \hat{z}(t)), \quad \dot{\hat{\theta}}_6 = 14\psi_6(t)(\ddot{\xi}_1 - \hat{z}(t)).
\end{aligned}$$

Результаты компьютерного моделирования для различных значений параметра $\bar{\theta}$ приведены на рис. 2–5. На рис. 2–3 приведены результаты моделирования для $\bar{\theta} = -0,5$. Из результатов моделирования видно, что $\hat{\theta}_2 \rightarrow -0,5$.

На рис. 4–5 приведены результаты моделирования при $\bar{\theta} = -0,75$. Из результатов моделирования видно, что $\hat{\theta}_2 \rightarrow -0,75$.

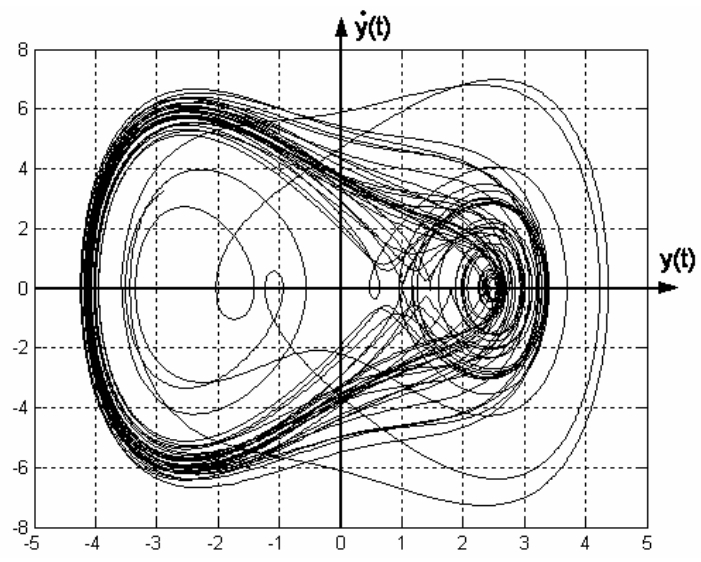


Рис. 2. Результаты моделирования при $\bar{\theta} = -0,5$. Фазовый портрет

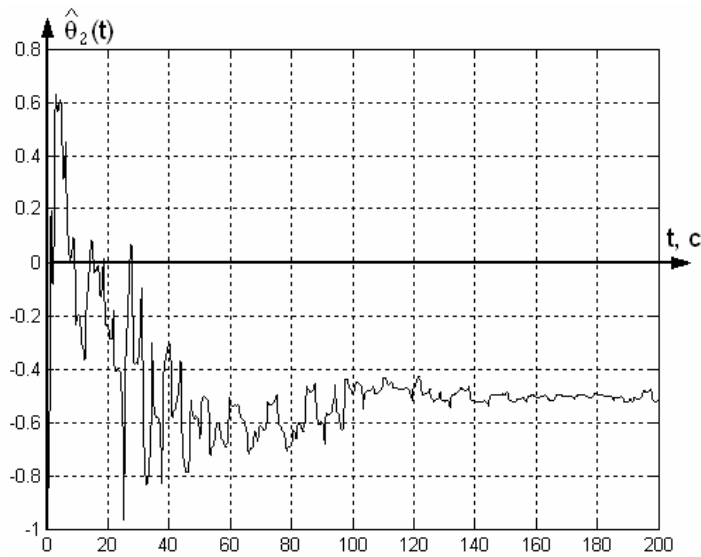


Рис. 3. Результаты моделирования при $\bar{\theta} = -0,5$. График функции $\hat{\theta}_2 = \hat{\bar{\theta}}$

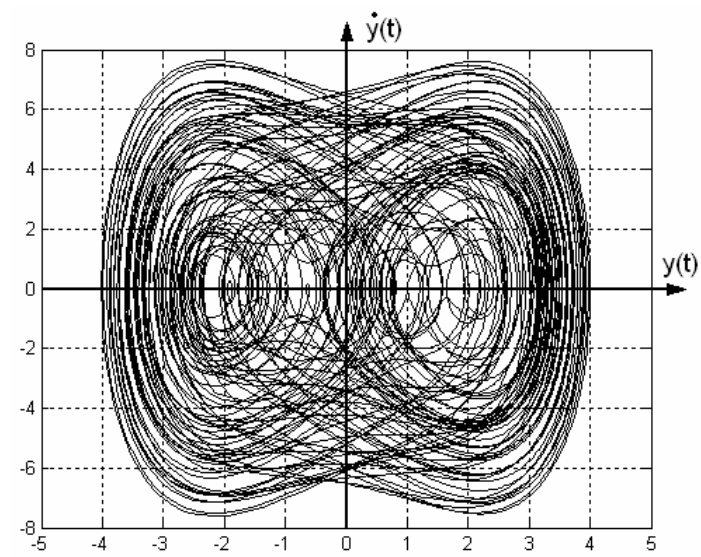


Рис. 4. Результаты моделирования при $\bar{\theta} = -0,75$. Фазовый портрет

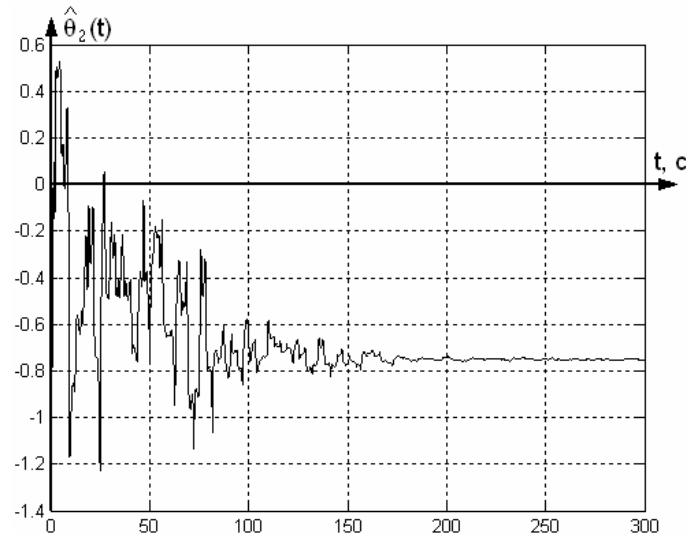


Рис. 5. Результаты моделирования при $\bar{\theta} = -0,75$. График функции $\hat{\theta}_2 = \hat{\bar{\theta}}$

Таким образом, результаты моделирования иллюстрируют работоспособность предлагаемой схемы адаптивной оценки неизвестного параметра $\bar{\theta}$ модели Дуффинга вида (1).

5. Заключение

С использованием адаптивного наблюдателя вида (20), (21) для хаотической системы Дуффинга решена задача восстановления неизвестного закодированного параметра $\bar{\theta}$. Данный результат, в отличие от известных аналогов, предусматривает измерение только выходного сигнала хаотической системы, а также позволяет находить неизвестный закодированный параметр $\bar{\theta}$ в условиях полной параметрической неопределенности модели (1).

Литература

1. Фрадков А.Л. Кибернетическая физика: принципы и примеры. СПб: Наука, 2003.
2. Fradkov A.L., Nijmeijer H., Markov A. Adaptive observer-based synchronization for communications // *Int. J. Bifurcations and Chaos*. 1997. 10, 12, P. 2807–2814.
3. Markov A.Yu., Fradkov A.L. Adaptive synchronization of coupled chaotic system // *Proc. Int. Conf. «Fractals and Chaos in Chemical Engineering»*, Rome, 1996, P. 153–154.
4. Fradkov A.L., Hill D. Exponential feedback passivity and stabilizability of nonlinear systems // *Automatica*, 1998, 6, P. 697–703.
5. Huijberts H., Nijmeijer H., Willems R. System identification in communication with chaotic system // *IEEE trans. Circ. Syst. – I*. 2000. V. 47, P. 800–808.
6. Nikiforov V.O., Andrievsky B.R. Adaptive observers for nonlinear nonpassifiable systems with applications to signal transmission // *Proc. 41th IEEE Conf. Decision and Control*, Las Vegas, 2002, 10–13 Dec.
7. Efimov D.V. Robust adaptive nonlinear partial observers for time-varying chaotic systems // *Proc. IEEE CDC 2004*, Atlantis, Paradise Island, Bahamas, 13–15 Dec.
8. Efimov D.V., Fradkov A.L. Adaptive partial observers with applications to time-varying chaotic systems // *Proc. IUTAM symp. On chaotic dynamics and control of systems and processes in mechanics*, Springer, 2005, P. 27–35.
9. Ефимов Д.В. Робастное и адаптивное управление нелинейными колебаниями. СПб: Наука, 2005.
10. Nikolis G., Prigogine I.R. Self-organization in non-equilibrium systems. Wiley, New York, 1977.

МОДЕЛИРОВАНИЕ И ОПТИМИЗАЦИЯ СУБЛИМАЦИОННОГО РОСТА ОБЪЕМНЫХ КРИСТАЛЛОВ НИТРИДА АЛЮМИНИЯ

Д.С. Базаревский

Научный руководитель – к.ф.-м.н., доцент А.С. Сегаль

В работе представлена детальная математическая модель сублимационного роста нитрида алюминия и ее упрощение на основе теории подобия. Рассчитанные на основе разработанной модели поля температуры в ростовом тигле, распределения скорости роста по поверхности кристалла и эволюции формы кристалла с хорошей точностью совпадают с экспериментальными данными. Исследованы основные закономерности массообмена между ростовым тиглем и окружающей средой, выработаны рекомендации по оптимизации процесса.

Введение

Нитрид алюминия (AlN) в настоящее время рассматривается как наиболее перспективный материал для производства подложек для полупроводниковых приборных структур на основе нитридов III группы – AlN, GaN, InN и их твердых растворов. Данные материалы обладают уникальным сочетанием физических свойств, которое определяет возможность создания на их основе высокоэффективных полупроводниковых приборов коротковолновой оптоэлектроники (синие, фиолетовые и ультрафиолетовые свето- и лазерные диоды), мощной высокочастотной электроники (мощные высокочастотные транзисторы) и «экстремальной» электроники, т.е. приборов, способных работать при высоких температурах и в химически агрессивных средах [1].

Однако, несмотря на очевидные преимущества III-нитридов по сравнению с традиционными полупроводниковыми материалами, развитие приборов нового поколения на их основе сдерживается отсутствием подходящих подложек. Широко используемые кремниевые, сапфировые, арсенид-галлиевые и менее распространенные карбид-кремниевые подложки характеризуются существенным рассогласованием свойств кристаллических решеток (постоянных решеток, коэффициентов термического расширения и т.д.) с III-нитридами, что приводит к снижению кристаллографического качества и избыточному образованию дефектов в выращенных на них III-нитридных гетероструктурах. В конечном счете, это приводит к ухудшению характеристик приборов и к быстрой их деградации [2].

Идеальным подложечным материалом для III-нитридных гетероструктур были бы сами III-нитриды. В настоящее время потребность в таких подложках частично удовлетворяется за счет GaN подложек, которые обычно производятся как квазиобъемные эпитаксиальные слои на сапфировых или арсенид-галлиевых подложках методом хлоридной газофазной эпитаксии [3, 4]. Данный метод обеспечивает скорость роста до 100–200 мкм/час в сочетании с плотностью дислокаций $\sim 10^7$ – 10^8 см⁻². Гораздо лучшее сочетание высокой скорости роста кристаллов и низкой концентрации дислокаций достигается в методе сублимационного роста объемных кристаллов AlN. В настоящее время данный метод продемонстрировал возможность роста объемных кристаллов AlN диаметром 1 дюйм со скоростями до 1 мм/час и с плотностью дислокаций до 500–1000 см⁻² [5]. Послеростовая обработка таких кристаллов позволяет получить высококачественные AlN подложки для III-нитридных гетероструктур.

Технология сублимационного роста AlN началась с работ Слэка с соавторами середины 1970-х годов [6, 7], которые впервые продемонстрировали возможность выращивания объемных монокристаллов AlN данным методом. В этих работах были исследованы режимы и механизмы сублимационного роста AlN, а также выделены основные проблемы, которые необходимо решить для создания эффективной промышленной технологии. В течение последующих 30 лет технология развивалась по пути решения

этих проблем в рамках лабораторных исследований усилиями нескольких научных групп из США, Германии и России.

Существенными проблемами технологии являются нахождение оптимальных ростовых условий и их поддержание при длительном росте крупных кристаллов, а также обеспечение оптимальной слабо выпуклой формы растущего кристалла. Прямое экспериментальное решение этих проблем затрудняется закрытостью ростового тигля и высокими температурами процесса, при этом моделирование выдвигается на передний план.

Впервые теоретическое исследование сублимационного роста AlN было проведено в [8]. В этой работе медленная реакция разложения двухатомного азота была признана основной лимитирующей стадией процесса. С учетом этого предположения была разработана простая одномерная модель, которая фактически не включала в рассмотрение эффекты массопереноса. В большинстве более поздних моделей лимитирующей стадией процесса считался перенос газовых компонент [9–12]. Первая модель этого типа была представлена в работе [9]. В ней поверхностное осаждение компонент представлялось посредством соотношения Герца-Кнудсена, в котором коэффициент прилипания для атомов Al принимался равным единице. С учетом того, что концентрация алюминия в газовой фазе мала, было получено явное соотношение для скорости роста AlN. Позднее этот подход был применен во многих других работах [10–12].

Для обеспечения предсказательного моделирования имеющиеся в литературе модели требуют существенного усовершенствования и расширения. Так, границы применимости подходов, в рамках которых рост лимитируется массопереносом или поверхностными кинетическими процессами, до сих пор остаются неясными. Первые неприменимы для низких значений давления [9, 13, 14], так как в этом случае состав газовой фазы далек от равновесного. В то же время в таких системах нельзя пренебрегать механизмами массопереноса, как это было сделано в [10]. Модель должна также учитывать все эффекты, которые могут стать важными при тех или иных условиях. К числу таких эффектов следует отнести испарение источника AlN с боковых и нижних поверхностей, массоперенос в областях между источником и кристаллом, массообмен с окружающим пространством через щели и поры в ростовом тигле и соответствующий перепад давления, постепенное изменение условий роста по мере эволюции формы источника и кристалла.

Модель сублимационного роста AlN, описывающая поверхностные кинетические эффекты, перенос компонент и влияние температуры и давления в рамках единого квазитермодинамического подхода, была разработана в [13, 14]. Однако в этих работах была рассмотрена только одномерная версия модели. В настоящей работе данный подход распространен на реальные двумерные задачи. В статье представлена комплексная модель сублимационного роста объемных кристаллов AlN, учитывающая все существенные физические эффекты, включая теплообмен, газовую динамику и поверхностную химию. На основании этой модели проведено исследование основных закономерностей, определяющих рост кристалла, в частности, рассмотрен массообмен между ростовым тиглем и окружающим пространством и эволюция формы источника и кристалла. По результатам расчетов выданы рекомендации по оптимизации ростовых условий и конструкции реактора.

Постановка задачи

Схема реактора для сублимационного роста AlN приведена на рис. 1. Цилиндрический ростовой тигель 1 расположен коаксиально в цилиндрической печи с резистивным нагревателем 2 в области с ниспадающим по вертикальной оси распределением температуры 3. Контроль температуры на крышке тигля осуществляется пирометром 4

через отверстие в изоляции 5. Это отверстие также обеспечивает радиационный отток тепла от крышки тигля и, как следствие, необходимое падение температуры от дна контейнера к его крышке. Поликристаллический источник AlN 6 расположен в области с большей температурой на дне тигля, в то время как монокристаллическая затравка крепится к более холодной крышке. Тигель установлен на стержне 8, который может вращаться, сглаживая возможные последствия неоднородного распределения температуры по углу. Внешние стенки 9 установки охлаждаются водой до комнатной температуры.

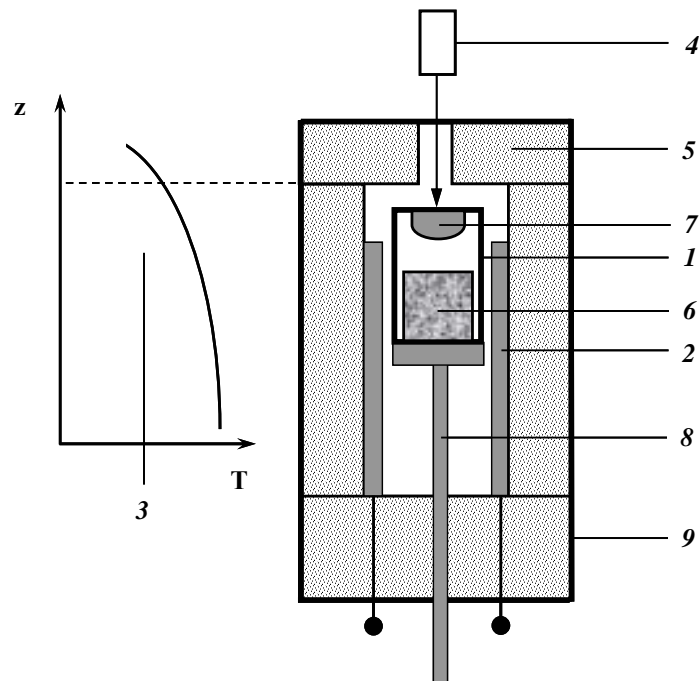


Рис. 1. Схема ростовой установки для сублимационного роста AlN

Источник AlN разлагается в высокотемпературной зоне с образованием двух газообразных компонентов – Al и N₂, которые переносятся к менее прогретому растущему монокристаллу, где осаждаются. Существенно, что в реальных ростовых процессах трудно обеспечить хорошую герметизацию тигля, при этом через щели между тиглем и внутренней областью печи (преимущественно это щель между крышкой тигля и его стенками) устанавливается определенный массообмен.

Модель процесса основана на следующих предположениях:

- в системе есть только две газообразные компоненты: Al и N₂ (малые посторонние примеси);
- скорость роста кристалла AlN зависит от локального состава пара и локальной температуры, но не зависит от ориентации кристалла (изотропный рост);
- испарение источника AlN происходит только с поверхности (используются плотные поликристаллические источники);
- эволюция формы источника и кристалла протекает значительно медленнее, чем процессы переноса (это позволяет описывать нестационарную эволюцию в квазистационарном приближении);
- излучение поверхности является серым;
- твердые тела непрозрачны;
- газ не поглощает и не рассеивает излучение;
- оптические свойства поверхности не зависят от температуры.

В рамках перечисленных предположений процессы переноса описываются полной системой уравнений Навье-Стокса:

$$\nabla \cdot (\rho \mathbf{V}) = 0, \quad (1)$$

$$\nabla \cdot (\rho C_i \mathbf{V} + \mathbf{J}_i) = 0, \quad i = Al, N_2, \quad (2)$$

$$\nabla \cdot (\rho \mathbf{V} \mathbf{V} + P \mathbf{E} + \mathbf{T}) = 0, \quad (3)$$

$$\nabla \cdot \left(\rho \mathbf{V} \sum_{i=Al, N_2} C_i H_i + \sum_{i=Al, N_2} \mathbf{J}_i H_i + \mathbf{q} \right) = 0, \quad (4)$$

$$P = R \rho T \sum_{i=Al, N_2} C_i / \mu_i, \quad (5)$$

$$\nabla \cdot \mathbf{q}_r = 0. \quad (6)$$

Здесь ρ – плотность газа, \mathbf{V} – вектор скорости газа, C_i – массовые доли газовых компонент, $\mathbf{J}_i = -\rho D \nabla C_i$ – диффузионные потоки газовых компонент (D – бинарный коэффициент диффузии для смеси Al/N₂), P – давление, \mathbf{E} – единичный тензор, $\mathbf{T} = -\eta(\nabla \mathbf{V} + \mathbf{V} \nabla)$ – тензор вязких напряжений (η – динамический коэффициент вязкости), $H_i = H_{0i} + \int_{T_0}^T C_{pi}(T) dT$ – парциальные энтальпии компонент смеси (H_{0i} – энтальпии образования при нормальных условиях, $C_{pi}(T)$ – удельные коэффициенты теплоемкости, T – температура, $T_0 = 298K$ – нормальная температура), $\mathbf{q} = -\lambda \nabla T$ – поток тепла (λ – коэффициент теплопроводности), μ_i – мольные массы компонент, \mathbf{q}_r – вектор лучистого потока.

Граничные условия для системы уравнений (1)–(6) на границе газа и твердых поверхностей записываются в следующем виде:

$$\rho \mathbf{V} \cdot \mathbf{n} = \sum_i G_i \mu_i, \quad (7)$$

$$\mathbf{V} \cdot \boldsymbol{\tau} = 0, \quad (8)$$

$$(\rho C_i \mathbf{V} + \mathbf{J}_i) \cdot \mathbf{n} = G_i \mu_i, \quad i = Al, N_2, \quad (9)$$

$$\left(\rho \mathbf{V} \sum_{i=Al, N_2} C_i H_i + \sum_{i=Al, N_2} \mathbf{J}_i H_i + \mathbf{q} + \mathbf{q}_r \right) \cdot \mathbf{n} = \mathbf{q}_s \cdot \mathbf{n}. \quad (10)$$

Здесь \mathbf{n} и $\boldsymbol{\tau}$ – соответственно единичные векторы в нормальном и тангенциальном направлениях, G_i – мольные потоки компонент на межфазной границе, $\mathbf{q}_s = -\lambda_s \nabla T_s$ – тепловой поток на границе (λ_s – коэффициент теплопроводности, T_s – температура твердого тела на границе). Поле температуры в твердых блоках подчиняется уравнению Лапласа, а в нагревателе, благодаря тепловыделению, – Пуассона. Постановка и стандартные методы численного решения таких задач описаны, например, в [15]).

На инертных поверхностях нормальные составляющие потоков компонент, очевидно, равняются нулю. На границах источника и кристалла AlN они могут быть найдены из квазитермодинамической модели. Следуя этой модели, нормальные составляющие потоков компонент на границе удовлетворяют соотношениям Герца-Кнудсена (см. [13, 14]):

$$G_i = \alpha_i(T) \beta_i(T) (P_i - P_i^e). \quad (11)$$

Здесь $\alpha_i(T)$ – температурно-зависимые коэффициенты прилипания компонент, $\beta_i(T) = (1/2\pi R \mu_i T)^{1/2}$ – коэффициенты Герца-Кнудсена, $P_i = R \rho T C_i / \mu_i$ – парциальные давления компонент, P_i^e – так называемые термодинамические давления компонент, представляющие собой многокомпонентные аналоги давления насыщенного одноком-

понентного пара. Разность $P_i - P_i^e = \Delta P_i$ может быть интерпретирована как скачок давления на герц-кнудсеновском слое. Два неизвестных термодинамических давления находятся из закона действующих масс и условия стехиометрического массообмена на межфазных границах, соответствующих поверхностной реакции $2Al + N_2 \rightleftharpoons 2AlN(\text{solid})$:

$$(P_{Al}^e)^2 P_{N_2}^e = K(T), \quad (12)$$

$$G_{Al} = 2G_{N_2}, \quad (13)$$

где $K(T)$ – температурно-зависимая константа равновесия приведенной реакции, которая определяется термодинамическими свойствами реагентов. Расчеты показывают, что эта константа равновесия с хорошей точностью приближается следующей аррениусовской аппроксимацией: $K(T) = 1.529 \cdot 10^{38} \exp(-1.502 \cdot 10^5 / T) Pa^3$. Коэффициент прилипания химически активных паров алюминия принимается равным единице. Для слабо-реактивного азота коэффициент прилипания был найден в [14, 16] также в виде аррениусовской аппроксимации:

$$\alpha_{N_2}(T) = 3.5 \cdot \exp(-30000/T) / [1 + 8 \cdot 10^{15} \exp(55000/T)]. \quad (14)$$

Несмотря на осевую симметрию, численная реализация поставленной задачи оказывается довольно затратной в смысле вычислительных ресурсов. Расчетная область состоит из большого числа блоков (от 100 до 200), размеры которых находятся в широком диапазоне значений (от $5 \cdot 10^{-5}$ м до $5 \cdot 10^{-1}$ м). Первое обстоятельство определяет необходимость обмена данными между блоками, что часто предполагает потерю точности при аппроксимации и интерполяции данных на границах блоков и замедляет сходимость. Необходимость достаточного сеточного разрешения блоков с малыми размерами увеличивает число элементов расчетной сетки. В каждой ячейке газовых блоков необходимо решать полную систему Навье-Стокса с учетом излучения. В условиях такой «ресурсоемкости» представляется целесообразным провести упрощение постановки задачи за счет ее расщепления на тепловую и масспереносную части.

Возможность такого расщепления обусловлена тремя обстоятельствами. Во-первых, низкая интенсивность массообмена между ростовым тиглем и окружающим газом позволяет не рассматривать перенос паров алюминия в области вне тигля. Во-вторых, слабая конвекция внутри тигля позволяет пренебречь конвективным переносом тепла внутри тигля по сравнению с теплопроводным и радиационным. В третьих, конвективный и теплопроводный теплообмен между охлаждаемой внешней стенкой установки и экранной изоляцией оказываются пренебрежимо малыми по сравнению с радиационным. Последнее обстоятельство представляется неочевидным, так как в данной области возникает система свободноконвективных вихрей, и интенсивность конвективного теплообмена может резко усиливаться. Характер сходимости решения в полной постановке задачи указывает на возникновение автоколебаний, а оценка числа Рэлея ($Ra \sim 10^8$) не только подтверждает это, но и говорит о возможном возникновении переходного турбулентного течения.

Для того чтобы сравнить интенсивность переноса тепла по разным каналам в этой области, мы оценили безразмерные критерии подобия: число Нуссельта Nu и радиационный критерий Rad . Оценка сверху числа Нуссельта по стандартной аппроксимации для развитой турбулентности из [17] дает значение $Nu \sim 30$, что значительно меньше радиационного критерия $Rad \sim 700$, показывающего интенсивность теплообмена излучением по сравнению с теплопроводностью. Это свидетельствует о преобладании лучистого теплообмена над конвективным теплопереносом и теплопроводностью. Во всех других областях установки это преобладание еще более выражено. Таким образом, рас-

чет теплообмена в установке сводится к решению уравнения теплопроводности в твердых непрозрачных блоках установки и расчету излучения в газовых областях.

Полученное в результате расчета теплообмена распределение температуры используется для определения температуры на границах ростового тигля. Это распределение крайне незначительно меняется в процессе роста кристалла. Поэтому однажды рассчитанное распределение температуры по стенкам тигля можно далее использовать как неизменное граничное условие для серии квазистационарных шагов по времени при моделировании массопереноса в тигле и эволюции фронтов кристаллизации и испарения.

Анализ термодинамического равновесия

Термодинамическое равновесие в системе AlN(твердый)-Al/N₂(газ) возможно в поле постоянной температуры при постоянном давлении. В этом случае распределение концентраций компонент однородно во всей газовой фазе, а потоки компонент и скорость газа во всех точках равны нулю. Равновесные парциальные давления тогда удовлетворяют закону действующих масс:

$$P_{Al}^2 P_{N_2} = K(T). \quad (15)$$

Существенно, что данное уравнение не определяет равновесный состав пара единственным образом, т.е. система имеет термодинамическую степень свободы. Обычно в качестве замыкающего соотношения используется равенство суммы парциальных давлений давлению в системе.

$$P_{Al} + P_{N_2} = P. \quad (16)$$

Система уравнений (15)–(16) сводится к одному кубическому уравнению относительно одного из парциальных давлений, которое решается аналитически.

В системе (15)–(16) предполагается, что давление внутри тигля – некоторая заданная величина. На практике это означает, что уровень давления поддерживается извне, что предполагает незамкнутость системы, а значит, равновесие, строго говоря, невозможно. Если система замкнута, то давление, устанавливающееся в ней, заранее неизвестно. Тогда, вместо замыкающего уравнения (16) следует использовать соотношение

$$\frac{P_{Al}}{\mu_{Al}} - \frac{2P_{N_2}}{\mu_{N_2}} = \Delta. \quad (17)$$

где Δ – некоторая постоянная для данной системы величина, так называемый дисбаланс. Дисбаланс сохраняется неизменным, если в замкнутой системе протекают реакции испарения и/или осаждения, так как в результате реакции на один моль выделяющегося или испаряющегося азота по стехиометрическому соотношению приходится два моля алюминия.

В замкнутой системе давление, определяемое уравнениями (15), (17), становится зависимой величиной. Зависимость давления от дисбаланса проиллюстрирована графиком на рис. 2. Минимум давления приходится на точку $\Delta = 0$, причем давление в этой точке составляет

$$P^*(T) = \frac{3}{2} [2K(T)]^{\frac{1}{3}}. \quad (18)$$

Давление в замкнутой системе, таким образом, не может упасть ниже этого минимального значения. Используя аррениусовскую аппроксимацию константы равновесия, можем записать аппроксимацию для минимального (далее, критического) давления

$$P^*(T) = 1.011 \cdot 10^{13} \exp\left(\frac{-5.007 \cdot 10^4}{T}\right).$$

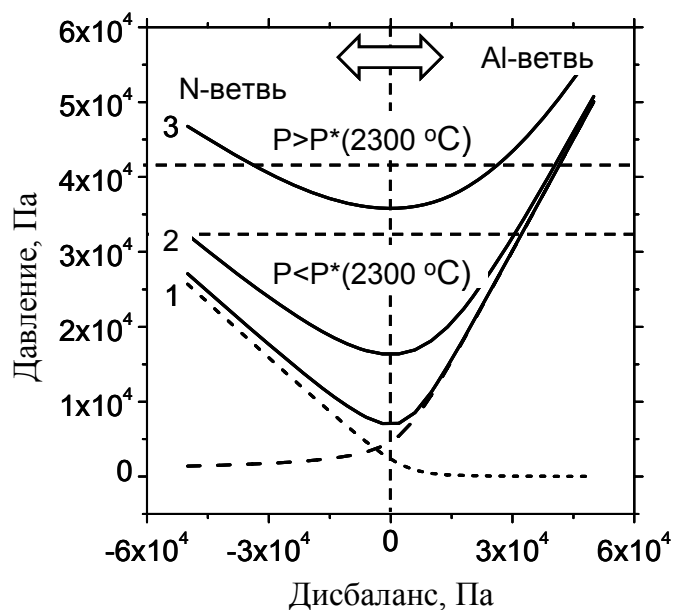


Рис. 2. Зависимости равновесных парциальных давлений Al (пунктир) и N₂ (мелкий пунктир) при температуре 2100°С, а также равновесных полных давлений при температурах 2100°С, 2200°С, и 2300°С (сплошные линии 1, 2 и 3, соответственно), от дисбаланса в замкнутом контейнере

Если система не замкнута, то через поры и щели в ростовом контейнере неизбежно возникает массообмен с окружающим пространством. Однако, если поле температуры однородно, а поры/щели достаточно малы, то это явление носит локальный характер, и в системе устанавливается состояние, близкое к равновесию, отвечающему давлению во внешнем пространстве. Парциальные давления компонент при этом образуют «почти» однородное поле, а их значения находятся из решения системы (15)–(16).

Как видно из графика рис. 2, эта система имеет два решения для давлений выше критического, одно решение для давления, равного критическому, и не имеет решений для значений давления меньше критического. В первом случае ($P > P^*$) два решения соответствуют равновесному пару, обогащенному по алюминию или азоту. Стоит отметить, что на практике обычно реализуется второй случай, когда мольная доля азота сильно превышает мольную долю алюминия.

Для случая, когда $P < P^*$, равновесие в системе невозможно. На практике это означает, что со всех поверхностей твердого нитрида алюминия идет интенсивное испарение. Это утверждение выполняется локально: если некоторый участок поверхности твердого AlN находится при температуре T и давлении, чье значение меньше критического давления, соответствующего данной температуре, то с данного участка поверхности происходит травление. Если давление в системе ниже, чем критическое давление для наиболее низкой температуры (критическое давление монотонно падает с падением температуры), то травление идет со всех поверхностей.

Верификация математической модели

На рис. 3 представлена зависимость температуры в центре крышки тигля от положения тигля на оси печи, рассчитанная с помощью разработанной модели (сплошная линия), в сравнении с экспериментальными данными, полученными с помощью цветового пирометра (точки). Здесь положение тигля отсчитывается от верхнего экрана (штриховая линия на рис. 1). Из графика видно, что расчетная линия более пологая, чем

та, на которой лежат экспериментальные точки. Такое расхождение может быть связано с некоторой неопределенностью данных по теплопроводности и оптическим свойствам материалов.

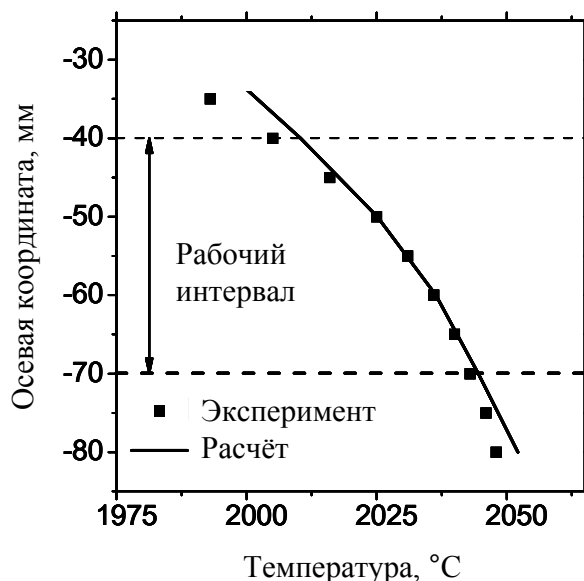


Рис. 3. Зависимость температуры на крышке тигля от осевого положения тигля

Как бы то ни было, достигнуто достаточно хорошее согласование расчетных и экспериментальных данных для положений тигля внутри рабочего диапазона, что создает хорошую базу для дальнейшего применения модели в изучении процесса сублиминационного роста AlN. Отметим, что в данном случае как эксперимент, так и расчеты проводились для пустого тигля, т.е. без источника и кристалла. Поле температуры и ее осевое распределение, однако, отличаются слабо (в пределах пяти градусов) от соответствующих распределений, полученных из расчета с учетом поликристаллического источника и кристалла (см. рис. 4). Наличие поликристаллического источника, таким образом, не только практически не сказывается на поле температуры вне тигля, но и незначительно влияет на температурное поле внутри него.

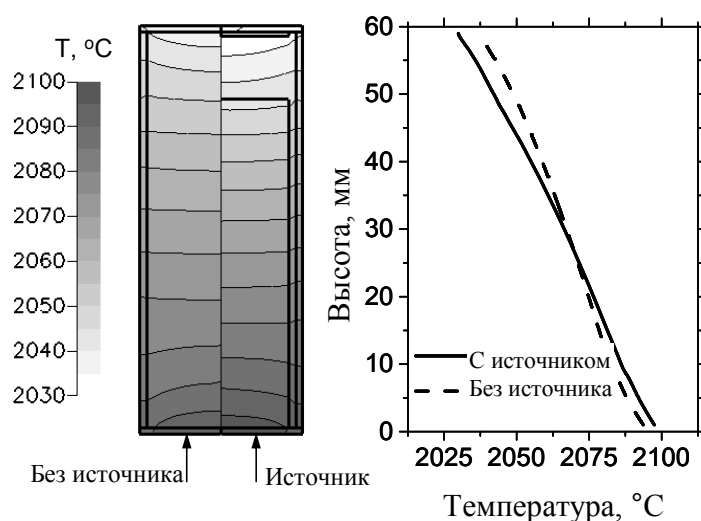


Рис. 4. Рассчитанные поле температуры в тигле и распределение температуры по оси тигля для пустого тигля и тигля с источником и затравкой

Рис. 5 иллюстрирует хорошее согласование расчетной и экспериментально наблюдаемой форм кристалла после двадцатичасового ростового процесса. Интересно, что выросший кристалл имеет грибовидную форму. Этот эффект обусловлен массообменом через щель под крышкой тигля. Давление в тигле для стационарного (в нашем случае квазистационарного) процесса всегда выше, чем давление снаружи тигля, поэтому возникает конвективный поток пара Al/N₂ через щель. Одновременно в обратном направлении идет диффузия алюминия и встречная диффузия азота. Локальное обеднение пара алюминием вблизи края растущего кристалла приводит к слабому травлению этой области.

Квазитермодинамическая модель описывает эти эффекты количественно. Кроме того, она позволяет получить простое выражение, связывающее скорость роста/травления с локальными температурой и составом пара. Действительно, если выразить термодинамические давления из (11) как $P_i^e = P_i - G_i / \alpha_i \beta_i$ и подставить их в (12), то, используя (13), после некоторых преобразований получим уравнение для полного мольного потока G на границе:

$$\frac{P_{Al}^2 P_{N_2}}{K(T)} \left(1 - \frac{2}{3} \frac{G}{\beta_{Al} P_{Al}} \right)^2 \left(1 - \frac{1}{3} \frac{G}{\alpha_{N_2} \beta_{N_2} P_{N_2}} \right) = 1. \quad (19)$$

Если состав пара равновесный, т.е. $P_{Al}^2 P_{N_2} = K(T)$, то уравнение (19) имеет очевидное решение $G=0$. Малое отклонение от равновесия порождает малый межфазный поток G . Линеаризуя уравнение (19) по малому параметру G , находим его приближенное решение:

$$G \approx \frac{P_{Al}^2 P_{N_2} / K(T) - 1}{4/(3\beta_{Al} P_{Al}) + 1/(3\alpha_{N_2} \beta_{N_2} P_{N_2})}. \quad (20)$$

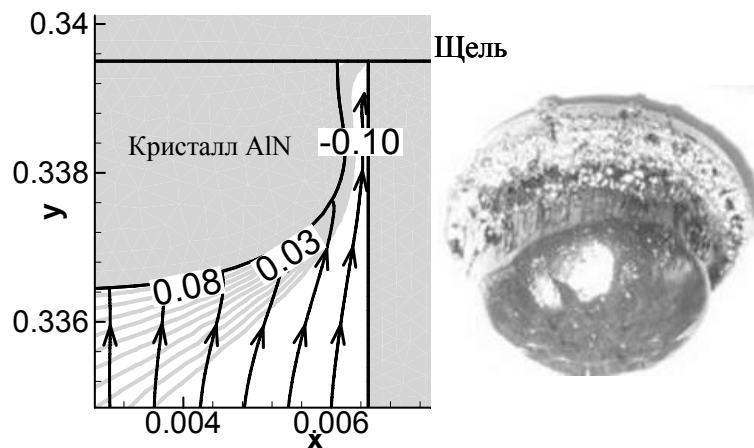


Рис. 5. Расчетная и экспериментальная формы кристалла после 20-часового процесса. Поле пересыщения и линии тока для расчета

Последнее соотношение показывает, что величина $S = P_{Al}^2 P_{N_2} / K(T) - 1$, которая может быть интерпретирована как локальное пересыщение, является движущей силой межфазного массообмена, при этом величина в знаменателе $R = 4/(3\beta_{Al} P_{Al}) + 1/(3\alpha_{N_2} \beta_{N_2} P_{N_2})$ может рассматриваться как кинетическое сопротивление. Положительное или отрицательное значение пересыщения свидетельствует, соответственно, о локальном осаждении или травлении AlN. Это проиллюстрировано рис. 5, где изображены линии тока и поле пересыщения вблизи поверхности растущего кристалла. Видно, что кристалл растет в области, где пересыщение положительно, и

травится там, где оно отрицательно. В терминах пересыщения могут быть объяснены многие другие эффекты, возникающие при сублимационном росте AlN.

Из рис. 5 видно, что абсолютное значение пересыщения гораздо меньше единицы. Малость пересыщения обеспечивается малым отклонением поля температуры от ее среднего значения в тигле. Интересно, что в этом случае линеаризацией и последующим преобразованием уравнений переноса и граничных условий можно свести решаемую систему к одному уравнению Лапласа относительно пересыщения с условиями III рода на ростовых поверхностях.

Обсуждение результатов

Для обеспечения достаточно высоких скоростей роста (порядка 100 мкм/час) поле температуры, представленное на рис. 4, непригодно. Для того чтобы получить результаты, более близкие к практическим, задаваемая на нагревателе мощность была повышена, значение температуры на крышке тигля при этом стало равным 2235 °С. Следует отметить, что поле температуры при этом не только не изменилось качественно, но сохранились практически неизменными значения перепадов температуры между источником и затравкой, между дном тигля и его крышкой, т.е. поле температуры равномерно сместилось примерно на 200 °С вверх.

С использованием рассчитанного поля температуры были получены зависимости скорости роста, коэффициента потерь и давления в тигле от степени его замкнутости и от внешнего давления. Для этого были проведены три серии расчетов в широком диапазоне внешнего давления с разными размерами щели под крышкой тигля: 100 мкм ширина и 1 мм длина (наиболее открытый тигель), 50 мкм ширина и 1 мм длина и 50 мкм ширина и 10 мм длина (наиболее замкнутый тигель).

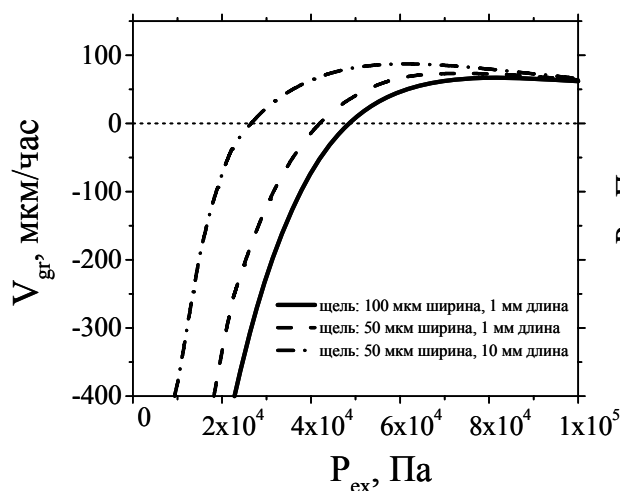


Рис. 6. Зависимость скорости роста от внешнего давления для разных размеров щели

На рис. 6 изображены графики зависимости скорости роста от внешнего давления. Критическое давление при данной температуре составляет $2.2 \cdot 10^4$ Па (напомним, что если давление в системе ниже критического, то травление происходит на всех границах AlN). В число дополнительных факторов, порождающих травление кристалла, как уже указывалось, входит встречная диффузия азота по щели. Приведенный график иллюстрирует данный эффект, например, для значения давления $3 \cdot 10^4$ Па: чем выше степень открытости тигля, тем ниже скорость роста. Зависимости скорости роста от давления имеют слабо выраженный максимум, значение которого также зависит от степени открытости тигля и лежит в промежутке от $5 \cdot 10^4$ Па до $9 \cdot 10^4$ Па.

Примечательно, что кристалл может расти, даже если значение давления вне тигля меньше критического. Для объяснения этого эффекта рассмотрим рис. 7, на котором приведен график зависимости давления внутри тигля от давления во внешней области. Из графика следует, что давление внутри тигля всегда больше давления снаружи. Разница эта тем больше, чем выше замкнутость тигля и чем меньше значение внешнего давления. Таким образом, если щель под крышкой тигля настолько узкая и длинная, что внутреннее давление превышает критическое, а эффект встречной диффузии азота локализован вблизи края затравки, то пересыщение и, следовательно, скорость роста на затравке – положительные.

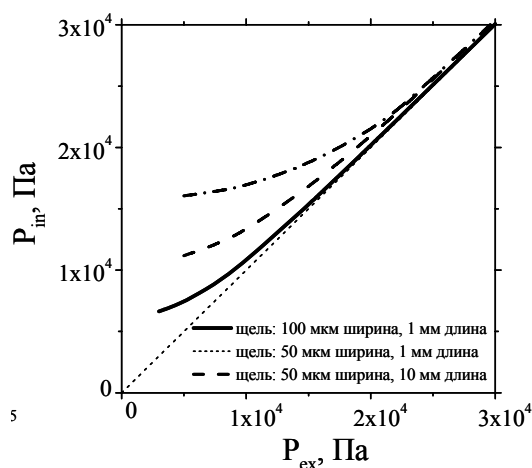


Рис. 7. Зависимость давления в тигле от внешнего давления для разных размеров щели

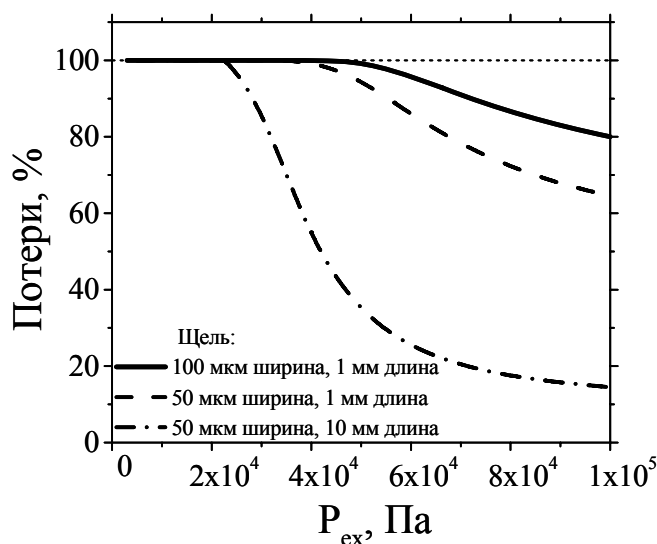


Рис. 8. Зависимость потерь от внешнего давления для разных размеров щели

Однако рост при низких давлениях и в плохо герметизированном тигле чрезвычайно неэкономичен. Это наглядно иллюстрирует рис. 8, на котором изображены графики зависимости потерь (потерями мы называем отношение массы пара, покинувшего тигель за некоторый промежуток времени, к массе испарившегося за то же время источника) от внешнего давления для ранее приведенных конфигураций щели). Одним из направлений оптимизации процесса является минимизация потерь. На практике удается

достичь их снижения до 5–10 %, однако характерным расчетным значением является 30%. Процесс проводят при давлениях $6 \cdot 10^4$ – $8 \cdot 10^4$ Па, поэтому основным инструментом минимизации потерь является герметизация тигля. Важным фактором минимизации потерь является также перепад температуры между источником и кристаллом: чем он больше, тем меньше коэффициент потерь. Однако при увеличении перепада температуры сильно увеличивается скорость роста, и качество кристалла может резко ухудшаться (может начаться рост поликристалла). Таким образом, оптимальными являются следующие ростовые условия: температура на крышке тигля 2200–2300 °С; давление $6 \cdot 10^4$ – $8 \cdot 10^4$ Па, максимально герметизированный тигель. Такие условия обеспечивают скорость роста порядка 100 мкм/час, затравливание и рост монокристалла, небольшие значения потерь.

Заключение

В результате проделанной работы выявлены основные физические эффекты, влияющие на теплоперенос в установке для сублимационного роста нитрида алюминия и массоперенос в тигле, оценена их взаимная интенсивность, разработан метод расчета. Хорошее совпадение расчетов и эксперимента позволило верифицировать модель. Проведенные несколько серий расчетов с вариацией основных ростовых параметров позволили получить важные для понимания процесса зависимости скорости роста и потерь, которые позволили выбрать оптимальные условия роста монокристалла.

Литература

1. Rutberg et al. Gallium Nitride: a Material Opportunity. 2001.
2. GaN and Related Alloys - 2002. MRS Symp. Proc. 743 (2002).
3. R.J. Molnar, W. Götz, L.T. Romano, N.M. Johnson, J. Cryst. Growth 178 (1997) 147.
4. R. Fornari, et al, Mat. Sci. Eng. B79 (2001) 159.
5. L.J. Schowalter et al. Bulk AlN crystal growth and substrate preparation: the native nitride alternative. Book of Int. Workshop on Bulk Nitride semicond. III, Zakopane, 2005, p. 49.
6. G.A. Slack, T.F. McNelly, J. Cryst. Growth 34 (1976) 263.
7. G.A. Slack, T.F. McNelly, J. Cryst. Growth 42 (1977) 560.
8. P.M. Dryburgh, J. Cryst. Growth 125 (1992) 65.
9. L. Liu, J.H. Edgar, J. Cryst. Growth 220 (2000) 243.
10. B. Wu, R. Ma, H. Zhang, M. Dudley, R. Schlessler, Z. Sitar, J. Cr. Gr. 253 (2003) 326.
11. V. Noveski, R. Schlessler, S. Mahajan, S. Beaudoin, Z. Sitar, J. Cr. Gr. 264 (2004) 369.
12. B. Wu, R. Ma, H. Zhang, V. Prasad, J. Cryst. Growth 266 (2004) 303.
13. S.Yu. Karpov, D.V. Zimina, Yu.N. Makarov, E.N. Mokhov, A.D. Roenkov, M.G. Ramm, Yu.A. Vodakov, Phys. Stat. Sol. A 176 (1999) 435.
14. A.S. Segal, S.Yu. Karpov, Yu.N. Makarov, E.N. Mokhov, A.D. Roenkov, M.G. Ramm, Yu.A. Vodakov, J. Cryst. Growth 211 (2000) 68.
15. C.R. Kleijn. In: “Computational Modeling in Semiconductor Processing”, Ed. M. Meyyappan, Artech House, Boston-London, 1995, p. 97.
16. S.Yu. Karpov, A.V. Kulik, A.S. Segal, M.S. Ramm, Yu.N. Makarov. Phys. Stat. Sol. 188 (2001) 763.
17. Hollands K.G.T. et al Heat Mass Transfer, 18, 879 (1975).

НАШИ АВТОРЫ

- Алеев Алексей Муратович** – студент кафедры оптико-электронных приборов и систем
- Анисимов Андрей Геннадиевич** – студент кафедры оптико-электронных приборов и систем
- Арановский Станислав Владимирович** – аспирант кафедры систем управления и информатики
- Араканцев Константин Геннадьевич** – студент кафедры оптико-электронных приборов и систем
- Арефьева Наталья Николаевна** – младший научный сотрудник кафедры оптики квантоворазмерных систем
- Базаревский Денис Станиславович** – аспирант кафедры компьютерных технологий
- Батшев Владислав Игоревич** – студент Московского государственного технического университета им. Н.Э. Баумана
- Березкина Анастасия Сергеевна** – студент кафедры оптики квантоворазмерных систем
- Блинников Андрей Алексеевич** – аспирант кафедры системы управления и информатики
- Богатинский Егор Маркович** – студент кафедры оптико-электронных приборов и систем
- Богданов Всеволод Викторович** – аспирант кафедры технологии приборостроения
- Бойков Владимир Иванович** – доцент кафедры систем управления и информатики
- Бондаренко Роман Владимирович** – аспирант кафедры технологии приборостроения
- Волков Александр Валерьевич** – студент кафедры оптических технологий
- Волков Игорь Сергеевич** – аспирант кафедры системы управления и информатики
- Ворона Алексей Михайлович** – аспирант кафедры оптико-электронных приборов и систем
- Денисов Артем Руфимович** – кандидат технических наук, сотрудник кафедры вычислительной техники Костромского государственного технологического университета
- Дехтерева Ксения Владимировна** – студент кафедры высшей математики
- Дьякова Надежда Вячеславовна** – студент кафедры высшей математики
- Ермолаева Елена Владимировна** – студент кафедры прикладной и компьютерной оптики
- Жуков Дмитрий Валерьевич** – студент кафедры оптико-электронных приборов и систем
- Иванов Александр Николаевич** – аспирант кафедры компьютеризации и проектирования оптических приборов
- Исаков Алексей Сергеевич** – аспирант кафедры систем управления и информатики
- Камнев Дмитрий Анатольевич** – аспирант кафедры систем управления и информатики
- Каряев Константин Викторович** – ассистент кафедры оптических технологий
- Ким Александр Александрович** – студент кафедры твердотельной оптоэлектроники
- Коровьяков Анатолий Николаевич** – доцент кафедры систем управления и информатики
- Кужаков Павел Викторович** – аспирант кафедры оптико-электронных приборов и систем

Ле Зуй Туан – аспирант кафедры прикладной и компьютерной оптики
Лесик Маргарита Алексеевна – студент кафедры общей физики НИИ лазерных исследований Санкт-Петербургского государственного университета
Лишкевич Антон Алексеевич – студент кафедры прикладной и компьютерной оптики
Лузина Наталья Петровна – аспирант кафедры проектирования компьютерных систем
Макаричев Глеб Вячеславович – студент кафедры оптических технологий
Масаева Людмила Александровна – студент кафедры оптических технологий
Мацерук Ксения Юрьевна – студент кафедры прикладной и компьютерной оптики
Мачихин Александр Сергеевич – студент кафедры оптико-электронных приборов научных исследований Московского государственного технического университета им. Н.Э. Баумана
Мерсон Алексей Дмитриевич – аспирант кафедры оптико-электронных приборов и систем
Мотылькова Марина Михайловна – аспирант кафедры систем управления и информатики
Немкова Анастасия Александровна – аспирант кафедры оптических технологий
Николаев Николай Анатольевич – аспирант кафедры систем управления и информатики
Орешков Андрей Александрович – аспирант кафедры систем управления и информатики
Пантюшин Антон Валерьевич – студент кафедры оптико-электронных приборов и систем
Полищук Григорий Сергеевич – аспирант кафедры системотехники оптических приборов и комплексов
Пруненко Юлия Константиновна – аспирант кафедры прикладной и компьютерной оптики
Пыркин Антон Александрович – студент кафедры системы управления и информатика
Ремицкая Анна Ярославовна – студент кафедры высшей математики
Репин Алексей Александрович – выпускник кафедры прикладной и компьютерной оптики
Син Сянмин – аспирант кафедры оптико-электронных приборов и систем
Соколов Дмитрий Александрович – аспирант кафедры информационно-навигационных систем
Суворова Ирина Юрьевна – студент кафедры компьютерной фотоники
Усик Александр Александрович – студент кафедры оптико-электронных приборов и систем
Фомина Юлия Николаевна – аспирант кафедры технологии приборостроения
Чан Куок Туан – аспирант кафедры прикладной и компьютерной оптики
Чупраков Сергей Александрович – аспирант кафедры системотехники оптических приборов и комплексов
Шарков Николай Александрович – сотрудник ФГУП ЦНИИ им. академика А.Н. Крылова
Шепелевич Андрей Николаевич – аспирант кафедры прикладной и компьютерной оптики
Шимко Александр Анатольевич – студент кафедры общей физики НИИ лазерных исследований Санкт-Петербургского государственного университета

СОДЕРЖАНИЕ

1. ОПТОТЕХНИКА	4
Лесик М.А., Шимко А.А. Использование фемтосекундного лазерного излучения для создания волноводных структур в халькогенидных стеклообразных полупроводниках.....	4
Арефьева Н.Н. Применение метода наноимпринта для единичного копирования полимерной френелевской и микрооптики.....	8
Березкина А.С. Влияние типа и концентрации красителя на формирование микроэлемента на торце оптического волокна.....	12
Батшев В.И. Геометрические и оптические свойства афокальной двухзеркальной системы.....	17
Араканцев К.Г. Исследование влияния вертикального градиента температуры воздушного тракта на погрешность измерений в авторефлексионной оптико-электронной системе.....	25
Богатинский Е.М. Анализ пространственного распределения энергии в планарной оптической равносигнальной зоне.....	33
Алеев А.М., Анисимов А.Г., Пантюшин А.В. Об основных погрешностях контроля соосности с помощью авторефлексионной оптико-электронной системы.....	39
Немкова А.А. Измерение показателя преломления однослойного просветляющего покрытия.....	45
Мерсон А.Д. Анализ схем построения систем измерения параметров угловой пространственной ориентации на основе явления анаморфозы.....	48
Ворона А.М. Анализ регулярного виньетирования в авторефлексионных углоизмерительных системах.....	52
Пруненко Ю.К. Исследование влияния технологических параметров на качество изображения призмной анаморфотной системы.....	58
Чупраков С.А. Свойства оптических систем, содержащих децентрированные планоидные элементы в области аберраций третьего порядка.....	63
Полищук Г.С. Искажения волнового фронта при изменении положения плоскости интерферограммы.....	72
Ермолаева Е.В. Варианты композиции зеркально-линзового объектива плананастигмата.....	77
Ермолаева Е.В. Схема контроля выпуклых несферических поверхностей вращения второго порядка методом анаберрационных точек.....	82
Ермолаева Е.В., Мацерук К.Ю., Репин А.А. Зеркально-линзовый концентрический объектив.....	86
Шепелевич А.Н. Параметрический синтез трехзеркальной оптической системы на основе базовой двухзеркальной.....	98
Каряев К.В. Разработка программного комплекса для нового поколения офтальмологических приборов. Прибор для исследования реакции глаза на внешние интенсивные засветки.....	105
Макаричев Г.В. Установка для исследования реакции глаза на внешние интенсивные засветки.....	110
Чан К.Т. Исследование влияния зазора в башмачной призме на качество изображения.....	114

Суворова И.Ю. Разработка и расчет оптической системы лазерной технологической установки 117

2. ОПТИЧЕСКОЕ ПРИБОРОСТРОЕНИЕ 122

Ле Зуй Туан. Разработка алгоритма и программы для расширения возможностей аппаратуры оценки качества изображения оптических систем 122

Ле Зуй Туан, Лишкевич А.А. Разработка и исследование компьютеризированного интерферометра с дифракционной решеткой на основе схемы Ронки и программного обеспечения для расшифровки интерферограмм 128

Иванов А.Н. Исследование муарового эффекта, возникающего при наложении частотных спектров объекта 137

Мачихин А.С. Однокамерная сканирующая стереоскопическая система для реконструкции трехмерной структуры объектов 142

Син Сянмин. Исследование контрольного элемента автоколлимационного угломера с гиперболической отражающей гранью 147

Жуков Д.В., Усик А.А. Адаптивный многопроходный алгоритм определения координат изображений точечных излучателей 151

Кужаков П.В. Сравнение нейросетевых алгоритмов при оптических методах коррекции сегментов отражающих поверхностей радиотелескопов 155

Ким А.А. Безапертурный рамановский наноспектрометр ближнего поля: демонстрация эффекта усиления оптического поля при регистрации рамановских сигналов 162

Волков А.В., Масаева Л.А. Разработка и создание конструкции рефлективного поляризирующего фильтра для ЖК-дисплеев, способствующего снижению потерь света 170

3. ТЕХНОЛОГИИ ПРИБОРОСТРОЕНИЯ, МЕХАТРОНИКА И РОБОТОТЕХНИКА 178

Соколов Д.А. Особенности построения инклинометра по диаметральной схеме на основе ДНГ 178

Фомина Ю.Н. Исследование алгоритмов оптимизации конфигурирования и распределения заказов на решении задач ТПП в среде виртуального предприятия 187

Богданов В.В., Бондаренко Р.В. Использование CMS-системы для разработки модуля расчета режимов резания 197

Богданов В.В. Интеграция систем автоматизированной подготовки производства в едином информационном пространстве 202

Денисов А.Р. Конструкторско-технологическое моделирование в условиях мелкосерийного машиностроительного производства 208

Лузина Н.П. Акустико-эмиссионный метод контроля при оценке степени деградации механических свойств и остаточного ресурса работоспособности трубных сталей 213

4. СИСТЕМНЫЙ АНАЛИЗ, МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ И УПРАВЛЕНИЕ В ТЕХНИЧЕСКИХ СИСТЕМАХ 219

Шарков Н.А. Алгоритмы имитационного моделирования и ситуационного анализа процессов противоаварийного управления сложными человеко-машинными системами в обучающих судовых тренажерах 219

Дьякова Н.В. Изучение эффективности обнаружения сигнала для многоканальной системы 230

Ремицкая А.Я. Марковские процессы и простейшие модели теории массового обслуживания. Компьютерное моделирование простейших моделей массового обслуживания	239
Дехтерева К.В. Моделирование тестов обнаружения сигнала в многоканальной системе и сравнение их свойств.....	249
Мотылькова М.М. Исследование режимов захвата и автосопровождения цели в системах пространственного слежения.....	260
Бойков В.И., Волков И.С., Коровьяков А.Н. Экспериментальное определение информативности аналоговых каналов микроконвертеров.....	268
Бойков В.И., Камнев Д.А., Орешков А.А. Оценка погрешности определения параметров интерференционного поля	276
Исаков А.С. Реализация наблюдателя состояний асинхронного двигателя с короткозамкнутым ротором в бездатчиковой системе векторного управления.....	280
Пыркин А.А. Управление в условиях запаздывания.....	287
Блинников А.А. Использование полиномов Котельникова в задаче аппроксимации модели термодинамического объекта.....	293
Арановский С.В., Николаев Н.А. Синтез адаптивного наблюдателя для хаотической системы Дуффинга	302
Базаревский Д.С. Моделирование и оптимизация сублимационного роста объемных кристаллов нитрида алюминия.....	309
НАШИ АВТОРЫ	321

**Научно-технический вестник СПбГУ ИТМО. Выпуск 38.
ИССЛЕДОВАНИЯ В ОБЛАСТИ ОПТИКИ, ПРИБОРОСТРОЕНИЯ
И УПРАВЛЕНИЯ. Труды молодых ученых / Главный редактор д.т.н.,
проф. В.Н. Васильев. – СПб: СПбГУ ИТМО, 2007. 326 с.**

**НАУЧНО-ТЕХНИЧЕСКИЙ ВЕСТНИК СПбГУ ИТМО
Выпуск 38**

**ИССЛЕДОВАНИЯ В ОБЛАСТИ ОПТИКИ, ПРИБОРОСТРОЕНИЯ
И УПРАВЛЕНИЯ.
Труды молодых ученых**

Главный редактор
доктор технических наук, профессор
В.Н. Васильев

Дизайн обложки В.А. Петров, А.А. Колокольников
Редакционно-издательский отдел СПбГУ ИТМО
Зав. РИО Н.Ф. Гусарова

Лицензия ИД № 00408 от 05.11.99.

Подписано в печать 10.10.07.

Заказ 1060. Тираж 100 экз.