

УПРАВЛЕНИЕ МАНИПУЛЯЦИОННЫМИ РОБОТАМИ С УПРУГИМИ ШАРНИРАМИ С ПРИМЕНЕНИЕМ ОБУЧЕНИЯ С ПОДКРЕПЛЕНИЕМ

Самородова М.Э.*, Каканов М.А*, Борисов О.И.*

**Факультет систем управления и робототехники, Университет ИТМО, Россия, 197101, Санкт-Петербург, Кронверкский пр., 49*

В робототехнике классическим является представление робота как жёсткой конструкции, однако данное упрощение динамики манипулятора стоит отнести к идеальной ситуации, которое оправдано только для медленного движения и малых взаимодействующих сил. На практике же у робототехнической системы присутствует механическая гибкость, основными причинами которой являются:

- уменьшение массы звеньев за счет использования легких материалов и утончения конструкции;
- использование податливых передаточных элементов.

Гибкость шарниров является обычным явлением для современных промышленных роботов, когда в качестве трансмиссий и передач применяются ремни, длинные валы, волновое или циклоидальные передачи. Под действием сил/крутящих моментов, возникающих при нормальной работе робота, эти компоненты, по своей сути, гибки, что влечёт за собой изменяющееся во времени смещение между положением приводов и положением ведомых звеньев. Без специального управляющего воздействия во время свободного движения наблюдается колебательное поведение, обычно небольшой величины, но относительно высокой частоты. Более того, некоторая форма неустойчивости (например, дребезжание) может возникать в задачах, связанных с контактом с окружающей средой.

Кроме того, в последнее время упругие элементы были сознательно добавлены в «мягких» роботах (например, SEA- или VSA-технологии) для физического взаимодействия человека с роботом и для достижения более естественной характеристики движения. Добавление дополнительной механической упругости в виде пружины гарантирует инерционную развязку между двигателем и ведомым звеном, тем самым уменьшая кинетическую энергию при нежелательных столкновениях с окружающей средой.

Добавление упругости в итоговую динамическую модель робота приводит к повышению его сложности как объекта управления, поэтому для робота с упругими шарнирами недостаточно классических методов управления. Перспективным методом управления может являться регулятор, построенный на основе расширенного наблюдателя, коэффициенты которого определяются с помощью обучения с подкреплением. Эффективность данного подхода обусловлена следующими факторами.

Системы управления с обратной связью зависят от информации об изменяющихся во времени состояниях системы. Во многих ситуациях такие измерения невозможны, и требуются альтернативные подходы. Одним из таких подходов является оценка переменных состояния с помощью добавления наблюдателя в систему. Фактически, наблюдателем является динамическая система с измерениями в качестве входных данных и оценкой состояний или возмущений в качестве выходных. Принципиальным отличием от многих других наблюдателей обладает расширенный наблюдатель состояния (ESO), который не зависит от точной математической модели системы и рассматривает все факторы, влияющие на работу системы, включая нелинейную динамику и неопределённости. ESO оценивает неопределённости вместе с состояниями системы, что позволяет подавлять или компенсировать возмущения.

Таким образом, ESO является эффективным инструментом для проектирования систем с неопределённостями, однако данный подход не позволяет достичь оптимального решения, поскольку нет точного знания о модели объекта. Введение в управление обучения с подкреплением (RL) позволит построить адаптивный контроллер, который ищет оптимальные решения управления в реальном времени методом проб и ошибок. Благодаря повторным попыткам поиска новых управляющих воздействий RL накапливает информацию, используемую для получения улучшенного контроля. Именно эта особенность RL позволяет адаптивным алгоритмам управления превосходить традиционные методы управления. Однако основное преимущество одновременно может служить и главным недостатком – гибкость RL может привести к многочисленным нежелательным стратегиям управления во время поиска оптимального решения, кроме того, помимо временной неоптимальности решения, может наблюдаться и неустойчивость системы, которая в реальных условиях нежелательна. Поэтому важно найти такой подход к построению системы управления с RL, который будет гарантировать стабильное поведение системы и точно отслеживать входной эталонный сигнал с течением времени.

Основная идея данной работы заключается в применении RL совместно с расширенным наблюдателем, который формирует робастные оценки состояния, с целью достижения субоптимального управления для манипуляционных роботов с упругими шарнирами.