

**ПРИМЕНЕНИЕ МАШИННОГО ЗРЕНИЯ И ТРАНСФОРМАЦИЯ ПОТОКА В
ЗВУКОВОЕ ИЗОБРАЖЕНИЕ ДЛЯ СЛАБОВИДЯЩИХ ЛЮДЕЙ**
Москинский Г.И. (Университет ИТМО), **Шарабанов А.Д.** (Самарский государственный
институт культуры)

Научный руководитель –к.б.н., доцент Папшев В.А.
(Самарский государственный технический университет)

Аннотация: Технологии машинного зрения и построения объемного звучания за последние годы продвинулись достаточно далеко, а технологии мобильных сетей 5G, запуск которых произойдет в ближайшее время, могут позволить слабовидящим людям в режиме онлайн передавать визуальную информацию через аудиальный канал. Это позволит открыть новые возможности и вывести на новый уровень жизнь таких людей. В данном докладе будет продемонстрированы концепция и варианты реализации аудиального зрения.

Введение. Развитие современных мобильных технологий, рост производительности и одновременное уменьшение размеров вычислительных устройств открывает новые возможности для решения проблем людей с ограниченными возможностями здоровья, таких как слабовидящие или слепые, которые ранее были недоступны, особенно если рассматривать несколько областей науки во взаимодействии.

Одновременное развитие таких направлений как машинное зрение, которое позволяет распознавать формы предметов, их цвет, назначение, направление движения, лица, габариты удаленных предметов, определять расстояние до множества точек, и программного обеспечения по моделированию объемного звучания, которое позволяет создавать глубину и ширину сцены, позволяет задуматься об устройстве, которое будет иметь возможность транслировать визуальную составляющую окружения в аудиальный канал человека и иметь гибкую возможность настройки глубины и ширины сцены, детальности описания, режимов представления (навигация, распознавание предметов, лиц и мимики, описание пространства, чтение текста).

Нами были изучены современные доступные устройства, помогающие слепым и слабовидящим людям ориентироваться в пространстве и решать ряд бытовых задач и методы работы этих устройств.

На основании изученных имеющихся на данный момент технических средств помощи слепым и слабовидящим людям их можно разделить на две группы:

1. помогающие с навигацией в пространстве;
2. помогающие с восприятием окружения (описание предметов, цветов, эмоций и пр.).

По способу взаимодействия их так же можно объединить в две группы:

1. Устройства с тактильным каналом передачи информации;

К таким устройствам относятся белая трость, брайлевский дисплей, устройство проектирования картинки на лоб

2. Устройства с аудиоканалом.

К этой группе относятся большинство устройств — определители цвета, измерители расстояний, часы, дальномеры, устройства чтения текста, устройства преобразующие 2D изображение в аудио образ.

На рынке представлены как зарубежные, так и отечественные разработки, осуществляющие преобразование видеоизображения в аудио образ. Среди них наиболее интересная с точки зрения реализации и продвинутая российская разработка The vOICe.

Основная часть. Прототип и принципы устройства, рассмотренного в докладе, основываются на имеющихся данных и возможностях машинного зрения и принимается с допущением работающего устройства. В периметр нашего доклада включается принцип

трансляции получаемых с помощью машинного зрения данных в аудио составляющую с использованием сцены и описательных характеристик.

Для навигации по заданной траектории в изначально неизвестном окружении с возможными препятствиями на данный момент успешно реализуются алгоритмы LSD SLAM с использованием одной камеры, которые отслеживают окружающее пространство и прогнозируют траекторию и при совпадении маршрута с координатами препятствия вносит поправки или производит информирование о невозможности альтернативного пути (например, закрытая дверь, забор и пр.).

В задачи машинного зрения так же входят: распознавание объектов, идентификация (индивидуального экземпляра), распознавание текста, оценка траектории движения, чтение штрих-кодов.

Сам прототип представляет из себя техническое устройство, состоящее из следующих компонентов: цифровая умная камера, включающая в себя: программное обеспечение обработки изображения, процессор, программное обеспечение машинного зрения, дальномер (для калибровки расстояний с камеры), специальные звуковые динамики.

При консультации со слепыми людьми было установлено, что для нормальной ориентации слепого человека в пространстве, он должен непрерывно оценивать окружающую обстановку при помощи звуков. Следовательно, его уши должны быть свободны. Использование закрытых наушников или наушников-капель не представляется возможным. Еще одним ограничением использования стандартных наушников является то, что с ними невозможно полноценно сформировать сцену с необходимой глубиной и шириной. С учетом данных ограничений была разработана конструктивная схема наушников с вынесенными на расстояние около 5 см вперед динамиками. Данная реализация оставляет свободными уши для восприятия окружающих звуков и дает возможность использовать аудио канал для информирования слепого человека об окружении, создавая сцену с достаточной глубиной.

Выводы. Проведя подобную эмуляцию, можно с уверенностью сказать, что предложенный способ использования наушников и цифровая обработка звука, переводящая пространственное положение объекта в аудио канал вполне точно описывает физическое положение предмета в пространстве и может быть использована для решения многих повседневных задач слепого человека.