

## **Раздел V. Мехатронные и робототехнические системы**

**В.А. Веселов, С.О. Ершов, О.С. Ипатов**

### **ПОВЫШЕНИЕ ЭФФЕКТИВНОСТИ СРЕДСТВ ПРОСТРАНСТВЕННОГО ОРИЕНТИРОВАНИЯ В ЧЕЛОВЕКО-МАШИННЫХ СИСТЕМАХ**

Активное развитие вычислительной техники и информационных технологий не только открывает перед исследователями новые возможности в области построения искусственных интеллектуальных систем, но также ставит новые вопросы о взаимодействии технических систем с человеком. Существует целый ряд задач, когда управление тем или иным объектом или процессом не может быть полностью доверено автоматике. Например, такая ситуация складывается при управлении сложными транспортными средствами в недетерминированной обстановке. Тем более нельзя игнорировать особенности человеческого восприятия сигналов при построении технических средств очувствления (ТСО) для слепых. Особенно важным становится учет человеческого фактора при создании информационных систем, действующих в режиме реального времени, поскольку самым «узким» местом при передаче информационного потока оказываются человеческие органы чувств.

В данной статье затронуты некоторые проблемы построения высокоинформативных ТСО для пространственного ориентирования инвалидов по зрению, когда проблема обеспечения информационного потока еще более осложняется отсутствием наиболее мощного в информационном отношении зрительного восприятия. В случае сложных комплексов, управляемых зрячим оператором, рассмотренные ниже системы можно рассматривать как дополнительный информационный канал, позволяющий повысить осведомленность человека об окружающей обстановке и(или) состоянии управляемого объекта.

#### **Анализ достижений в рассматриваемой области**

История активного развития электронных ТСО для слепых насчитывает около полувека. Наиболее популярными путями получения первичной пространственной информации первое время являлись ультразвуковая (УЗ) и световая (в т.ч. лазерная) локация. Но в последнее десятилетие наметился заметный рост интереса к телевизионным системам, вызванный тем, что новые относительно недорогие вычислительные средства позволяют осуществлять в реальном времени обработку видеоизображения и преобразование его в звуковые или тактильные сигналы, доступные восприятию слепого. По объему получаемой информации эти системы сегодня вне конкуренции. Проблемы, связанные с их миниатюризацией и удешевлением, - дело недалекого будущего.

Если в XX веке большая часть исследований по рассматриваемой тематике велась относительно обособленно, то сейчас идет процесс общего осмысления опыта, накопленного в мировой практике, и поиск оптимальных путей дальнейшего развития на стыках различных дисциплин. Но серьезной систематизации материала на данный момент еще не проведено. В данной работе предпринята очередная попытка обобщения и анализа экспериментального материала, хотя бы в несколько ограниченной области. Сразу оговоримся, что ТСО, подразумевающие хирургические вмешательства в человеческий организм (имплантируемые стимуляторы сетчатки глаза или зрительного кортекса мозга), рассматриваться не будут.

Основной вывод, к которому приходят сегодня исследователи, заключается в том, что компромиссные решения следует искать в комплексном подходе, используя в едином ТСО различные физические принципы получения пространственной информации. Добавим, что комплексный подход необходим и в плане способов представления информации человеку. Но сначала попытаемся хотя бы частично очертить зоны повышенного интереса в обширном поле накопленного в мировой практике материала.

- Первый аспект – получение пространственной информации.

Как уже было отмечено, растет рейтинг телевизионных систем, причем на данный момент развиваются как однокамерные, так и стереоскопические системы. В таблице приведены краткие характеристики некоторых телевизионных ТСО, позволяющие представить основные направления развития в данной области на сегодня. Основное ограничение телевизионных ТСО – зависимость от уровня освещенности, в частности, проблема засветки солнечным излучением. Еще одна проблема – игнорирование прозрачных (стеклянных) объектов. Забегая вперед, отметим, что ультразвуковые локаторы свободны от указанных недостатков, хотя имеют низкую разрешающую способность по угловым координатам и меньшую дальность действия. Выигрыш от объединения двух путей выглядит весьма перспективным.

*Краткие характеристики некоторых телевизионных ТСО*

Телевизионные ТСО	Автор	Способ представления информации
Однокамерные	Paul Bach-y-Rita, University of Wisconsin Medical School's orthopedics and rehabilitation department	Матрица тактильных стимуляторов на спине или матрица электро-стимуляторов на языке
	Peter Meijer, Netherlands	Стереофонические много-тональные сигналы
	Tadayoshi Shioyama, Department of Mathematics and Applied Physics in Kyoto University	Речевые команды и сообщения
Двухкамерные	J.L.Gonzalez Mora, Universidad de La Laguna, Spain	Система объемного звучания, головные телефоны
	Yoshihiro Kawai, Tsukuba Electro-technical Laboratory, Japan	Объемные акустические образы, использование звукопроводности костей черепа

- Второй аспект – обработка исходных данных и построение модели среды. Этот процесс во многом определяется тем, каким образом информация будет поставляться пользователю.

В концепциях, предусматривающих максимальное использование ресурсов человеческого мозга, задача сводится к преобразованию визуальной информации в иную модальность (звуковые или тактильные образы) с минимальными потерями и искажениями, т.е. речь идет о передаче «сырой», минимально обработанной информации. Такую направленность имеют работы Р.Meijer и Р.Bach-y-Rita.

В стереоскопических системах акцент делается на построении 3-мерной (3D) модели среды, что требует гораздо больших вычислительных затрат и высокой произво-

длительности процессорной системы. При таком подходе неизбежно возникает вторая задача: преобразование полученной 3D-модели в невизуальные 3D-образы. В проектах J.L.Gonzalez Mora и Y.Kawai [1] решения сходны: формируются виртуальные акустические объемные образы, в которых 3D-объекты представлены в виде совокупности точечных источников звука, покрывающих воображаемые поверхности объектов, причем «точки», относящиеся к различным объектам, звучат по-разному, чтобы человек мог их слышать одновременно и различать направления на источники звука.

Последняя из рассматриваемых концепций имеет ориентацию на высокоорганизованную городскую инфраструктуру. Речь идет об исследованиях T.Shioyama, который, анализируя видеообразы от одиночной камеры, распознает, например, наличие пешеходных переходов, заранее точно зная размеры и форму стандартной дорожной разметки (белых полос на асфальте), а затем геометрически вычисляет ширину дороги в этом месте. Производится также распознавание сигналов светофора. Круг задач, решаемых таким образом, очень узок, но ценность подхода в том, что он представляет собой попытку организации распознавания образов в реальном времени при движении пользователя с естественной скоростью.

Во всех рассмотренных подходах первичная информация от видеокамеры оказывается явно избыточной с точки зрения пропускной способности невизуальных каналов восприятия. Попытки полностью предоставить задачу фильтрации человеческого органам чувств и мозгу (как это делалось в первых ТСО P.Bach-y-Rita) приводили к излишним потерям полезных сведений о среде. При распознавании образов полностью отсекается вся информация, не соответствующая имеющейся базе данных (шаблонов). Компромиссным путем можно считать выделение контуров изображений локальных объектов и передача их пользователю в более контрастном виде. Такие решения применяются в той или иной мере во всех ТСО, рассмотренных в таблице, кроме системы T.Shioyama.

- Третий аспект – представление полученной информации слепому через осязание и слух.

Тактильные матрицы, используемые P.Bach-y-Rita, могут иметь конструкцию в виде набора игольчатых вибраторов, воздействующих на достаточно широкие участки кожи человека, или представлять собой небольшую пластину электрических стимуляторов, накладываемую на язык. Таким образом, 2-мерный видеообраз преобразуется в 2-мерный же тактильный. Более простая разновидность тактильных датчиков – инерционные вибраторы, кодирующие несложную информацию посредством различной интенсивности и переменных ритмов вибраций.

Более эффективен в отношении пропускной способности слуховой канал восприятия. Возможны два принципиально различных подхода: использование многотональных звуковых сигналов или речевых сообщений.

В системе P. Meijer [2] основной объем информации передается посредством полифонических образов. Имеет смысл рассмотреть используемый здесь принцип преобразования подробнее, поскольку на текущий момент система P. Meijer, названная «VOICE», признается многими специалистами одной из наиболее эффективных. Видеокадр разбивается на вертикальные столбцы. Каждый столбец преобразуется в сочетание (сумму) звуковых сигналов, частота каждого из которых тем выше, чем выше расположен соответствующий сегмент (пиксел) видеоизображения, а амплитуда зависит от яркости соответствующего пиксела, то есть угол возвышения кодируется частотой, а яркость – громкостью сигнала. Причем одновременно «озвучивается» только один столбец. Происходит сканирование видеокadra (с периодом около 1 секунды) слева направо и последовательное «звучание» столбцов. В начале каждого кадра формируется короткий привязочный звуковой импульс (щелчок), кроме того, стереофонический сигнал формируется так, чтобы

кажущееся направление на источник звука совпадало с действительным направлением на соответствующий участок поля обзора. На рис.1 схематично показано формирование сигнала для 1-го столбца видеокadra, разбитого на 8 столбцов и 8 строк и имеющего 3 градации яркости.

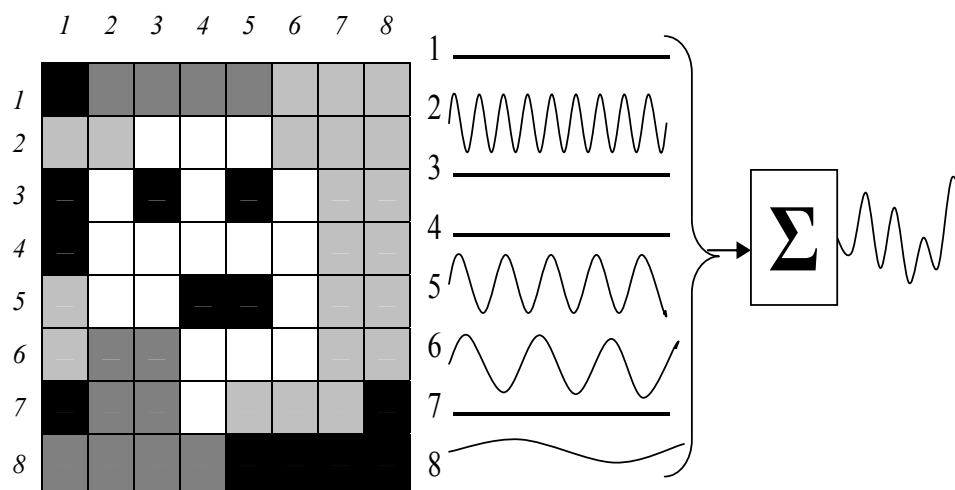


Рис.1. Упрощенная диаграмма построения звукового образа в системе «VOICE» (для 1-го столбца)

Речевые сообщения в ТСО «VOICE» носят лишь вспомогательный характер, например при сообщении (по запросу пользователя) о преобладающем цвете объекта. Интересен используемый в «VOICE» способ управления режимами представления информации: используется система распознавания речевых команд пользователя, поступающих через микрофон.

Формирование звуковых образов в ТСО J.L.Gonzalez Mora и Y.Kawai осуществляется посредством систем объемного звучания, индивидуально настраиваемых под особенности восприятия конкретного человека. Различие заключается в конструкции акустических систем. В ТСО J.L.Gonzalez Mora специальные головные телефоны в значительной степени блокируют восприятие обычных звуков внешнего мира, а Y.Kawai применяет стереофоническую систему, использующую звукопроводность костей черепа и оставляющую уши пользователя свободными. Общий недостаток обоих ТСО – в низкой надежности определения координат объектов пользователем: даже в упрощенных лабораторных экспериментах при оценке направления на одиночный виртуальный источник звука процент ошибок очень высок.

Речь в качестве основного способа передачи данных наиболее целесообразна в системах с распознаванием образов, именно этим путем идет T.Shioyama.

Использование полифонических образов имеет преимущество перед речевым представлением в смысле повышения плотности информационного потока, поскольку эти образы могут синтезироваться с рациональным использованием полного звукового диапазона с равномерным распределением информационной нагрузки по всему спектру. Кроме того, дополнительную информацию здесь несет и уровень громкости сигнала. Речь же использует частотный диапазон неравномерно, поэтому даже при максимально ускоренном воспроизведении сообщений общее количество звуковой информации меньше. Скомпенсировать этот недостаток можно лишь за счет разработки очень эффективного, емкого по содержанию, но компактного специального

словаря пространственных терминов, характеризующих объект и его главные признаки. А для реализации такого словаря, даже если бы он был составлен, необходим мощный быстродействующий аппарат распознавания образов. Ожидать эффективного применения такого подхода можно, вероятно, лишь в условиях хорошо организованной инфраструктуры, прежде всего, в городских условиях, когда высока концентрация упорядоченных и стандартных объектов. В недетерминированной среде ТСО, поставляющие «сырую», не переработанную информацию (звуковую или тактильную), обладают более значительным потенциалом.

- Четвертый аспект – обучение пользователя.

Сегодня многие университеты мира занимаются изучением пластичности мозга при обучении человека интерпретации искусственных невизуальных образов среды. Установлено, что в результате постоянной практики мозг адаптируется к восприятию этой информации. Более того, выдвинуто принципиально важное утверждение, что мозгу для восприятия и обработки пространственных образов в конечном итоге неважно, по какому из каналов, зрительному, слуховому или тактильному, поступила информация. Речь, разумеется, идет о человеке, прошедшем курс специального обучения. Сам процесс адаптации мозга при этом во многом аналогичен тому, который имеет место при освоении зрения новорожденными детьми. Практический вывод, который можно сделать на основании этого открытия, следующий: наилучшим образом проблему замещения зрения можно решить при обеспечении максимально достижимого информационного потока через невизуальные каналы восприятия человека, причем достаточно длительное обучение является неотъемлемой частью вопроса.

### Структура комплексной системы технического зрения

На основании представленного выше анализа предлагается комплексный подход к задаче построения высокоинформативного ТСО. Обобщенная структурная схема системы приведена на рис.2.

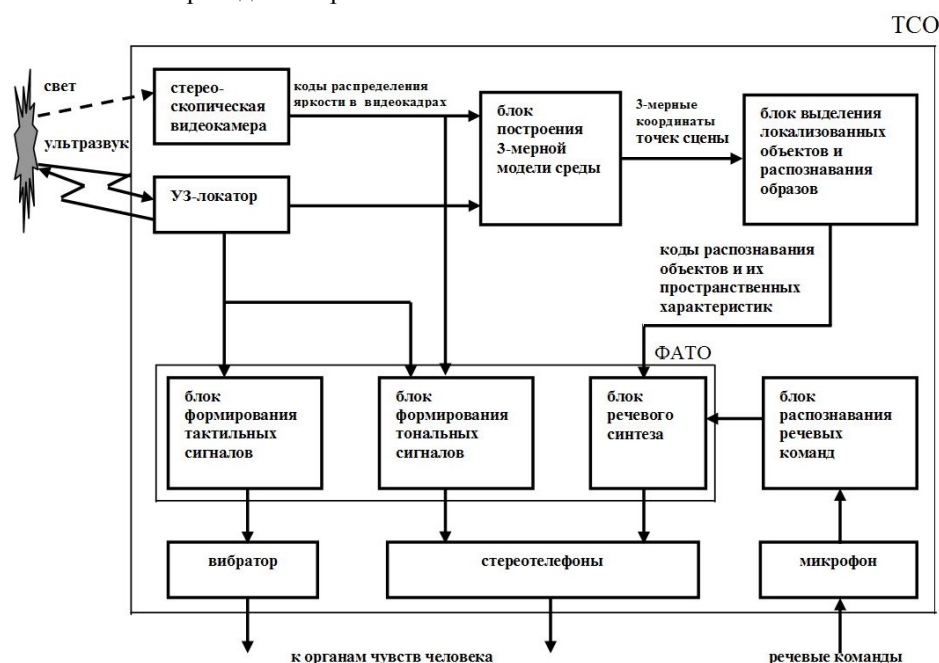


Рис.2. Блок-схема высокоинформативного ТСО

В части получения исходной пространственной информации необходимо обеспечить высокую разрешающую способность по углам и дальности, а также максимальную надежность обнаружения объектов. С этой целью одновременно применена стереоскопическая видеокамера и ультразвуковой локатор, отвечающий за уточнение расстояния до ближайшего объекта и формирование экстренных сигналов об опасном сближении с препятствием.

В части обработки видеообразов исходные данные направляются по двум параллельным потокам:

- построение 3-мерной модели среды с последующим распознаванием образов;
- прямое преобразование 2-мерных видеообразов в звуковые «кадры» с использованием концепции Р.Мејџет (в блоке формирования тональных сигналов).

УЗ-локатор одновременно выполняет функции уточнения координат дальности для построения 3D-модели и предоставляет информацию для выработки тактильных или акустических сигналов тревоги в формирователь аудио-тактильных образов (ФАТО).

В части представления информации пользователю акцент в данной системе сделан на акустический канал (речевые сообщения и полифонические образы). Тактильный датчик (вибратор) применен лишь как канал экстренного предупреждения об опасности. Степень приближения к объекту может кодироваться, например, интенсивностью вибраций.

Приоритетами при формировании аудио-тактильных образов управляет блок распознавания речевых команд, получающий инструкции от человека через микрофон. Должны быть предусмотрены несколько режимов работы ТСО. Например, если по умолчанию система вырабатывает полифонические образы, то для получения речевых сообщений пользователь должен подать голосом соответствующую команду. Вибратор может иметь несколько режимов, вплоть до отключения по предпочтению пользователя на текущий момент. Скорость подачи информации человеку, уровень громкости и другие параметры выходных сигналов ТСО тоже мобильно управляются пользователем. Для этого он должен предварительно выработать удобный ему словарь кратких инструкций и занести их в память блока распознавания речевых команд. Наиболее эффективное распознавание речи обеспечивается благодаря процессу «обучения» распознающего аппарата. Причем результат, очевидно, будет наилучшим при настройке на голос и особенности дикции конкретного пользователя. Таким образом, функция «обучаемости» рассматриваемого блока должна быть заложена в данном ТСО.

Высокий уровень сложности задач, параллельно решаемых в данной системе, и требование функционирования в реальном времени приводят к необходимости реализации ТСО в виде многопроцессорной системы. Количество процессоров и распределение задач между ними в данной работе не рассматривается. Для переносного ТСО, удобного при использовании слепыми в повседневной практике, необходима минимизация габаритов и веса устройства. Поэтому возникает задача построения специального микрокомпьютера. В плане конструкции данное ТСО целесообразно размещать на голове человека в виде специальных очков, налобной повязки или головного убора. В этом случае камеры, телефоны, микрофон и тактильный датчик могут быть размещены в едином конструктиве.

### Заключение

Широкий спектр научно-исследовательских задач, возникающих при разработке рассматриваемого ТСО, может быть решен лишь совместными усилиями нескольких коллективов. В СПб БГТУ «Военмех» им. Д.Ф.Устинова исследования по данному направлению ведутся уже 15 лет (начало было положено в 1991-1992гг., когда под руководством В.А.Веселова и В.Г.Кузнецова была выполнена НИР «Создание прибора очки для слепых»). В апреле 2005г команда БГТУ выиграла грант по проекту «Комплексная разработка индивидуальных технических

средств ориентирования для слепых». На данный момент, помимо БГТУ, в работе по реализации изложенных выше идей готовы участвовать ГНУ ЦНИИ РТК, СПб Центр медико-социальной реабилитации для инвалидов по зрению и некоторые европейские ученые, в том числе Р.Meijer и представители Технического университета Твенте (Голландия).

#### БИБЛИОГРАФИЧЕСКИЙ СПИСОК

1. Kawai Y., Kobayashi M., Minagawa H., Miyakawa M., Tomita F. A Support System for Visually Impaired Persons Using Three-Dimensional Virtual Sound // *Int. Conf. Computers Helping People with Special Needs* (ICCHP 2000), Karlsruhe, Germany, July 17-21, 2000, pp. 327-334.
2. Meijer P.B.L. An Experimental System for Auditory Image Representations // *IEEE Transactions on Biomedical Engineering*, Vol. 39, No. 2, Feb 1992, pp. 112-121.

**Т. Акинфиев, Р. Фернандес, М. Армада**

#### **ПРИВОД С ДВОЙНЫМ ИЗМЕНЕНИЕМ ПЕРЕДАТОЧНОГО ОТНОШЕНИЯ И УПРАВЛЕНИЕ ЭТИМ ПРИВОДОМ**

В статье обсуждается интеллектуальный привод с непрерывно меняющимся передаточным отношением и двойными свойствами. Такой тип привода особенно эффективен при использовании в машинах, рабочие элементы которых совершают старт-стопные движения, причем движения вперед и назад характеризуются различными величинами масс (моментов инерции) подвижных частей, как например, в промышленных или в шагающих роботах. Обсуждаются различные алгоритмы управления указанным приводом, в том числе универсальный алгоритм адаптивного управления. Был разработан, изготовлен и испытан специальный прототип привода. Прототип робота с таким приводом находится в стадии изготовления. Сравнение с классическим приводом показывает высокую эффективность разработанного привода.

Различными авторами было показано [1-4], что использование приводов с плавно меняющимся передаточным отношением оказывается эффективным как в квази-статическом, так и в динамическом режимах. Так, в начале и конце движения такие приводы имеют высокое передаточное отношение, обеспечивая высокое ускорение рабочих элементов. В средней же части траектории эти приводы имеют низкое передаточное отношение, что позволяет достигать высоких скоростей перемещения. Важно отметить, что приводной мотор имеет практически постоянную скорость, что дает возможность получить высокий КПД. Все эти положительные эффекты проявляются при правильной настройке приводов. Однако в некоторых случаях не удастся осуществить такую настройку. Это относится к ситуации, когда один и тот же привод должен перемещать в одном направлении тело большой массы, а в другом направлении – тело малой массы. Так, например, в шагающих роботах один и тот же привод осуществляет сначала перемещение легкой ноги робота, а потом – тяжелого тела робота. В промышленных роботах один и тот же привод осуществляет перемещение руки робота с тяжелым грузом, а потом – движение руки робота без груза. В этих случаях оптимальная настройка на один режим движения оказывается далеко не оптимальной для другого режима движения.

#### **Цель работы**

Цель работы – повышение эффективности привода для таких старт-стопных движений, при которых один и тот же привод должен осуществлять два различных