

А.Л. РОНЖИН, А.А. КАРПОВ  
**СРАВНИТЕЛЬНЫЙ АНАЛИЗ ФУНКЦИОНАЛЬНОСТИ  
ПРОТОТИПОВ ИНТЕЛЛЕКТУАЛЬНЫХ ПРОСТРАНСТВ**

---

*Ронжин А.Л., Карпов А.А. Сравнительный анализ функциональности прототипов интеллектуальных пространств.*

**Аннотация.** В статье представлен обзор прототипов интеллектуальных пространств, предназначенных для поддержки проведения научно-образовательных мероприятий и оснащенных средствами автоматической записи выступлений участников. На основе анализа применяемых в них средств обработки аудиовизуальных сигналов и реализованных пользовательских сервисов была предложена оригинальная классификация прототипов интеллектуальных пространств. Также описана специфика разработанного интеллектуального зала и его отличия от рассмотренных прототипов.

**Ключевые слова:** интеллектуальное пространство, обработка аудиовизуальных сигналов, автоматизация записи выступлений.

*Ronzhin A.L., Karpov A.A. A comparative analysis of smart space prototypes functionality.*

**Abstract.** A survey of smart space prototypes intended for scientific and educational meetings and facilitated by means of automatic speech recording is represented in the paper. Analysis of the used audio visual signal processing means and realized user services allowed us to propose a classification of smart space prototypes. The peculiarities of the developed smart meeting room and its distinction features from the considered prototypes are described.

**Keywords:** smart space, audiovisual signal processing, automation of speech recording.

---

**1. Введение.** Применение интеллектуальных информационных технологий в сфере бизнеса и образования, в том числе, при проведении распределенных мероприятий и для автоматизации процесса записи выступлений на совещании приобретает важную роль в связи с повышением мобильности людей и необходимостью контроля качества принимаемых решений [4, 6, 12, 21]. Развитие научной парадигмы интеллектуального пространства сформировало на сегодняшний день несколько моделей интеллектуального окружения, обслуживающего пользователей, находящихся в некотором ограниченном пространстве: интеллектуальной комнате, доме, лекционной аудитории, зале совещаний [10, 12, 19]. Разработка средств захвата и обработки аудиовизуальных сигналов, позволяющих бесконтактно оценить текущую ситуацию в наблюдаемом помещении, является одним из основных направлений исследований в этой области.

При проектировании интеллектуальных залов для проведения совещаний, лекций и научно-образовательных мероприятий наиболее активно сейчас используются следующие методы аудиовидеобработки: 1) определение и слежение за перемещением пользователей на ос-

нове видеомониторинга [13]; 2) определение положения, распознавание и идентификация лиц [15]; 3) определение местоположения источника звука [7, 23]; 4) распознавание речи [5, 1]; 5) идентификация говорящего [17]; 6) синтез речи [2]. Применение перечисленных методов и их комбинации позволяет разработать средства для автоматической записи выступлений участников, организации телетрансляций, протоколирования и архивирования аудиовизуальной записи по окончании мероприятия.

В следующем разделе рассмотрено несколько прототипов интеллектуальных залов, описано примененное в них математическое и программно-аппаратное обеспечение. В третьем разделе предложена классификация интеллектуальных залов, учитывающая их оснащение и назначение. Четвертый раздел описывает специфику разработанного интеллектуального зала по сравнению с проанализированными прототипами.

**2. Сравнительный анализ прототипов интеллектуальных пространств.** При подготовке данного обзора были проанализированы научные работы последних лет, описывающие прототипы интеллектуальных пространств (в большинстве случаев это интеллектуальные залы совещаний), которые были разработаны и используются в ходе научно-исследовательских проектов в организациях, представленных в таблице 1. Рассмотрим функциональность перечисленных прототипов интеллектуальных залов, включающая в себя их оснащение, а также методы, необходимые для реализации сервисов информационной поддержки проведения и автоматизации хода мероприятия. Сенсорное оснащение интеллектуального зала может быть разделено на два типа: основное, записывающее события во всем пространстве зала, и вспомогательное – для более детального анализа определенных зон интереса, выявленных с помощью основного оборудования. В качестве сенсорного оборудования основного типа, применяемого в интеллектуальных залах, обычно используется потолочная панорамная камера и массив микрофонов, при помощи которых ведется аудиовизуальный мониторинг.

Вспомогательным оборудованием являются отдельный микрофон и камера с высоким разрешением, обслуживающие различные зоны в зале; камера с функциями наклона, поворота и масштабирования (PTZ), задействованная для наведения на лица конкретных участников в зале и записи их выступлений; встроенный микрофон персональной веб-камеры, установленной на конференц-столу и др. Каждый из рассмотренных типов оборудования указан в единственном экземе-

пляре, необходимом для проектирования минимальной конфигурации системы аудиовизуального мониторинга интеллектуального зала. С увеличением размера помещения, числа участников, числа зон наблюдения, требований, предъявляемых к протоколированию проводимых мероприятий, допускается возможность масштабирования программно-аппаратных ресурсов, позволяющая настроить весь комплекс оборудования под конкретные условия эксплуатации.

Таблица 1. Список организаций, исследующих прототипы интеллектуальных залов.

№	Организация	Город, Страна	Обозначение	URL
1.	Политехнический университет Каталонии	Барселона, Испания	UPC	<a href="https://imatge.upc.edu/web/?q=node/106">https://imatge.upc.edu/web/?q=node/106</a>
2.	Технический институт Тренто	Тренто, Италия	ITC	<a href="http://www.fbk.eu/">http://www.fbk.eu/</a>
3.	Научно-исследовательский институт человеческих факторов	Делфт, Нидерланды	TNO	<a href="http://www.amiproject.org/">http://www.amiproject.org/</a> , <a href="http://www.tno.nl">http://www.tno.nl</a>
4.	Центр информационных технологий университета Гонконга	Гонконг, Китай	ITSC	<a href="http://www.ust.hk/itsc/">http://www.ust.hk/itsc/</a>
5.	Исследовательский институт Idiap	Мартиньи, Швейцария	Idiap	<a href="http://www.idiap.ch/scientific-research/smart-meeting-room">http://www.idiap.ch/scientific-research/smart-meeting-room</a>
6.	Национальный институт стандартов и технологий США	Гейтерсберг, США	NIST	<a href="http://www.nist.gov/smartSPACE/">http://www.nist.gov/smartSPACE/</a>
7.	Петрозаводский государственный университет	Петрозаводск, Россия	ПетрГУ	<a href="http://petsu.ru/">http://petsu.ru/</a> , <a href="http://oss.fruct.org/wiki/SmartRoom">http://oss.fruct.org/wiki/SmartRoom</a>
8.	Санкт-Петербургский институт информатики и автоматизации Российской академии наук	Санкт-Петербург, Россия	СПИИРАН	<a href="http://www.spiiras.nw.ru/">http://www.spiiras.nw.ru/</a>

Перечень сенсорного и мультимедийного оборудования залов и применяемые методы обработки аудиовизуальных сигналов представлены в таблице 2. Большинство интеллектуальных залов оснащено аудио- и видеозаписывающим оборудованием, а также современными устройствами ввода/вывода мультимедийной информации.

Таблица 2. Перечень оборудования и методов, применяемых в прототипах интеллектуальных залов.

№	Зал	Оборудование	Методы
1.	UPC	6 стационарных видеокамер; PTZ камера с функциями наклона, поворота и масштабирования; панорамная потолочная камера; 4 отдельных микрофона; 3 массива микрофонов Т-образной формы; массив микрофонов типа Mark III.	1) распознавание лиц; 2) идентификация лиц; 3) обнаружение объектов; 4) аудиовизуальное определение и слежение за активностью участников мероприятий; 5) определение поз участников мероприятий
2.	ITC	4 стационарные камеры; PTZ камера; 4 отдельных микрофона; 7 массивов микрофонов Т-образной формы; Массив микрофонов типа Mark III; проектор; проекционный экран.	1) орфографическое транскрибирование; 2) трехмерное слежение за перемещением пользователей интеллектуального зала; 3) определение лиц пользователей; 4) идентификация пользователей; 5) оценка положения и ориентации головы; 6) определение речевой активности и распознавание речи; 7) определение неречевых акустических событий.
3.	TNO	6 стационарных видеокамер; 2 круговых массива микрофонов; проектор; проекционный экран.	1) идентификация текстовых заметок на доске; 2) определение и аннотирование активности участников мероприятий; 3) распознавание жестов; 4) распознавание эмоций участников.
4.	ITSC	стационарная камера; проектор; проекционный экран.	1) определение и слежение за изменениями на проекционном экране; 2) определение и слежение за перемещением лектора; 3) определение лица лектора.
5.	Idiap	7 стационарных видеокамер; 2 круговых массива микрофонов; проектор; проекционный экран.	1) идентификация текстовых заметок на доске; 2) определение и аннотирование активности участников мероприятий; 3) распознавание жестов; 4) распознавание эмоций участников.

№	Зал	Оборудование	Методы
6.	NIST	5 стационарных камер; 4 отдельных микрофона; 3 массива микрофонов Т-образной формы; массив микрофонов типа Mark III; проектор; проекционный экран; электронная доска для заметок.	1) распознавание речи; 2) идентификация говорящего участника; 3) распознавание лиц участников; 4) определение и разделение источников звука; 5) нормализация аудиоканала.
7.	Петргу	проектор; проекционный экран; сенсорная доска, климатические датчики	1) определение клиентских устройств участников мероприятия; 2) персонализированное обслуживание участников мероприятия.
8.	СПИИРАН	2 стационарные камеры; 2 PTZ камеры; панорамная потолочная камера; 10 персональных веб-камер со встроенными микрофонами; 2 отдельных микрофона; 3 массива микрофонов Т-образной формы; плазменная панель с сенсорной насадкой; проектор; проекционный экран.	1) определение речевой активности и распознавание речи; 2) определение неречевых акустических событий; 3) идентификация говорящего участника; 4) определение положения источника звука; 5) определение и слежение за перемещением участников; 6) определение лиц участников; 7) определение поз участников.

Для обработки аудиовизуальных сигналов, записываемых оборудованием, расположенным внутри зала, применяются методы анализа изображений (распознавание и идентификация лиц, определение и слежение за участниками, определение их поз), анализа звука (определение речевой активности, определение и разделение источников звука, идентификация говорящего участника, распознавание речи и неречевых акустических событий), комбинированные методы (оценивание положения и ориентации головы выступающего, многомодальное распознавание эмоций). Теперь на примерах рассмотрим оснащение некоторых залов.

В университетах UPC, ITS и NIST [16, 20] установлено два различных типа видеокамер (стационарная и PTZ видеокамера). В интеллектуальном зале UPC дополнительно установлена панорамная потолочная камера. Кроме того, перечисленные залы оснащены персональными микрофонами и двумя различными массивами микрофонов с Т-

образной формой и типа «Mark III». В университетах TNO, Idiap интеллектуальные залы оборудованы стационарными видеокамерами и круговыми массивами микрофонов, применяются методы идентификации текстовых заметок на доске, определения и аннотирования деятельности участников мероприятий, распознавания жестов и эмоций участников мероприятий [20, 21]. В интеллектуальном зале ITSC, предназначенном для проведения образовательных мероприятий, установлена только одна стационарная камера, реализованы методы определения и слежения за изменениями на проекционном экране, определения и слежения за перемещением лектора и определения его лица [14]. Прототип интеллектуального зала Петрозаводского государственного университета разрабатывается с 2012 года, на данный момент зал оснащен презентационным, аудиовидеозаписывающим оборудованием, климатическими датчиками, в нем реализованы методы персонализированного обслуживания участников научно-образовательных мероприятий, использующих свои мобильные устройства для управления презентационным оборудованием [3, 18].

Интеллектуальный зал СПИИРАН предназначен для проведения научных и образовательных мероприятий, его оснащение и применяемые методы обработки аудиовизуальных данных кратко описаны в следующем разделе и представлены в работах [6, 8, 11].

**3. Классификация прототипов интеллектуальных пространств.** По результатам анализа характеристик приведенных выше прототипов была предложена классификация интеллектуальных залов, учитывающая следующие аспекты: тип проводимого в зале мероприятия; число участников, которое вмещает зал; сенсорное и мультимедийное оборудование, встроенное в зал; пользовательские сервисы, реализованные в зале. На рисунке 1 представлена структура классификации интеллектуальных залов.



Рис. 1. Классификация прототипов интеллектуальных залов

В зависимости от назначения и готовности зала можно выделить четыре типа проводимых в зале мероприятий: исследовательское совещание, рабочее совещание, лекция, конференция. Первый тип мероприятия имеет заранее известный для участников сценарий и применяется для создания тестовых баз данных и экспериментальной проверки разработанных методов и программных средств. Рабочее совещание в рассматриваемых залах обычно проводится за круглым столом с небольшим числом участников. К третьему типу относятся, как правило, учебные занятия, в которых заранее определен один выступающий. Четвертый тип – наиболее сложный с точки зрения оценки аудиовизуального мониторинга ситуации в зале, так как число выступающих, слушателей и удаленных участников может варьироваться.

В работе [22] введена классификация залов по числу участников, где залы, вмещающие до двадцати участников определены как малые, средние – около 50 человек и большие – с числом участников свыше 100. Большинство рассмотренных интеллектуальных залов вмещают менее десяти человек (UPC, ИТС, TNO, Idiap, NIST), к средним по вместимости следует отнести залы в ИТС, ПетрГУ, СПИИРАН.

При оснащении интеллектуальных залов используются вычислительные, сетевые ресурсы, активационное оборудование для управления силовыми приборами и средствами жизнеобеспечения. Кроме того, обязательным условием является наличие проектора и другой

мультимедийной презентационной техники. В предложенной классификации акцент сделан на сенсорное оборудование, необходимое для аудиовизуального мониторинга, то есть в первую очередь отдельные микрофоны и видеокamеры. Также применяются массивы микрофонов, обеспечивающие локализацию источника звука, и персональные камеры со встроенными микрофонами для записи речи конкретного участника, расположенного за конференц-столом.

Также интеллектуальные залы можно классифицировать по реализованным в них пользовательским сервисам. На рисунке 1 показаны десять основных типов сервисов, которые чаще всего применяются для автоматизации проведения мероприятий и обслуживания участников, присутствующих в зале. Кроме того, актуальны сервисы, обеспечивающие интеграцию клиентских устройств с презентационным оборудованием интеллектуального зала, для участников, расположенных как внутри зала, так и за его пределами.

**4. Сервисы интеллектуального зала СПИИРАН.** При проектировании интеллектуального зала были учтены эргономические аспекты расположения мультимедийного, аудио- и видеозаписывающего оборудования так, чтобы обеспечить охват и обслуживание наибольшего числа участников. В состав сенсорного оборудования зала входят четыре типа видеокамер (две стационарные, две PTZ, одна панорамная потолочная камера, десять персональных веб-камер) два типа аудиозаписывающего оборудования (два отдельных потолочных микрофона и три массива микрофонов T-образной формы) [9]. При реализации системы мониторинга участников интеллектуального зала на основе распределенной обработки аудиовизуальных сигналов были использованы как существующие методы обработки цифровых данных (сегментации изображений, вычисление и сравнение гистограмм распределения цветов и т.д.), так и разработанные собственные оригинальные методы, в частности, метод регистрации участников мероприятий, метод аудиовизуальной записи их выступлений [9, 11].

Метод автоматизации регистрации участников мероприятия за счет применения многоэтапной процедуры бимодального анализа зон зала позволяет найти присутствующих в нем участников и получить аудиовизуальные данные, необходимые для идентификации присутствующих. Метод аудиовизуальной записи выступлений участников в интеллектуальном зале использует бимодальный (по аудио и видео) подход к определению активного диктора, обеспечивающий своевременный автоматический захват и синхронность аудио- и видеопотоков данных выступающего в зале участника. На основе предложенных ме-

тодов был создан комплекс программных средств аудиовизуального мониторинга и определения ситуации в интеллектуальном зале, обеспечивающий обработку аудиовизуальных потоков данных от набора сетевых камер и массивов микрофонов для определения и слежения за перемещением участников, их регистрацию и запись выступлений в ходе мероприятия в интеллектуальном зале. В таблице 3 приведено сравнение разработанного зала и рассмотренных выше прототипов.

Таблица 3. Реализованные сервисы в прототипах интеллектуальных залах.

Прототипы интеллектуальных залов	Сервисы, реализованные в интеллектуальных залах									
	Интернет трансляция мероприятий	Создание тестовых аудиовизуальных баз данных	Подключение удаленных участников мероприятий	Автоматическое формирование мультимедийного отчета	Протоколирование хода мероприятия	Идентификация говорящего участника	Интеграция устройств участников с системой сервисов интеллектуального зала	Удаленное управление оборудованием зала	Автоматическая регистрация участников	Автоматическая запись выступлений участников
UPC	+	+	+		+			+		
ITC	+	+	+		+	+	+	+		
TNO	+	+	+		+					
ITSC	+			+	+					+
Idiap	+	+		+	+					
NIST	+	+	+		+	+				+
ПетрГУ	+		+				+	+		
СПИИРАН	+	+	+	+	+	+	+	+	+	+

В таблице 3 представлены основные сервисы, необходимые для проведения и поддержки различного типа мероприятий, записи и трансляции мероприятий, автоматического формирования мультимедийного отчета о прошедшем событии, обеспечивающие возможность взаимодействия пользователей на расстоянии и предоставления записанных мультимедийных данных.

Анализ данных, представленных в таблице 3, показывает, что сервисы записи, трансляции, протоколирования мероприятий, а также создание тестовых аудиовизуальных баз данных реализованы почти во всех рассматриваемых интеллектуальных залах. Однако сервисы, служащие для автоматизации хода проведения мероприятий, такие как «Автоматическая запись выступлений участников», «Идентификация говорящего участника», «Интеграция устройств участников с системой сервисов интеллектуального зала» и «Автоматическое формирование мультимедийного отчета» реализовано лишь в половине из перечисленных интеллектуальных залов, а сервис «Автоматическая регистрация участников» был апробирован только в разработанном зале.

**5. Заключение.** Разработка интеллектуальных залов совещаний, а также реализация сервисов поддержки проведения мероприятий на основе естественного и ненавязчивого способа взаимодействия между пользователями и интеллектуальным пространством является предметом исследований ряда текущих российских и зарубежных проектов. Анализ существующих прототипов показал, что целью большинства реализованных средств интеллектуальных залов является накопление исследовательских данных об активности пользователей и ситуации в помещении, которые далее могут применяться для анализа деятельности участников мероприятий и формирования моделей автоматического управления информационными и техническими системами.

### Литература

1. *Кипяткова И.С., Карпов А.А.* Аналитический обзор систем распознавания русской речи с большим словарем // Труды СПИИРАН. 2010. Вып. 12. С. 7-20.
2. *Лобанов Б.М., Цирульник Л.И., Железны М., Крноул З., Ронжин А., Карпов А.* Система аудиовизуального синтеза русской речи // Информатика. Минск, Беларусь, 2008. № 4 (20). С. 67-78.
3. *Ломов А.А., Корицун Д.Ж.* Операция подписи для приложений в интеллектуальных пространствах платформы Smart-M3 // Труды СПИИРАН. 2012. Вып. 23. С. 439-458.
4. *Мещеряков Р.В., Балацкая Л.Н., Чойнзонов Е.Л.* Специализированная информационная система поддержки деятельности медицинского учреждения // Информационно-управляющие системы. 2012. № 5. С. 51-56.
5. *Ронжин А.Л.* Топологические особенности морфофонемного способа представления словаря для распознавания русской речи // Вестник компьютерных и информационных технологий, № 9, 2008, С. 12-19.
6. *Ронжин А.Л., Будюков В.Ю., Ронжин Ал.Л.* Технологии формирования аудиовизуального интерфейса системы телеконференций // Автоматизация и современные технологии. № 5. 2011, С. 20-26.
7. *Ронжин А.Л., Карпов А.А.* Сравнение методов локализации пользователя многомодальной системы по его речи // Известия вузов. Приборостроение. 2008. Т. 51, № 11. С. 41-47.

8. *Ронжин Ал.Л., Будков В.Ю., Ронжин Ан.Л.* Формирование профиля пользователя на основе аудиовизуального анализа ситуации в интеллектуальном зале совещаний // Труды СПИИРАН. 2012. Вып. 23. С. 482-494.
9. *Ронжин Ал.Л., Ронжин Ан.Л.* Система аудиовизуального мониторинга участников совещания в интеллектуальном зале // Доклады ТУСУРа, № 1 (22), ч. 1, 2011, С. 153-157.
10. *Юсупов Р.М., Ронжин А.Л.* От умных приборов к интеллектуальному пространству // Вестник Российской Академии Наук: научный и общественно-политический журнал, Т. 80, Вып. 1, 2010, С. 45-51.
11. *Юсупов Р.М., Ронжин А.Л., Прищепина М.В., Ронжин Ал.Л.* Модели и программно-аппаратные решения автоматизированного управления интеллектуальным залом // Автоматика и телемеханика, 2011, № 7. С. 39-49.
12. *Aldrich F.* Smart Homes: Past, Present and Future / Inside the Smart Home // Ed. Harper R. London: Springer-Verlag, 2003. pp. 17-39.
13. *Calonder M., Lepetit V., Fua P.* BRIEF: Binary Robust Independent Elementary Features. Computer Vision – ECCV'10. 2010. pp. 778-792.
14. *Chou H., Wang J., Fuh C., Lin S., Chen S.* Automated lecture recording system // International Conference on System Science and Engineering, 2010. pp. 167-172.
15. *Ekenell H. K., Fischer M., Jin Q., Stiefelhagen R.* Multi-modal Person Identification in a Smart Environment // Proc. of the Computer Vision and Pattern Recognition, CVPR '07. 2007. pp. 1-8.
16. *Fillinger A., Hamchi I., Degré S., Diduch L., Rose T., Fiscus J., Stanford V.* Middleware and Metrology for the Pervasive Future // IEEE Pervasive Computing Mobile and Ubiquitous Systems. Vol. 8, num. 3, 2009. pp. 74-83,
17. *Imseng D.; Friedland G.* Tuning-Robust Initialization Methods for Speaker Diarization // IEEE Transactions on Audio, Speech, and Language Processing, vol.18, no.8, 2010. pp. 2028-2037.
18. *Kadirov R., Cvetkov E., Korzun D.* Sensors in a Smart Room: Preliminary Study // Proc. 12th Conf. Open Innovations Framework Program FRUCT, Oulu, Finland, 2012. pp. 37-42.
19. *Lampi F.* Automatic Lecture Recording. Dissertation. The University of Mannheim, Germany. 2010. 229 p.
20. *Nakashima H., Aghajan H. K., Augusto J. C.* Handbook of Ambient Intelligence and Smart Environments // Springer. 2010. 1294 p.
21. *Renals S., Hain T., Boulard H.* Recognition and Understanding of Meetings The AMI and AMIDA Projects // Proc. of the IEEE Workshop on Automatic Speech Recognition and Understanding, ASRU'07, 2007. pp. 238-247.
22. *Rui Y., Gupta A., Grudin J., He L.* Automating lecture capture and broadcast: Technology and videography // Multimedia Systems, Vol. 10, 2004, pp. 3–15.
23. *Zhang C., Yin P., Rui Y., Cutler R., Viola P., Sun X., Pinto N., Zhang Z.* Boosting-Based Multimodal Speaker Detection for Distributed Meeting Videos // IEEE Trans. on Multimedia, Vol.10, No.8, 2008, pp. 1541-1552.

**Ронжин Александр Леонидович** — младший научный сотрудник лаборатории речевых и многомодальных интерфейсов СПИИРАН. Область научных интересов: технологии интеллектуального пространства, аудиолокализации, техническое зрение. Число научных публикаций — 30. [ronzhinal@iias.spb.su](mailto:ronzhinal@iias.spb.su); СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; п.т. +7(812)328-7081, факс +7(812)328-7081.

**Ronzhin Alexander Leonidovich** — junior researcher, Laboratory of Speech and Multimodal Interfaces, SPIIRAS. Research interests: smart space, sound source localization, computer vision. The number of publications — 30. ronzhinal@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-7081, fax +7(812)328-7081.

**Карпов Алексей Анатольевич** — канд. техн. наук, доцент, старший научный сотрудник лаборатории речевых и многомодальных интерфейсов СПИИРАН. Область научных интересов: автоматическое распознавание речи, многомодальные интерфейсы, аудиовизуальное распознавание и синтез речи. Число научных публикаций — 150. karpov@iias.spb.su; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-7081, факс +7(812)328-7081.

**Karpov Alexey Anatolyevich** — PhD, associative professor, senior researcher, Laboratory of Speech and Multimodal Interfaces, SPIIRAS. Research interests: automatic speech recognition, multimodal interfaces, audio-visual speech recognition. The number of publications — 150. karpov@iias.spb.su; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-7081, fax +7(812)328-7081.

**Поддержка исследований.** Данное исследование поддержано Минобрнауки РФ (ФЦП «Исследования и разработки», госконтракт № 11.519.11.4025), фондом РФФИ (проект № 12-07-31128-МОЛ\_А) и Грантом Президента РФ (проект № МК-1880.2012.8).

Рекомендовано лабораторией речевых и многомодальных интерфейсов, заведующий лабораторией Ронжин Ан.Л., д-р техн. наук, доцент.  
Статья поступила в редакцию 31.01.2013.

## РЕФЕРАТ

### *Ронжин Ал.Л., Карпов А.А.* Сравнительный анализ функциональности прототипов интеллектуальных пространств.

Развитие научной парадигмы интеллектуального пространства сформировало на сегодняшний день несколько вариантов прототипов интеллектуальной среды, обслуживающих пользователей, находящихся в некотором ограниченном пространстве: интеллектуальной комнате, доме, лекционной аудитории, зале совещаний. Разработка средств захвата и обработки аудиовизуальных сигналов, позволяющих бесконтактно оценить текущую ситуацию в наблюдаемом помещении, является одним из основных направлений исследований в этой области. В статье представлен обзор прототипов интеллектуальных пространств, предназначенных для проведения научно-образовательных мероприятий и оснащенных средствами автоматической записи выступлений участников. На основе анализа применяемых в них средств обработки аудиовизуальных сигналов и реализованных пользовательских сервисов была предложена своя классификация прототипов интеллектуальных пространств. Также описана специфика разработанного интеллектуального зала и его отличия от восьми других исследовательских прототипов, расположенных в университетах и институтах Европы, США, Китая и России. Анализ существующих прототипов показал, что целью большинства реализованных средств интеллектуальных залов является накопление исследовательских данных об активности пользователей и ситуации в помещении, которые могут дальше применяться для анализа деятельности участников мероприятий и формирования моделей автоматического управления информационными и техническими системами.

## SUMMARY

### *Ronzhin A.L., Karpov A.A.* **A Comparative Analysis of Smart Space Prototypes Functionality.**

Today the development of scientific paradigm of smart space forms the several variants of smart environment prototypes, which serve users located inside a limited premise: smart room, home, lecture and meeting room. The development of means of capture and processing for audio visual signals, which less contact estimate the current situation in the observed premise, is a one of the main direction of the research in this domain. A survey of smart space prototypes intended for scientific and educational meetings and facilitated by means of automatic speech recording is represented in the paper. Analysis of used audio visual signal processing means and realized user services allows us to propose a classification of smart space prototypes. The peculiarities of the developed smart meeting room and its distinction features from the eight research prototypes located in universities and institutes of Europe, USA, China, and Russia are described. The analysis of the existent prototypes showed that the purpose of the most realized means of smart meeting rooms is the accumulation of research data about user activities and situation inside the premises, which could be further used for study of meeting participant activity and creation of models for automatic control of information and technical systems.