

А.Л. Ронжин, В.Ю. Будков, А.Л. Ронжин  
**ФОРМИРОВАНИЕ ПРОФИЛЯ ПОЛЬЗОВАТЕЛЯ НА ОСНОВЕ  
АУДИОВИЗУАЛЬНОГО АНАЛИЗА СИТУАЦИИ В  
ИНТЕЛЛЕКТУАЛЬНОМ ЗАЛЕ СОВЕЩАНИЙ**

---

*Ронжин А.Л., Будков В.Ю., Ронжин А.Л. Формирование профиля пользователя на основе аудиовизуального анализа ситуации в интеллектуальном зале совещаний.*

**Аннотация.** Рассматривается проблема персонализированной настройки оборудования интеллектуального зала и формирования профиля пользователя на основе многоканальной обработки аудио- и видеопотоков, регистрирующих текущую ситуацию и поведение участников мероприятия в зале совещаний. Оценивание предпочтений по использованию оборудования, пользовательскому интерфейсу, роли и активности участников во время мероприятий позволяет автоматизировать процессы подготовки интеллектуального зала, управления мультимедийным презентационным и записывающим оборудованием в ходе выступлений. С помощью разработанной системы аудиовидеолокализации выступающих в ходе нескольких совещаний в интеллектуальном зале в автоматическом режиме было сделано 212 записей. Накопленные экспериментальные данные позволили оценить места в зале, с которых чаще всего задаются вопросы. Точность наведения видеокамеры на выступающего в зоне презентаций, а также в рядах кресел оценивалась по размеру и положению его лица в кадре на протяжении всей съемки и в среднем составила 90%.

**Ключевые слова:** Компьютерное зрение, интеллектуальное пространство, видеомониторинг, определение диктора, обработка аудиовизуальных сигналов, профиль пользователя.

*Ronzhin A.L., Budkov V.Yu., Ronzhin A.L. User profile forming based on audiovisual situation analysis in smart meeting room.*

**Abstract.** This paper discusses the problem of personal adjustment of smart room devices and forming of a user profile based on processing of multichannel audio and video streams, which register of the current situation and meeting participants behavior in the meeting room. Estimation of preference of device usage, user interface, participant role and their activity during meeting allows us to automate the processes of smart room preparation as well as manage multimedia presentation and record devices during events. 212 records were made during several meetings in the smart room with the help of the developed system of audio and video speaker localization. The accumulated experimental data allowed us to estimate the places in the room, from which the participants asked questions most of the time. The accuracy of camera pointing on speaker in the presentation zone as well as in the rows of sits estimated by participant's face size and its position in frame during whole recording equals 90% approximately.

**Keywords:** Computer vision, smart space, video monitoring, speaker detection, audiovisual signal processing, user profile.

---

**1. Введение.** Интеллектуальное пространство – это среда, автоматически изменяющая состояние встроенного оборудования и сервисов для удовлетворения потребностей пользователей в текущей ситуации.

Интеллектуальный зал совещаний является одним из частных случаев реализации интеллектуального пространства. Аудиовизуальное, сенсорное оборудование является одной из главных компонент интеллектуального зала и применяется для анализа ситуации, оценки местоположения участников, их роли и статуса. Анализ истории поведения участников мероприятий служит для определения их предпочтений, создания модели профиля участника мероприятия (ПУМ). Автоматизация подготовки и управления мультимедийным оборудованием во время мероприятий достигается за счёт аудиовизуального анализа ситуации и персонифицированной оценки поведения каждого участника. Далее рассмотрим некоторые варианты ПУМ и их применение для различных задач в области информационных технологий.

Типичный профиль участника включает в себя информацию, описывающую: (1) контактные данные, профессия, место работы; (2) социальное окружение; (3) характеристики используемых мобильных устройств и персональных сервисов (например, тип коммуникационных интерфейсов, разрешение дисплея, поддерживаемое устройством); (4) расписание мероприятий.

В работе [1] рассмотрена теория систем социального взаимодействия, описывающая особенности совместной деятельности малых групп, а также методологию наблюдения, индексирования и анализа мероприятий с небольшим количеством участников. Данная теория в основном ориентирована на выявление и интерпретацию микроуровневых действий участников: вербальных и невербальных, направленных на улучшение работоспособности группы.

Практический подход к системе социального взаимодействия предложила Гарвардская компания SCG [2], были разработаны компьютерные средства для анализа опросного листа Bales SYMLOG, количественной оценки взаимодействия групп во время проведения мероприятий.

Теория систем социального взаимодействия также активно применяется при исследовании и разработке систем совместной деятельности распределённых групп людей. Проведение телеконференций или распределённых мероприятий является одним из видов такого взаимодействия. Анализ и описание поведения участника локальных, распределённых и гибридных мероприятий чаще всего производится на основе систем обработки и анализа контекста, в которых профиль участника структурирован и описан, в том числе при помощи онтологий [3,4].

Онтология контекста модели профиля, представленная в работе [5], состоит из классов, описывающих типы участников и их характе-

ристики. В профиле должны содержаться данные, необходимые для персонализированной настройки сервисов, реализованных в интеллектуальном пространстве [6]. Авторы работы [5] разделяют данные в профиле участника мероприятия на два типа:

1. статические, к которым относятся тип участника и его характеристики, такие как персональные данные, характеристики (т.е. профессия, родной язык и т.д.), возможности (например, способности), физическое состояние (т.е. кровяное давление и т.д.), когнитивное состояние (например, настроение) и так далее;

2. динамические, к которым относятся поведение участника, его интересы, предпочтения, местоположение и другие характеристики [7].

При реализации распределенных мероприятий профили участников могут храниться в мобильном устройстве клиента и/или на сервере в интеллектуальном зале. У одного и того же пользователя может быть несколько профилей в зависимости от его текущего местоположения или активности.

В работе [8] описана распределенная система Easy-Meeting для интеллектуального зала совещаний, состоящая из набора интеллектуальных агентов, сервисов, устройств и сенсоров, которые выполняют операции для решения общей задачи – обеспечения участников мероприятий сервисами и информацией, релевантными текущей ситуации. Также система отвечает за регистрацию и сохранение данных о: (1) местоположении участников мероприятия; (2) расписании событий мероприятия; (3) выступлениях, запланированных в мероприятии; (4) профилях выступающих участников; (5) статусе мероприятия.

Современные системы для интеллектуальных залов совещаний включают в себя автоматический захват и индексирование рукописных заметок, сделанных на электронной доске [9], слежение за перемещением и предоставление поддержки участникам мероприятий [10], перенос графического пользовательского интерфейса (GUI) [11], определение и слежение за докладчиками [12], распознавание и синтез речи [13, 14].

**2. Модель профиля участника мероприятия.** Рассмотренные выше варианты моделей ПУМ, а также способы получения и накопления данных об участниках мероприятий, были использованы при разработке системы аудиовидеолокализации выступающих и модели ПУМ, внедренных в интеллектуальном зале СПИИРАН. Разработанный интеллектуальный зал предназначен для проведения образовательных и научных мероприятий, таких как лекция, совещание, теле-

конференция и т.д. В связи с широким набором возможных мероприятий некоторые характеристики участников, например роль и статус, могут меняться. На рисунке 1 представлена структура данных профиля пользователя интеллектуального зала СПИИРАН [15].

Представленную структуру можно разделить на две категории данных: введенные пользователем и накапливаемые автоматически. Рассмотрим первую категорию, состоящую из подкатегорий «Личные данные» и «Дополнительные данные». Личные данные содержат персональные характеристики пользователя, а также его интересы и предпочтения. Кроме того, к ним относится информация об учетной записи пользователя интеллектуального зала. В подкатегории «Дополнительные данные» выделяют три типа дополнительной информации: (1) контактные данные; (2) служебные данные; (3) физиологические особенности.



Рис. 1. Модель ПУМ.

В ходе проведения мероприятий в интеллектуальном зале автоматически накапливаются данные об участниках, находящихся в зале, и используемом ими мультимедийном презентационном оборудовании. В случае распределенных мероприятий собирается информация о характеристиках клиентских устройств удалённых участников. Предпочтения содержат информацию о наиболее часто используемых сервисах, номерах кресел, которые участник занимал при участии в различных мероприятиях, а также описание оборудования, необходимого участнику, например, для выступления на мероприятии.

**3. Модель мероприятия.** Особенности настройки оборудования для конкретного мероприятия, а также история поведения участников во время их проведения в разработанной системе управления интеллектуальным залом хранятся в модели мероприятия, представленной на рисунке 2. Она включает в себя две категории данных: заранее заданные и текущие. Под «Заранее заданными данными» подразумеваются те данные, которые были определены заранее организаторами. Эта категория включает в себя подкатегорию «Тип мероприятия», в которой перечислены образовательные и научные мероприятия, проводимые в интеллектуальном зале.

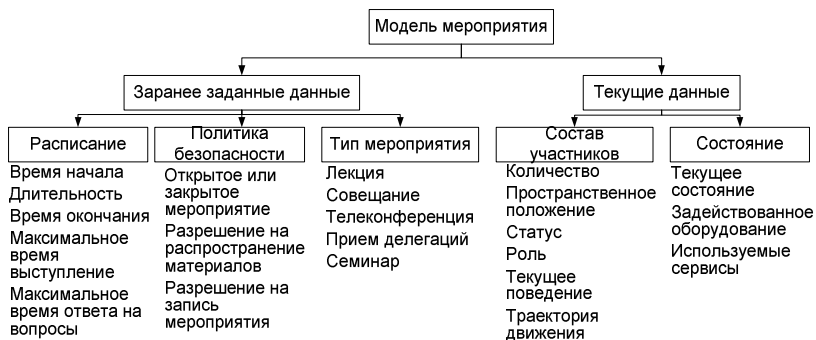


Рис. 2. Структура модели мероприятия.

Подкатегория «Расписание» описывает временные характеристики мероприятия, такие как: время начала, окончания и длительность мероприятия, а также максимальное время, отведённое для выступления, каждому участнику, и максимальное время дискуссии по окончании доклада, а также содержит информацию о секциях, на которые может быть разделено мероприятие, о кофе-брейках и т.д. Подкатегория «Политика безопасности» включает в себя пункт, определяющий открытое или закрытое проводится мероприятие, а также информацию о разрешении или запрете на распространение материалов и запись мероприятия.

Категория «Текущие данные» содержит информацию, дополняющуюся и изменяющуюся в процессе проведения мероприятия, она включает в себя параметры, описывающие состав участников, например, их текущее местоположение, количество и их передвижения по залу. Также в рассматриваемой категории описывается состояние мероприятия, которое содержит информацию о текущей части мероприя-

тия, а также о задействованном оборудовании и используемых сервисах.

Проанализируем параметры, описывающие участников мероприятия, более детально. Пространственное положение участников включает в себя данные об их текущих координатах и нахождение в одной из специфических зон интеллектуального зала. Кроме того, рассматриваемая категория содержит информацию о номере кресла в случае присутствия в нем участника. Статус участника выбирается из трех вариантов (локальный, удаленный, отсутствует на мероприятии) при помощи анализа информации, полученной от системы аудиовизуального мониторинга интеллектуального зала. Если участник присутствует на мероприятии, то у него есть роль, имеющая три разновидности: (1) основной докладчик; (2) докладчик; (3) слушатель.

Текущее поведение участника включает в себя три подкатегории: (1) тип движения; (2) тип аудиоактивности; (3) тип управления оборудованием. Каждая из представленных подкатегорий может принимать различные варианты, например, тип движения варьируется в зависимости от скорости перемещения участника и принимает значения: сидит, стоит, идет, бежит. Вторая подкатегория принимает один из трех вариантов типа аудиоактивности: (1) говорит; (2) шепчет; (3) шумит. Тип управления состоит из трех параметров: (1) наименование оборудования; (2) тип применяемой модальности; (3) тип действия.

Траектория движения представляет собой набор последовательных координат перемещения объекта, индекс зон, в которых он находился, скорость движения, местное время для каждой координаты, а также этап проведения мероприятия.

Рассмотрим процесс получения личных данных профиля пользователя, состоящий из нескольких этапов. На первом этапе выполняется получение информации об участнике с помощью системы аудиовизуального мониторинга, осуществляющей регистрацию участников мероприятия. Регистрация участников, сидящих в рядах кресел (зоне кресел), выполняется за счёт совместной работы многоканальной системы видеомониторинга (МСВ), многоканальной системы определения источника звука (МСА), многомодальной системы управления интеллектуального зала (МСУИЗ). Данные системы выполняют фотографирование крупным планом лиц участников и запись аудиохарактеристик их речи, а также запись и распознавание личных данных в режиме диалога с использованием технологии синтеза речи и аватара для визуализации мимики лица [16].

МСА выполняет захват аудиопотоков с трёх массивов микрофонов, вычисление координат источника звука в зале и оценку границ речевого сообщения. Полученные данные о координатах источника звука усредняются по трём массивам микрофонов, а затем передаются в МСУИЗ.

Получение визуальных данных МСВ осуществляет в три этапа [16], на каждом из них задействована отдельная видеокамера. В начале первого этапа считывается кадр с панорамной камеры, установленной на потолке в центре зала, после его обработки принимается решение о наличии участника в кресле. По окончании обработки всех кресел со стационарной камеры, охватывающей ряды кресел полностью, считывается кадр, в котором производится поиск лиц участников в зонах возможного появления лица, соответствующих занятым креслам. Далее осуществляется наведение интеллектуальной PTZ камеры на лицо каждого участника так, чтобы в кадре его лицо было снято крупным планом. Если лицо найдено, в МСУИЗ передается файл с фотографией участника и номер его кресла. В противном случае МСВ выполняет повторный поиск лица участника в рассматриваемом кресле, после того, как будут сфотографированы все остальные участники.

После получения всей аудиовизуальной информации МСУИЗ выполняет распознавание аудиоданных, записанных при проведении диалога с участниками мероприятия. По окончании данной операции система осуществляет создание и заполнение модели ПУМ для каждого из зарегистрированных участников, а также обновление имеющихся профилей участников.

При регистрации и сохранении данных участников в модели мероприятий, описывающих текущую ситуацию в интеллектуальном зале, участвуют системы МСА и МСВ, выполняющие обработку аудиовизуальной информации.

**4. Эксперименты.** Данные об участниках собирались в ходе имитации мероприятий, когда пользователи по заданному сценарию проводили совещание, а также реальных семинаров, когда участники были оповещены о регистрации их поведения, но это не влияло на ход запланированной ими деятельности в интеллектуальном зале.

Для проведения мероприятий по заданному сценарию была выбрана группа людей, численность которой варьировалась от 5 до 10 человек. План проведения мероприятий состоял из двух этапов. На первом этапе участникам эксперимента предлагалось занять любые кресла, после этого в интеллектуальный зал входил дополнительный участник, выступающий в роле лектора. Далее «лектор» задавал по-

очередно десять вопросов аудитории без назначения конкретного лица для ответа на вопрос. Кроме того, вопрос формулировался таким образом, чтобы любой из участников смог дать ответ на вопрос. После ответа на все заданные вопросы участники эксперимента выходили из интеллектуального зала.

На втором этапе при проведении имитации мероприятия участникам предлагалось сесть на любое другое кресло, на котором они ещё не сидели. Остальная часть эксперимента аналогична первому этапу. Основной целью данного экспериментального исследования было выявление мест, которые занимали участники, чаще всего проявлявшие активность во время мероприятия. На рисунке 3 представлена схема интеллектуального зала, на которой штриховыми линиями отмечены кресла, в которых участники проявляли наибольшую активность при ответах на вопросы.

На рисунке 3 показан конференц-стол, за которым размещаются до 10 участников, проекционный экран и плазменная панель, кроме того, в правой части зала расположены ряды кресел, вмещающие до 32 участников. Кроме того, в интеллектуальном зале СПИИРАН установлено аудиовизуальное сенсорное оборудование, подробное описание расположения оборудования и его применение представлено в работе [17]. Большинство кресел обозначено штриховыми линиями различного типа, означающим интенсивность активности участников мероприятий, сидящих в представленных креслах. Кресла, которые в процессе проведения мероприятий оказались пустыми, оставлены не заштрихованными.

Теперь рассмотрим мероприятия, проводимые без заранее составленного сценария и списка вопросов для участников. Особенностью такого мероприятия является наличие повышенного уровня аудиощумов в зале, реплик, не относящихся к обсуждаемой теме, а также перекрывающейся речи в аудиосигнале, одновременно произнесенной несколькими активными участниками и/или докладчиком [16]. Наличие данных факторов может повлиять на точность работы системы определения положения источника звука и привести появлению ложных источников звуков.

В автоматическом режиме было записано 212 событий, содержащих реплики участников мероприятий во время дискуссии. В ходе экспертной проверки было выявлено, что 7% событий содержат посторонний шум или речь выступающего (например, лектора), а 93% содержат реплики сидящих в зале участников. В результате было обна-



ружено, что активные участники чаще всего занимают места в левой половине рядов кресел.

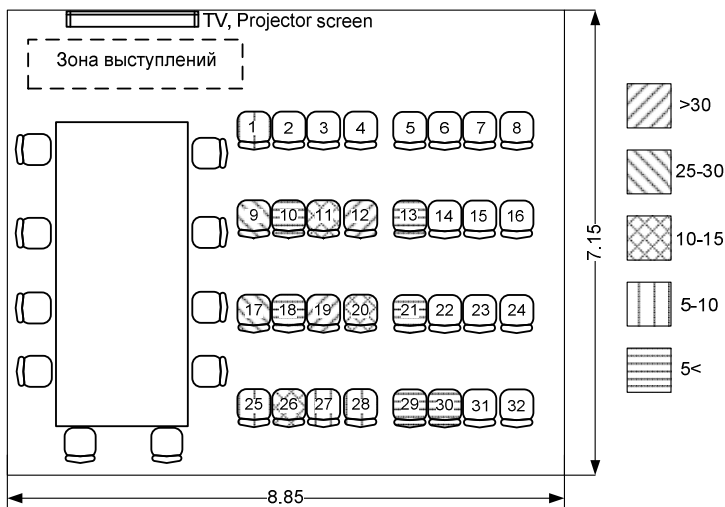


Рис. 3. Схема аудиоактивности участников мероприятий в процессе дискуссии.

Накопленные визуальные данные были применены для экспериментальной оценки эффективности работы системы видеомониторинга. Точность наведения видеокамеры на выступающего в зоне презентаций, а также в рядах кресел оценивалась по размеру и положению его лица в кадре на протяжении всей съемки. Анализ фотографий, сделанных в автоматическом режиме наведения, показал, что система определяет участников, сидящих в зоне кресел, и выполняет наведение камеры на выступающего докладчика с точностью 91% и 90% соответственно. Также были проведены эксперименты по оценке формирования аудиовизуальных записей, содержащих выступления докладчиков и высказывания активных участников во время дискуссии [18].

**5. Заключение.** Для автоматизации управления оборудованием в интеллектуальном зале СПИИРАН была разработана модель ПУМ, состоящая из двух категорий данных: введенные участником и автоматически накапливаемые, а также модель мероприятия, содержащая данные о текущей ситуации и заранее заданные параметры. Накопление данных производится системами автоматической обработки аудио- и видеосигналов, осуществляющих анализ ситуации в процессе

проведения мероприятия, а также слежение за перемещением участников и запись выступлений участников.

Данное исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта № 12-07-31128-МОЛ\_А, а также гранта Президента РФ (МД-501.2011.8).

## Литература

1. *Ellis C. A.* Project NEEM: Agent Based Meeting Augmentation // Technical report CU-CS-937-02, University Of Colorado at Boulder, Department of Computer Science, 2002. P. 1-20
2. *Bales R. F., Cohen S. P., Williamson S. A.* SYMLOG: A System for the Multiple Level Observation of Groups // The Free Press, N.Y., – 1979. – 537 p.
3. *Кузнецов О. П., Суховеров В. С., Шитилина Л. Б.* Онтологии в современных информационных системах // Датчики и Системы, №8, 2011. – С. 67-77.
4. *Chen H., Finin T., Joshi A.* The SOUPA Ontology for Pervasive Computing // Whitestein Series in Software Agent Technologies and Autonomic Computing, 2005, – P. 233-258.
5. *Qin W., Zhang D., Shi Y., Du K.* Combining User Profiles and Situation Contexts for Spontaneous Service Provision in Smart Assistive Environments // F.E. Sandnes et al. (Eds.): UIC 2008, LNCS 5061, 2008. – P. 187–200.
6. *Yu S., Al-Jadir L., Spaccapietra S.* Matching user's semantics with data semantics in location-based services // In Proceedings of the Workshop on Semantics in Mobile Environments 2005. – P. 147-152.
7. *Carmagnola F.* The five ws in user model interoperability // In Proceedings of the IUI 2008 Workshop on Ubiquitous User Modeling. 2008. – P. 1-6.
8. *Chen H., Perich F., Chakraborty D., Finin T., Joshi A.* Intelligent agents meet semantic web in a smart meeting room // In Proceedings of the 3rd International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS '04), 2004. – P. 854–861.
9. *Hammond T., Gajos K., Davis R., Shrobe H.* An agent based system for capturing and indexing software design meetings // In Proceedings of the International Workshop on Agents in Design (WAID'02), Cambridge, MA, 2002. – P. 24–26.
10. *Coen M. H.* Design principles for intelligent environments // In Proceedings of AAAI/IAAI, 1998. – P. 547–554.
11. *Bennett F., Richardson T., Harter A.* Teleporting -Making Applications Mobile // In Proceedings of Workshop on Mobile Computing Systems and Applications, Santa Cruz, 1994. – P. 82–84.
12. *Tambe M., Scerri P., Pynadath D. V.* Adjustable autonomy for the real world // In Proceedings of AAAI Spring Symposium on Safe Learning Agents. 2002. – P. 171–228.
13. *Фархадов М.П.* Распознавание речи в системах массового обслуживания населения // Труды СПИИРАН. Вып. 19, – 2011. – С. 65–86.
14. *Мещеряков Р.В.* Структура систем синтеза и распознавания речи // Известия ТПУ, № 5, – Т. 315, – 2009. – С. 121-126.
15. *Юсупов Р.М., Ронжин Ан.Л.* От умных приборов к интеллектуальному пространству // Вестник Российской Академии Наук: научный и общественно-политический журнал, Том 80, – Вып. 1, – 2010. – С. 45-51.
16. *Ронжин Ал.Л., Ронжин Ан.Л.* Система аудиовизуального мониторинга участников совещания в интеллектуальном зале // Доклады ТУСУРа, № 1 (22), – Ч. 1, – 2011. – С. 153-157.

17. *Ронжин Ан.Л., Будков В.Ю.* Технологии поддержки гибридных e-совещаний на основе методов аудиовизуальной обработки // Вестник компьютерных и информационных технологий, № 4, – 2011. – С. 31-35
18. *Ронжин Ал.Л.* Способы оценивания систем видеомониторинга ситуации в зале совещаний // Информационно-измерительные и управляющие системы, № 11, – Т. 9, – 2011. – С. 12-16.

**Ронжин Александр Леонидович** — младший научный сотрудник лаборатории речевых и многомодальных интерфейсов СПИИРАН. Область научных интересов: технологии интеллектуального пространства, аудиолокализации, техническое зрение. Число научных публикаций — 25. [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 — 25. [ronzhinal@iias.spb.su](mailto:ronzhinal@iias.spb.su); SPIIRAS, 39, 14-th line, St. Petersburg, 199178, Russia; office phone +7(812)328-7081, fax +7(812)328-7081.

**Будков Виктор Юрьевич** — младший научный сотрудник лаборатории речевых и многомодальных интерфейсов СПИИРАН. Область научных интересов: автоматическое распознавание речи, многомодальные интерфейсы, аудиовизуальное распознавание речи. Число научных публикаций — 40. [budkov@iias.spb.su](mailto:budkov@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.

**Budkov Viktor Yurievich** — junior researcher, Laboratory of Speech and Multimodal Interfaces, SPIIRAS. Research interests: automatic speech recognition statistical language models. The number of publications — 40. [budkov@iias.spb.su](mailto:budkov@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.

**Ронжин Андрей Леонидович** — док. техн. наук, заведующий лаборатории речевых и многомодальных интерфейсов СПИИРАН. Область научных интересов: автоматическое распознавание речи, многомодальные интерфейсы, аудиовизуальное распознавание речи. Число научных публикаций — 170. [ronzhin@iias.spb.su](mailto:ronzhin@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 Andrey Leonidovich** — Dr. Tech. Sci., Assoc. Prof., Head of the Laboratory of Speech and Multimodal Interfaces, SPIIRAS. Research interests: automatic speech recognition, multimodal interfaces, audio-visual speech recognition. The number of publications — 170. [ronzhin@iias.spb.su](mailto:ronzhin@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 выполнено при финансовой поддержке РФФИ в рамках научного проекта № 12-07-31128-МОЛ\_А, а также в рамках федеральной целевой программы «Исследования и разработки по приоритетным направлениям развития научно-технологического комплекса России на 2007–2013 годы» (ГК № 11.519.11.4025).

Рекомендовано лабораторией речевых и многомодальных интерфейсов, заведующий лабораторией Ронжин Ан.Л., д-р техн. наук, доцент.  
Статья поступила в редакцию 01.10.2012.

## РЕФЕРАТ

*Ронжин Ал.Л., Будков В.Ю., Ронжин Ан.Л.* **Формирование профиля пользователя на основе аудиовизуального анализа ситуации в интеллектуальном зале совещаний.**

В статье представлен обзор методов и систем формирования профиля пользователя интеллектуального пространства. Составлена типичная модель профиля, включающая в себя информацию, описывающую: (1) контактные данные, профессия, место работы; (2) социальное окружение; (3) характеристики используемых мобильных устройств и персональных сервисов (например, тип коммуникационных интерфейсов, разрешение дисплея, поддерживаемое устройством); (4) расписание мероприятий.

Кроме того, рассматривается проблема персонифицированной настройки оборудования интеллектуального зала и формирования профиля пользователя на основе многоканальной обработки аудио- и видеопотоков, регистрирующих текущую ситуацию и поведение участников мероприятия в зале совещаний. Оценивание предпочтений по использованию оборудования, пользовательскому интерфейсу, роли и активности участников во время мероприятий позволяет автоматизировать процессы подготовки интеллектуального зала, управления мультимедийным презентационным и записывающим оборудованием в ходе выступлений. С помощью разработанной системы аудиовидео-локализации выступающих в ходе нескольких совещаний в интеллектуальном зале в автоматическом режиме было сделано 212 записей. Накопленные экспериментальные данные позволили оценить места в зале, с которых чаще всего задаются вопросы. Точность наведения видеокамеры на выступающего в зоне презентаций, а также в рядах кресел оценивалась по размеру и положению его лица в кадре на протяжении всей съемки и в среднем составила 90%.

## SUMMARY

### ***Ronzhin Al.L., Budkov V.Yu., Ronzhin An.L. User profile forming based on audiovisual situation analysis in smart meeting room.***

The paper proposes a survey of methods and systems of smart room user profile formation. The typical user profile model is composed, which includes information, which describe: 1) contacts, profession, job; 2) social environment; 3) characteristic of mobile devices and personal services, which are used (for example, type of communication interface, screen size and etc.); 4) meetings schedule.

In addition, this paper discusses problem of personal adjustment of smart room devices and forming of a user profile based on processing of multichannel audio and video streams, which register current situation and meeting participants behavior in meeting room. Estimation of preference of device usage, user interface, participant role and their activity during meeting allows us to automate the processes of smart room preparation as well as to manage multimedia presentation and record devices during events. 212 records were made during several meetings in the smart room with the help of the developed system of audio and video speaker localization. The accumulated experimental data allowed us to estimate the places in the room, from which the participants asked questions most of the time. The accuracy of camera pointing on speaker in the presentation zone as well as in the rows of sits estimated by participant's face size and its position in frame during whole recording equals 90% approximately.