

В.Ю. БУДКОВ, М.В. ПРИЩЕПА
**РАЗРАБОТКА МОДЕЛЬНО-АЛГОРИТМИЧЕСКОГО
ОБЕСПЕЧЕНИЯ УПРАВЛЕНИЯ РАСПРЕДЕЛЕННЫМИ
МОДУЛЯМИ ИНТЕЛЛЕКТУАЛЬНОГО ЗАЛА**

Будков В.Ю., Прищепка М.В. Разработка модельно-алгоритмического обеспечения управления распределенными модулями интеллектуального зала.

Аннотация. Способность приложений анализировать текущие условия эксплуатации, в том числе, текущее состояние пользователя, физического окружения, вычислительных ресурсов, и динамически адаптировать сценарий взаимодействия с пользователем является одним из главных требований при разработке прототипов интеллектуальных пространств. С увеличением решаемых задач и числа обслуживаемых пользователей становится затруднительным держать под контролем множество программных и аппаратных модулей, задействованных в интеллектуальном пространстве, поэтому необходимо математическое обеспечение и программные средства, реализующие контроль совместной работы распределенных модулей. В статье описывается разработанная структура модели управления распределенными модулями интеллектуального зала и используемые многомодальные интерфейсы для обеспечения естественности человеко-машинного взаимодействия.

Ключевые слова: многомодальные интерфейсы, интеллектуальное пространство, веб-интерфейсы, обработка аудиовизуальных сигналов.

Budkov V.Yu., Prishchepa M.V. Development of models and algorithms for control of distributed modules in the smart meeting room.

Abstract. The software ability to analyze current operating conditions, including a current state of a user, a physical environment, computing resources, as well as dynamically adaptation of the scenario of interaction to the user, is a one of the main requirements in the development of smart space prototypes. Control of a set of software and hardware modules involved in the smart space is difficult with increasing of tasks and number of users, so mathematical and software tools, which implement control of distributed modules, are required. The paper describes the structure of the model for distributed modules control in a smart meeting room as well as used multimodal interfaces for natural human-computer interaction.

Keywords: multimodal interfaces, smart space, web interfaces, audiovisual signal processing.

1. Введение. Впервые понятие окружающего интеллектуального пространства было использовано сотрудниками компании Philips [1] для описания прототипа интеллектуальной гостиной комнаты со встроенными устройствами, обеспечивающей проактивное ненавязчивое персонализированное обслуживание находящихся в ней людей. Затем в материалах Европейской комиссии по прогнозу научно-технических результатов на 2010 год были отмечены три основных группы технологий, с которыми связано развитие систем окружающего интеллекта: повсеместные вычисления, повсеместные коммуникации и дружелюбные пользовательские интерфейсы [2]. В России тер-

мин повсеместные вычисления впервые упоминался в 2008 году в работе [3], тем не менее, работы по интеграции сетей пользовательских устройств на основе многоагентного подхода, а также разработка многомодальных пользовательских интерфейсов, обеспечивающих естественное взаимодействие с пользователем, начались гораздо раньше [4,5]. Необходимость привлечения и развития трех упомянутых выше базовых групп технологий обоснована тем, что все без исключения устройства интеллектуального пространства снабжены вычислительными ресурсами для связи и распределения данных между устройствами необходимы мощные сетевые ресурсы и, наконец, для обслуживания пользователей в ненавязчивом режиме требуются интерфейсы, использующие естественные для человека способы коммуникации.

Так как под окружающим интеллектуальным пространством понимается глобальное единое пространство, то его создание в ближайшее время затруднительно, поэтому сейчас ведутся исследования по разработке отдельных менее масштабных прототипов интеллектуальных пространств, например, «умная» комната, «умный» дом, «умный» город [6,7]. Также существующие прототипы можно разделить в зависимости от области их применения, например: умная больничная палата, интеллектуальный зал совещаний, умный поезд.

Несмотря на популярность «умных» технологий в большинстве случаев их интеллектуальность касается только автоматизации средств взаимодействия между объектами, а взаимодействие между пользователем и клиентским устройством осуществляется на основе сложившихся контактных способов ввода данных и управления, а также графического интерфейса. Хотя одной из основных целей разработки окружающего интеллектуального пространства является ненавязчивое естественное бесконтактное взаимодействие с пользователем. Для реализации такого типа взаимодействия необходимо анализировать поведение пользователей в режиме реального времени с помощью наборов гетерогенных типов датчиков. Учитывая естественную асинхронность естественных модальностей, а также возможные задержки в передаче и обработке аудиовизуальных потоков, проблема синхронизации совместной работы распределенных модулей отвечающих за регистрацию, распознавание и синтез естественных модальностей (речь, жесты, положение тела), а также управления информационно-управляющими сервисами в интеллектуальном пространстве является актуальной.

2. Распределенные модули интеллектуального зала. Разработанный вариант интеллектуального зала представляет собой систему аппаратно-программных модулей, активационных и коммутирующих

устройств, мультимедийных средств и аудиовизуальных сенсоров, функционально взаимосвязанных между собой и собранных в единое устройство, представляющее собой интеллектуальное помещение, обеспечивающее участников мероприятия необходимыми сервисами в автоматическом режиме. Осведомленность комплекса о пространственном положении участников, их текущих действиях, роли в текущем мероприятии и их предпочтениях помогает более точно предсказать намерения и потребности участников. Моделирование контекста, извлечение, синхронизация и распределение знаний остаются наиболее важными задачами при проектировании интеллектуального пространства. Разработанные технологии для автоматической обработки аудиовизуальных данных были успешно внедрены в интеллектуальном зале для обеспечения естественного взаимодействия с оборудованием. Среди наиболее важных технологий, примененных в разработанном техническом решении, следует отметить: автоматическое распознавание русской речи, голосовую идентификацию диктора, локализацию источников звука, определение положения и слежение за движущимися объектами и лицами людей, определение позы человека [7,8].

Для автоматизации управления презентационным, аудио-, видео- и активационным оборудованием было предложено модельно-алгоритмическое обеспечение анализа текущего поведения участников и состояния устройств в интеллектуальном пространстве. Многомодальное управление оборудованием как внутри помещения, так и удаленно реализовано через веб-интерфейс, графическая компоновка которого учитывает характеристики дисплея клиентского устройства [9].

Анализ модельно-алгоритмического обеспечения многоканальной обработки аудиопотоков, использованного при разработке многомодального интерфейса к интеллектуальному пространству приведен в работах [8,10]. Совместная работа технологий аудиовизуальной обработки сигналов снабжает систему управления помещением данными о текущей обстановке в помещении, поведении пользователей, а также обеспечивает дистанционное распознавание голосовых команд за счет анализа пространственно-временной, ситуативной информации и предпочтений пользователей [11,12].

На рисунке показана функциональная схема работы интеллектуального зала, основными модулями которого являются: центральный сервер с многоядерной архитектурой, организующий взаимосвязь всех модулей и серверов (сервер аудиолокализации, сервер видеонаблюдения, сервер веб-трансляции мероприятий и сервер презентаций). Кроме того, к центральному серверу через коммутирующее устройство

подключены активационные устройства и мультимедийное оборудование. На центральный сервер передаются данные с серверов видеонаблюдения, аудиолокализации и презентаций.



Рис. Функциональная схема распределенных модулей интеллектуального зала.

Сервер видеонаблюдения отвечает за обработку данных, поступающих с IP-видеокамер. К серверу аудиолокализации подключены многоканальные платы аудиозахвата, обслуживающие массивы микрофонов. Данные о текущей звуковой активности в зале и координаты источников звука с этого сервера поступают на центральный сервер, который передает их и на сервер видеонаблюдения для наведения видеокамер на текущего выступающего.

Данные, поступающие с веб-камер, обрабатываются на сервере веб-трансляции мероприятия с целью оптимизации компоновки веб-интерфейса для удаленного участника и передачи на клиентское устройство аудиовизуального контента, позволяющего наиболее эффективно воспринять текущую ситуацию в зале. Сервер веб-трансляции обрабатывает потоки данных от удаленных пользователей, подключенных через веб-интерфейс, и передает их на центральный сервер. Кроме того, удаленный пользователь может переключать состояние оборудования интеллектуального зала и наблюдать за текущей обстановкой в нем с помощью модуля удаленного управления оборудованием. На центральном сервере данные о состоянии устройств по-

мещения обрабатываются и отображаются с помощью многомодальной системы управления интеллектуальным залом.

Сервер презентации предназначен для управления проектором и презентациями в ходе проведения мероприятий, в том числе, переключение слайдов, а также выполнение заранее заданного сценария показа нескольких презентаций докладчиков мероприятия. Этот сервер передает данные о состоянии проектора на центральный сервер.

Основное преимущество разработанного интеллектуального зала заключается в обеспечении участников мероприятий информационно-управляющими сервисами с многомодальным интерфейсом на основе анализа текущей ситуации в помещении и учета предпочтений участников предшествующих мероприятий. Разработка средств эффективного взаимодействия человека с компьютером сегодня является одним из приоритетных направлений развития информатики и искусственно-интеллекта в рамках создания критических для РФ технологий обработки, хранения, передачи и защиты информации. Это связано с тем, что уже сейчас вычислительная техника не используется в полной мере из-за отсутствия полноценного, привычного человеку, интерфейса для ввода информации и взаимодействия пользователя с компьютером. Изучение различных комбинаций многомодальных интерфейсов для взаимодействия с окружающим интеллектуальным пространством помогает разрешить фундаментальные вопросы человеко-машинного взаимодействия, а созданные модели будут востребованы при разработке пользовательских интерфейсов для новых приложений в области безопасности, медицины, робототехники, логистики и других научных направлений.

Кроме окружающих интеллектуальных пространств многомодальные интерфейсы получили широкое распространение в робототехнике. Особое внимание ученых и инженеров направлено на развитие обслуживающих роботов, предоставляющих ассистивные, информационные, обучающие, развлекательные услуги. Разработанные в ходе создания интеллектуального зала модели и способы человеко-машинного взаимодействия используются также в мобильном информационно-справочном роботе [13,14]. Разработанный информационно-справочный робот представляет собой мобильную подвижную систему, которая содержит сеть интеллектуальных аппаратно-программных модулей, активационных устройств, мультимедийных средств и аудиовизуальных сенсоров. Основная задача робота — обеспечение пользователей необходимой справочной информацией и сервисами на основе автоматического анализа окружающей обстановки.

3. Заключение. Разработанный интеллектуальный зал представляет собой систему распределенных, функционально взаимосвязанных между собой аппаратно-программных модулей, которые включают в себя активационные и коммутирующие устройства, сенсоры и различное мультимедийное оборудование. Совместная работа технологий аудиовизуальной обработки сигналов снабжает систему управления интеллектуальным пространством данными о текущей обстановке в помещении. Многомодальное управление оборудованием возможно как внутри помещения, так и удаленно, через веб-интерфейс, что повышает удобство пользования данной системой и позволяет предоставить посетителю различные информационно-управляющие сервисы с учетом текущей ситуации в помещении и предпочтений участников предшествующих мероприятий. В настоящее время ведется разработка системы автоматического контроля работы программных модулей, позволяющей в автономном режиме восстанавливать их работоспособность в случае их аварийного завершения.

Литература

1. *Zelkha E., Epstne B.* From Devices to Ambient Intelligence. Digital Living Room Conference, June 1998.
2. *Ducatel K., Bogdanowicz M., Scapolo F., Leijten J., Burgelman, J-C.* ISTAG - Scenarios of Ambient Intelligence in 2010 // European Commission Community Research. Feb. 2001.
3. *Городецкий В.И., Карсаев О.В., Самойлов В.В., Серебряков С.В.* Агентская платформа для повсеместных вычислений // Информационные технологии и вычислительные системы. Вып. 4. М.: Наука, 2008. – С. 51–60.
4. *Городецкий В.И., Карсаев О.В.* Многоагентная система планирования и составления расписаний: разработка распределенной базы знаний // Известия Южного федерального университета. Технические науки. 2001. Т. 22. № 4. С. 25-38.
5. *Карпов А.А., Ронжин А.Л., Ли И.В., Шалин А.Ю.* Речевые технологии в многомодальных интерфейсах // Труды СПИИРАН. 2004. Т. 1. № 2. С. 183-193.
6. *Aldrich F.* Smart Homes: Past, Present and Future // Inside the Smart Home / Ed. Harper R. London: Springer-Verlag, 2003. – pp. 17–39.
7. *Юсупов Р.М., Ронжин А.Л.* От умных приборов к интеллектуальному пространству // Вестник Российской академии наук. 2010. Т. 80. № 1. С. 45-51.
8. *Ronzhin A.L., Budkov V.Yu.* Multimodal Interaction with Intelligent Meeting Room Facilities from Inside and Outside // Springer-Verlag Berlin Heidelberg, S. Balandin et al. (Eds.): NEW2AN/ruSMART 2009, LNCS 5764, 2009, pp. 77–88.
9. *Ронжин А.Л., Будков В.Ю., Ронжин Ал.Л.* Технологии формирования аудиовизуального интерфейса системы телеконференций // Автоматизация и современные технологии. № 5. 2011, С. 20-26.
10. *Ронжин А.Л., Карпов А.А., Казиров И.А.* Особенности дистанционной записи и обработки речи в автоматах самообслуживания // Информационно-управляющие системы, Вып. 42, т. 5. СПб., ГУАП, 2009, С. 32–38.

11. *Ронжин Ал.Л., Ронжин Ан.Л.* Система аудиовизуального мониторинга участников совещания в интеллектуальном зале // Доклады ТУСУРа, № 1 (22), часть 1, 2011, С. 153-157.
12. *В.Ю. Будков, М.В. Прищепа, С.В. Глазков.* Контекстно-осведомленная система управления оборудованием и веб-трансляции мероприятий из интеллектуального зала // Вестник БГУ, Выпуск 9, 2011, С. 75-82.
13. *М.В. Прищепа, В.Ю. Будков, А.Л. Ронжин.* Система интеллектуального управления мобильным информационно-справочным роботом // Информационно-управляющие системы, 2010, № 6, С 2-6.
14. *Прищепа М.В.* Разработка профиля пользователя с учетом психологических аспектов взаимодействия человека с информационным мобильным роботом // Труды СПИИРАН. 2012. Вып. 21. С. 56–70.

Будков Виктор Юрьевич — младший научный сотрудник лаборатории речевых и мультимодальных интерфейсов СПИИРАН. Область научных интересов: автоматическое распознавание речи, мультимодальные интерфейсы, аудиовизуальное распознавание речи. Число научных публикаций — 40. budkov@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. Научный руководитель — д.т.н., доцент А.Л. Ронжин.

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; SPIIRAS, 39, 14-th Line V.O., St. Petersburg, 199178, Russia; office phone +7(812)328-7081, fax +7(812)328-7081. Scientific adviser — Dr. Tech. Sci., Assoc. Prof. A.L. Ronzhin.

Прищепа Мария Викторовна — младший научный сотрудник лаборатории речевых и мультимодальных интерфейсов СПИИРАН. Область научных интересов: модели взаимодействия пользователей с информационными роботами, разработка персонализированных стратегий человеко-машинного диалога, адаптация технологий аудио- и видеопереработки для их применения на подвижной платформе. Число научных публикаций — 30. prischepa@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. Научный руководитель — д.т.н., доцент А.Л. Ронжин.

Prischepa Maria Viktorovna — junior researcher, Laboratory of Speech and Multimodal Interfaces SPIIRAS. Research interests: models of interaction between user and information robot, development of personified strategies for human-machine interaction. The number of publications — 30. prischepa@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. Scientific adviser — Dr. Tech. Sci., Assoc. Prof. A.L. Ronzhin.

Поддержка исследований. Данное исследование выполнено при финансовой поддержке РФФИ в рамках научных проектов № 12-08-31520-мол_а и № 12-07-31201-мол_а.

Рекомендовано лабораторией речевых и мультимодальных интерфейсов, заведующий лабораторией Ронжин Ан.Л., д-р техн. наук, доцент.
Статья поступила в редакцию 15.10.2012.

РЕФЕРАТ

Будков В.Ю., Прищепа М.В. **Разработка модельно-алгоритмического обеспечения управления распределенными модулями интеллектуального зала.**

В статье описывается разработанная структура модели управления распределенными модулями интеллектуального зала и используемые многомодальные интерфейсы для обеспечения естественности человеко-машинного взаимодействия. Способность приложений анализировать текущие условия эксплуатации, в том числе, текущее состояние пользователя, физического окружения, вычислительных ресурсов, и динамически адаптировать сценарий взаимодействия с пользователем является одним из главных требований при разработке прототипов интеллектуальных пространств. С увеличением решаемых задач и числа обслуживаемых пользователей становится затруднительным держать под контролем множество программных и аппаратных модулей, задействованных в интеллектуальном пространстве, поэтому необходимо математическое обеспечение и программные средства, реализующие контроль совместной работы распределенных модулей. Изучение различных комбинаций многомодальных интерфейсов для взаимодействия с окружающим интеллектуальным пространством помогает разрешить фундаментальные вопросы человеко-машинного взаимодействия, а созданные модели будут востребованы при разработке пользовательских интерфейсов для новых приложений в области безопасности, медицины, робототехники, логистики и других научных направлений.

SUMMARY

Budkov V.Yu., Prischepa M.V. **Development of models and algorithms for control of distributed modules in the smart meeting room.**

The paper describes the structure of the model for distributed modules control in the smart meeting room as well as used multimodal interfaces for natural human-computer interaction. The software ability to analyze current operating conditions, including a current state of a user, physical environment, computing resources, as well as dynamically adaptation of the scenario of interaction to a user, is a one of the main requirements in the development of smart space prototypes. Control of a set of software and hardware modules involved in the smart space is difficult with increasing of tasks and number of users, so mathematical and software tools, which implement control of distributed modules, are required. The study of various combinations of multimodal interfaces for interaction with the intellectual space helps to resolve the fundamental issues of human-computer interaction and developed models could be useful for new application of interfaces in the field of security, medicine, robotics, logistics and other scientific fields.