

УДК 004.896

Архитектура и программные средства системы сопровождения распределенных мероприятий*

В.Ю. БУДКОВ¹, М.В. ПРИЩЕПА², А.Л. РОНЖИН³, А.И. САВЕЛЬЕВ⁴

¹ 199178, РФ, г. Санкт-Петербург, 14 линия В.О., 39, Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург, к.т.н., e-mail: budkov@iias.spb.su

² 199178, РФ, г. Санкт-Петербург, 14 линия В.О., 39, Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург, к.т.н., e-mail: prischepa@iias.spb.su

³ 199178, РФ, г. Санкт-Петербург, 14 линия В.О., 39, Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург, к.т.н., e-mail: ronzhinal@iias.spb.su

⁴ 199178, РФ, г. Санкт-Петербург, 14 линия В.О., 39, Санкт-Петербургский институт информатики и автоматизации Российской академии наук, Санкт-Петербург, e-mail: saveliev@iias.spb.su

В данной работе рассмотрена архитектура, модули обработки данных и алгоритмы работы системы сопровождения распределенных мероприятий. Представленная система на основе анализа состояния аудио-, видео- и презентационного оборудования выбирает мультимедийный контент, необходимый удаленному участнику для восприятия хода дискуссии, а также формирует отчетные материалы по результатам мероприятия по заданным шаблонам. В статье описана архитектура разработанной системы, состоящей из 5 основных модулей, обрабатывающих аудиовизуальную информацию, и управляющего сервера. Приведен пример переключения режимов модулей и синхронизации аудио- и видеоконтента, отправляемого удаленному участнику, в зависимости от событий, поступающих с модулей аудиолокализации, видеомониторинга и модулей управления мультимедийными устройствами. Рассмотрена организация взаимодействия блоков обработки многоканальных аудиовизуальных сигналов, задействованных в системе сопровождения мероприятий для формирования мультимедийного отчета о мероприятии и истории поведения участников. Описана структура основной части реляционной БД, используемой для хранения данных, получаемых от различных программно-аппаратных модулей, и информации по проводимым мероприятиям и их участникам. Представлены результаты экспериментального тестирования, показавшие, что время отображения неверного контента составило около 3% от всего времени заседания. Разработанная система реализована в интеллектуальном зале СПИИРАН и применяется при проведении распределенных совещаний.

Ключевые слова: интеллектуальное пространство, сопровождение распределенных мероприятий, видеоконференцсвязь, дистанционное обучение, компьютерное зрение, аудиовизуальная обработка данных, многомодальные интерфейсы, веб-приложение

ВВЕДЕНИЕ

Организация совместной работы и обсуждений между территориально распределенными участниками с использованием мультимедийных Интернет-приложений и мобильных устройств становится реальной альтернативой традиционным совещаниям. Интернет-приложения для проведения телеконференций и дистанционного обучения E-meeting, E-lecture становятся всё более популярными при проведении распределенных мероприятий, вебинаров [1, 2, 3]. Такие системы позволяют сэкономить на транспортных расходах, использовать персонифицированную форму обучения, а также предоставляют удобные средства поиска и доступа к информации.

* Статья получена 06 мая 2013 г.

Работа выполнена при поддержке Министерства образования и науки Российской Федерации, проект № 7.559.2011, гос. рег. номер НИР 01201255056

Для выявления основных проблем существующих систем сопровождения распределенных мероприятий был проведен их сравнительный анализ по пяти типам характеристик: 1) входные модальности, используемые для анализа и записи поведения участников во время проведения; 2) основные типы выходных данных, которые могут быть использованы при взаимодействии с пользователем системы; 3) основные виды оборудования, которые могут использоваться при проведении распределенных мероприятий; 4) сервисы обработки аудиовизуальных данных, записанных на мероприятии; 5) дополнительные возможности систем сопровождения. Проанализированные системы Webinar.ru, Cisco WebEx, Openmeetings, WebHuddle, Meetecho обладают широкими функциональными возможностями в области телекоммуникаций, однако вопросам автоматической обработки речи, анализа поведения участников во время диалога и другим требованиям, предъявляемым к информационным системам сопровождения распределенных мероприятий, начали уделять внимание относительно недавно [4, 5, 6].

Поскольку системы видеоконференцсвязи требуют больших сетевых ресурсов даже для передачи видео между двумя участниками, то поддержка многопользовательских видеоконференций является крайне затруднительной задачей [7, 8, 9]. Для уменьшения объема данных, передаваемых в ходе видеоконференцсвязи, в работе [4] используется автоматический способ определения текущего говорящего и его потокам мультимедийных данных выставляется наибольший приоритет при передаче остальным участникам. Таким образом, идентификация, диаризация дикторов, а также другие методы обработки речи и анализа лица человека применяются для автоматизации телекоммуникационных сервисов [10, 11, 12].

1. АРХИТЕКТУРА СИСТЕМЫ СОПРОВОЖДЕНИЯ РАСПРЕДЕЛЕННЫХ МЕРОПРИЯТИЙ В ИНТЕЛЛЕКТУАЛЬНОМ ЗАЛЕ СОВЕЩАНИЙ

Разработанная архитектура системы веб-трансляции мероприятий (СВТМ) состоит из пяти основных программных комплексов и управляющего сервера. На рис. 1 представлены все шесть модулей, которые отмечены цифрами. Первый комплекс – система управления мультимедийным оборудованием (СУМО), которая объединяет модули, управляющие мультимедийными приложениями и оборудованием, служащим для записи поведения пользователей и отображения презентационных данных. Второй комплекс – многоканальная система обработки персональных веб-камер (МСОПВ), которая захватывает и обрабатывает аудио- и видеопотоки с камер. Третий комплекс служит для хранения аудио- и видеоданных с мероприятия. Четвертый комплекс представляет собой базу данных, которая включает в себя информацию о совещании. Номером шесть на рис. 1 отмечен сервер управления мероприятиями (СУМ), который получает и анализирует данные из всех других модулей и предоставляет информацию для веб-системы отображения (ВСО), которая отмечена как номер пять. Система ВСО включает модули, которые передают мультимедийный контент удаленным участникам. Система управления контентом (СУК) включает третий, пятый и шестой комплексы.

Первый комплекс СУМО отвечает за работу мультимедийных устройств. Система управления умной доской (СУД) позволяет пользователям использовать плазменную панель с сенсорным экраном для рисования и записи заметок. Система управления презентациями (СУП) отвечает за загрузку, отображение и переключение слайдов презентации. Многоканальная система аудиолокализации (МСА) дает информацию об аудиоактивности в умной комнате. Многоканальная система видеомониторинга (МСВ) отвечает за обработку и запись видеопотоков, поступающих от камер, направленных на аудиторию, ведущего и участников, сидящих в зоне стульев.

МСОПВ состоит из персональных модулей веб-камер ПМВ, которые управляют работой персональных веб-камер расположенных за конференц-столом, а также сервера управления ПМВ – СПМВ, который обрабатывает аудио- и видеоданные, поступающие с этих модулей.

Аудиофайлы в формате *wav* и видеофайлы в формате *avi*, которые были получены от персональных камер и обработаны СУМ (изменен формат, разрешение и имя файла), изображения с МСВ, СУП, СУД и ПМВ, сохраняются в файловом хранилище.

База данных мероприятия реализована с помощью сервера MySQL и включает две таблицы: (1) основные сведения обо всех запланированных мероприятиях; (2) информация о текущем совещании, которая включает в себя некоторые данные для системы отображения совещания.

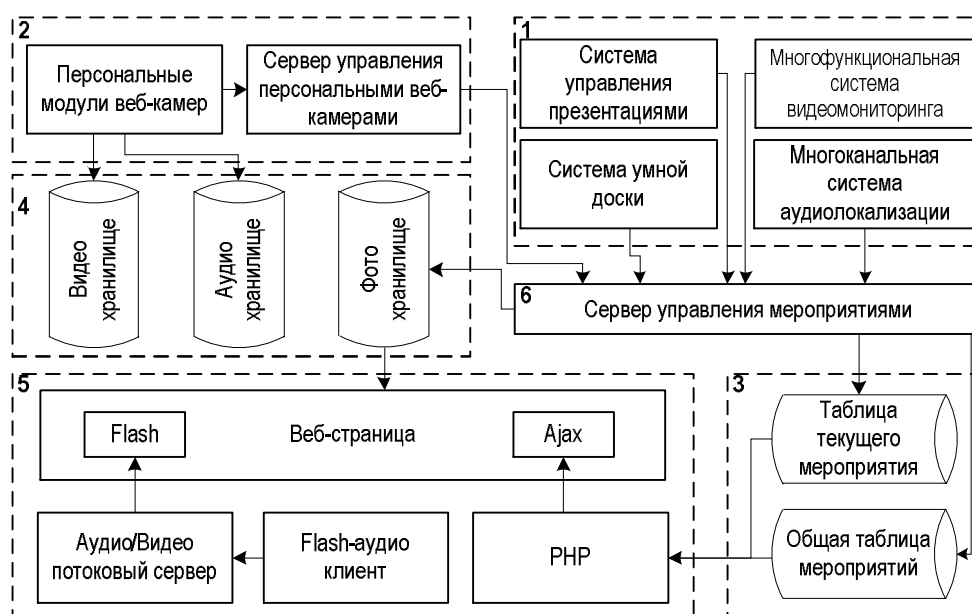


Рис. 1. Архитектура системы сопровождения распределенных мероприятий в интеллектуальном зале совещаний

ВСО представлена в виде веб-страницы с несколькими формами. Информация о наполнении этих форм получается на основе технологии AJAX. Передача аудиоданных реализована на основе потокового RTMP сервера и технологии Adobe Flash.

СУМ получает и анализирует данные от всех модулей, а также выбирает аудио- и видеоконтент для ВСО. Этот анализ основан на логико-временной модели [13].

Программные модули СВТМ были установлены на нескольких персональных компьютерах, объединенных в одну локальную сеть. Связь между модулями реализована на основе передачи строковых сообщений по UDP-протоколу.

Работа системы, отвечающей за веб-трансляцию и ее компонентов, зависит от текущей ситуации в комнате. На рис. 2 показан пример переключения режимов компонентов и синхронизации аудио- и видеоконтента в зависимости от событий, поступающих с модулей аудиолокализации, видеомониторинга и модулей управления мультимедийными устройствами. СУК управляет выбором мультимедийного контента, который доступен для удаленного участника совещания.

События, которые сгенерированы СУМ, влияют на работу системы веб-трансляции мероприятия. Эти события показаны в правой части рис. 2 по оси времени. Представленные события могут быть разделены на четыре типа по следующим критериям: (1) время; (2) деятельность основного докладчика; (3) деятельность сидящих участников; (4) использование презентационных устройств.

Первое событие, показанное на схеме, – «20 минут до встречи». После этого события (событие E1), система веб-трансляции переходит в режим подготовки и запускает модули МСА, МСВ и МСОПВ. Кроме того, система отправляет сообщения для опускания экрана и включения света. Логотип мероприятия отображается на веб-странице. После обнаружения аудиоактивности в комнате (событие E2) в модуле МСА запускается режим определения звуковой активности.

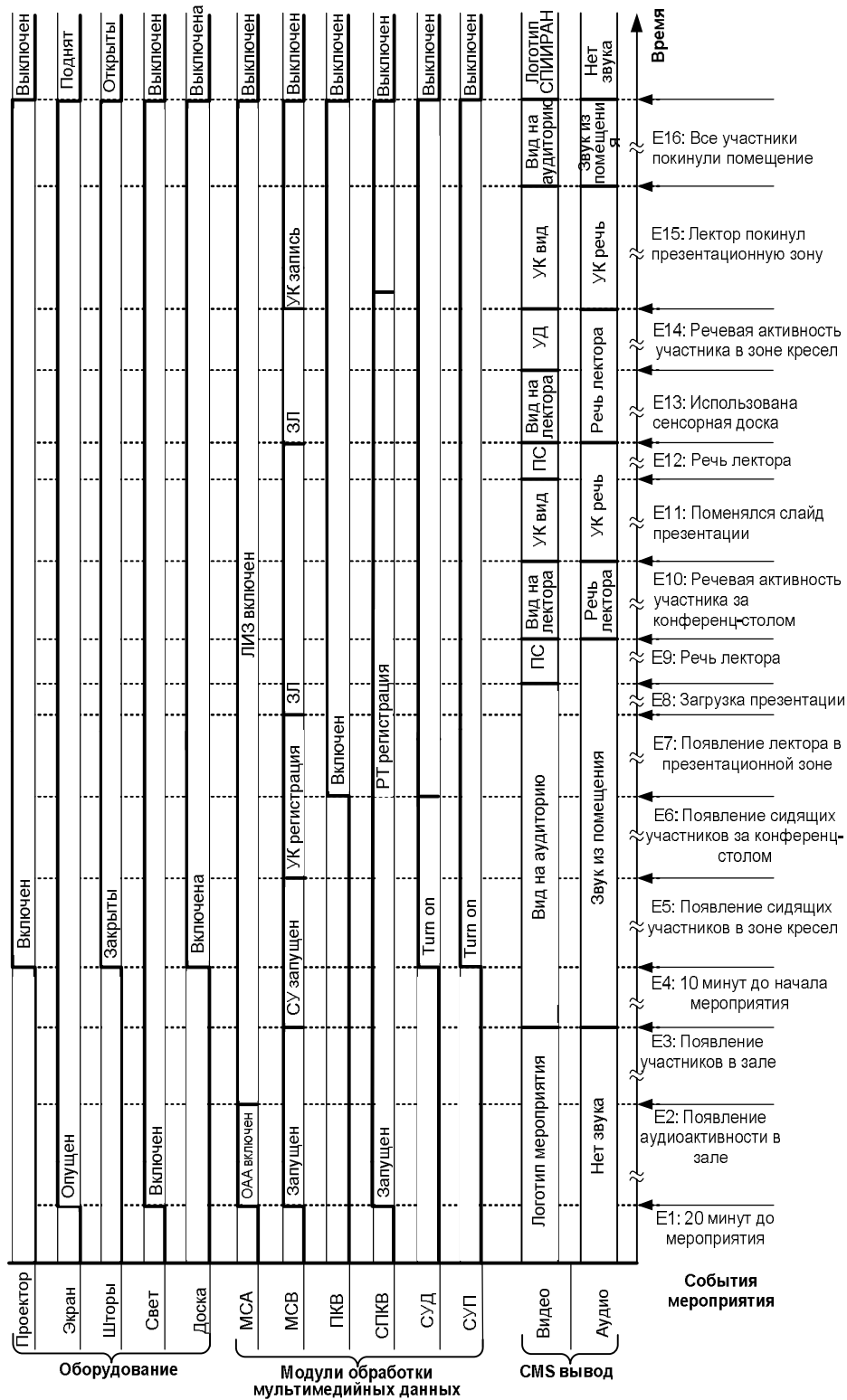


Рис. 2. Пример обработки событий, возникающих при проведении мероприятия

При появлении участника в комнате (событие E3) МСВ переходит в режим «слежения за участниками» (СУ) и на веб-страницу выводится изображение с общим видом комнаты, а также аудиопоток с микрофона, в котором анализируется все пространство комнаты. Когда до начала заседания остается около десяти минут (событие E4), включается проектор, плазменная панель с сенсорным экраном и модули СУД, РС.

Если появляются участники в зоне стульев (событие E5), МСВ начинает работать в режиме «регистрация участников» (РУ). Если есть участники, сидящие за конференц-столом (событие E6), запускаются модули ПМВ. Если МСВ обнаруживает основного докладчика (событие E7) и МСА обнаруживает его аудиодейтельность (события E9, событие E12), то кадры с камеры, направленной на него, отображаются на веб-странице.

Система начинает передачу аудиопотока с микрофона, с помощью которого производится анализ зоны презентаций. Когда презентация полностью загружена (событие E8), ее первый слайд отображается на веб-странице. Если аудиоактивность была обнаружена за конференц-столом (событие E10), то система получает аудиопоток с персональной веб-камеры, расположенной перед участником.

Когда слайд был изменен (событие E11), изображение с новым слайдом передается на веб-страницу. Если участник использует сенсорную доску (событие E13), то изображение слайдов на веб-странице изменяется на изображения с SBS. Если в зоне стульев была обнаружена аудиодейтельность (событие E14), то PTZ-камера фокусируется на говорящем участнике и MVPS записывает его речь. В то же время на веб-странице отображаются кадры с PTZ-камеры и записывается аудиопоток, полученный с микрофона в этой зоне. Когда основной докладчик покидает зону презентации (событие E15), для удаленных участников отображается общий вид комнаты.

Когда все участники совещания выходят из комнаты (событие E16), то все модули и устройства выключаются. После окончания совещания на веб-странице отображается только логотип.

Рассмотрим особенности функционирования и организацию взаимодействия блоков обработки многоканальных аудиовизуальных сигналов, задействованных в системе сопровождения мероприятий для формирования мультимедийного отчета о мероприятии и истории поведения участников.

В схеме взаимодействия, представленной на рис. 3, показано четыре типа используемого аудио- и видеозаписывающего оборудования: 1) персональные веб-камеры, установленные на конференц-столе; 2) отдельные микрофоны (МС), предназначенные для записи речи выступающих; 3) распределенная система видеокamer, состоящая из пяти камер, установленных на стенах и потолке интеллектуального зала; 4) массивы микрофонов (МА), применяемые для определения координат источника звука и выявления текущего выступающего или активного диктора. Считанные с сенсорного оснащения зала аудиовизуальные сигналы обрабатываются в трех блоках: 1) аудиовизуального мониторинга интеллектуального зала; 2) аудиовизуальной записи выступлений участников; 3) автоматической регистрации участников мероприятий. Далее рассмотрим назначение и состав каждого из перечисленных блоков.

Блок аудиовизуального мониторинга предназначен для бимодального анализа пространства зала для определения местоположения участников, слежения за их перемещением, определение сидящих участников, выявления текущего активного диктора за счет применения методов сегментации изображений, кадрирования областей изображения, детекции лиц участников, определения координат источника звука. В результате полученные данные передаются в блок структурирования, где происходит их объединение и оценка для выявления текущей ситуации в зале, которая в дальнейшем используется для настройки мультимедийного оснащения зала при помощи блока многомодального управления оборудованием интеллектуального зала.

Блок аудиовизуальной записи выступлений участников мероприятий включает в себя три подблока, показанных на рис. 3. Данное разделение связано с конфигурацией интеллектуального зала, где в правой части зала расположены ряды кресел, предназначенных для размещения 32 человек, а в левой части расположен коференц-стол, за которым могут сидеть до 10 участников, вследствие чего анализ пространства с сидящими участниками был разделен

между двумя типами устройств захвата видеосигнала – персональными веб-камерами и распределенной системой видеокамер. Сегментация аудиосигналов, записанных с отдельных микрофонов, производится за счет применения метода определения границ речи, при этом на каждую реплику участника может быть записано несколько аудиофайлов. Все записанные данные синхронизируются и объединяются в единые аудиовизуальные файлы, содержащие выступление участника, и записываются в базу данных.

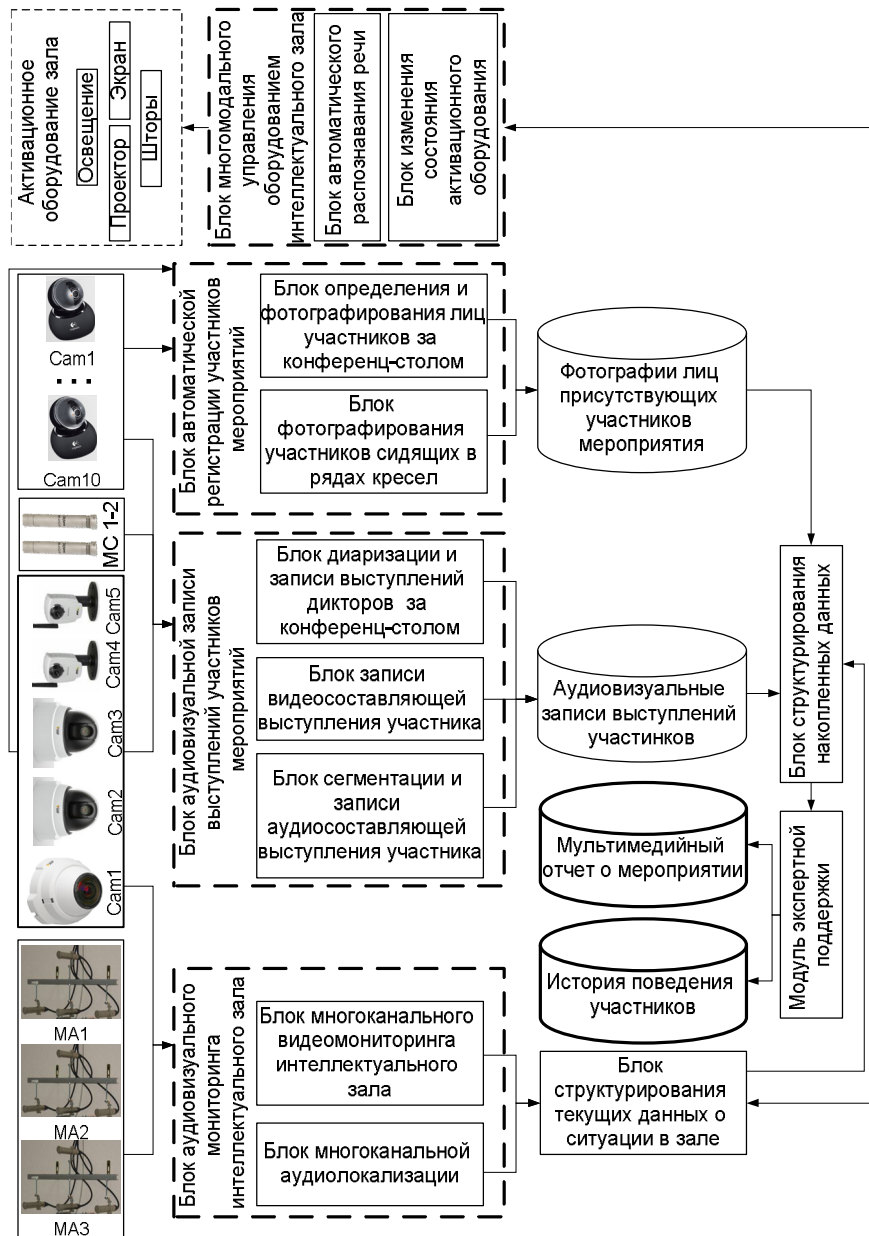


Рис. 3. Схема алгоритма взаимодействия блоков обработки аудиовизуальных данных при формировании мультимедийного отчета о мероприятии

Рассмотрим функционирование блока записи выступлений в трех различных ситуациях. Когда лектор (докладчик) рассказывает материал и на этом этапе мероприятия не возникает

дискуссии, аудиосоставляющая его выступление записывается при помощи отдельного микрофона, а видеоряд записывает с наведенной на него камеры с функциями наклона, поворота и масштабирования, входящей в состав распределенной системы камер. Такое же оборудование и операции считывания и обработки данных задействованы в ситуации, когда в процессе дискуссии активный участник, сидящий в одном из кресел, расположенных в правой части зала, задает вопрос. В ситуации, когда проводится совещание и все его участники расположены за конференц-столом или когда один из сидящих в одном из кресел, расположенных в левой части зала, задает вопрос, одновременно задействованы все три типа сенсорного оборудования. Подтверждения данных о говорящем участнике, выявленных при помощи устройств, задействованных в описанных выше ситуациях, производится за счет обработки аудиовизуальных данных, считанных с веб-камер, установленных на конференц-столе, при помощи метода определения и слежения за лицом участника и блока диаризации дикторов, который подробно описан в следующих разделах диссертационной работы.

Рассмотрим процессы обработки данных в блоке автоматической регистрации участников, сидящих за конференц-столом. Обнаружение присутствия участника выполняется за счет обработки потока видеоданных, считанных с веб-камер, основанной на определении наличия лица в кадре. В случае обнаружения производится формирование управляющей команды к камере с функциями наклона, поворота и масштабирования для ее наведения крупным планом на лицо участника и сохранение кадра. В результате, данные, полученные в ходе двухэтапного фотографирования, записываются в базу данных и в дальнейшем используются при формировании мультимедийного отчета о мероприятии.

Данные, поступающие с блока аудиовизуального мониторинга в блок многомодального управления оборудованием зала, применяются для настройки мультимедийного оснащения и изменения состояния активационных устройств в зависимости от поведения участников, их предпочтений и текущей ситуации.

По окончании мероприятия производится структуризации всех накопленных аудиовизуальных данных, в процессе которой выполняется присвоение данным идентификационного номера, соответствующего номеру участника, поведение которого они описывают. После окончания автоматической обработки данных осуществляется их экспертная проверка на наличие ошибок в процессе объединения и ошибок, возникших на этапе записи. Все данные, прошедшие проверку, далее используются для формирования мультимедийного отчета, а также для создания или обновления статистических данных, составляющих историю поведения участников на мероприятиях.

2. БАЗА ДАННЫХ СИСТЕМЫ СОПРОВОЖДЕНИЯ РАСПРЕДЕЛЕННЫХ МЕРОПРИЯТИЙ

Для хранения и последующей обработки данных, получаемых от различных программно-аппаратных модулей, была разработана структура реляционной базы данных, основные таблицы которой в виде ER-диаграммы показаны на рис. 4. Черная точка на конце связи между таблицами, обозначенная буквой Р, показывает отношение один ко многим.

В таблице «EventSection» хранится информация о секциях на проводимом мероприятии. В ней содержатся внешние ключи, ссылающиеся на мероприятие, расписание секции, расположение помещения, в котором проводится секция, ведущий данной секции, а также информация о начале и конце секции. Таблица «EventParticipants» содержит информацию об участниках данной секции и внешние ключи, ссылающиеся на секцию, тип участия в мероприятии, а также презентационные данные и положение в списке выступлений.

Полная информация о пользователях системы хранится в таблице «UserData». Она содержит данные об имени пользователя, его ученой степени, должности, а также контактные данные, место жительства и другую информацию. В таблицу «EventLog» заносятся записи о событиях, произошедших во время мероприятия, вместе с идентификатором устройства, которое сгенерировало данное событие. При наступлении какого-либо события в базу заносится запись, содержащая набор данных: тип события; его данные; устройство, от которого поступило событие; пользователь, который явился инициатором данного события. Таблица

«Device» содержит информацию об устройствах и видах передаваемых ими данных. Связь «многие ко многим» между таблицами «EventSection» и «Device» реализована через дополнительную таблицу. Таблица «RecordTypes» содержит информацию о типах записей. После проведения мероприятия данные из таблицы «EventLog» позволяют сгенерировать мультимедийный отчет и получить доступ к материалам, записанным во время его проведения. Генерация материалов производится по заранее сформированным шаблонам отчетов, в которые заносятся аудиовизуальные, графические, текстовые данные и статистические данные по событиям, автоматически накапливаемым в ходе совещания.

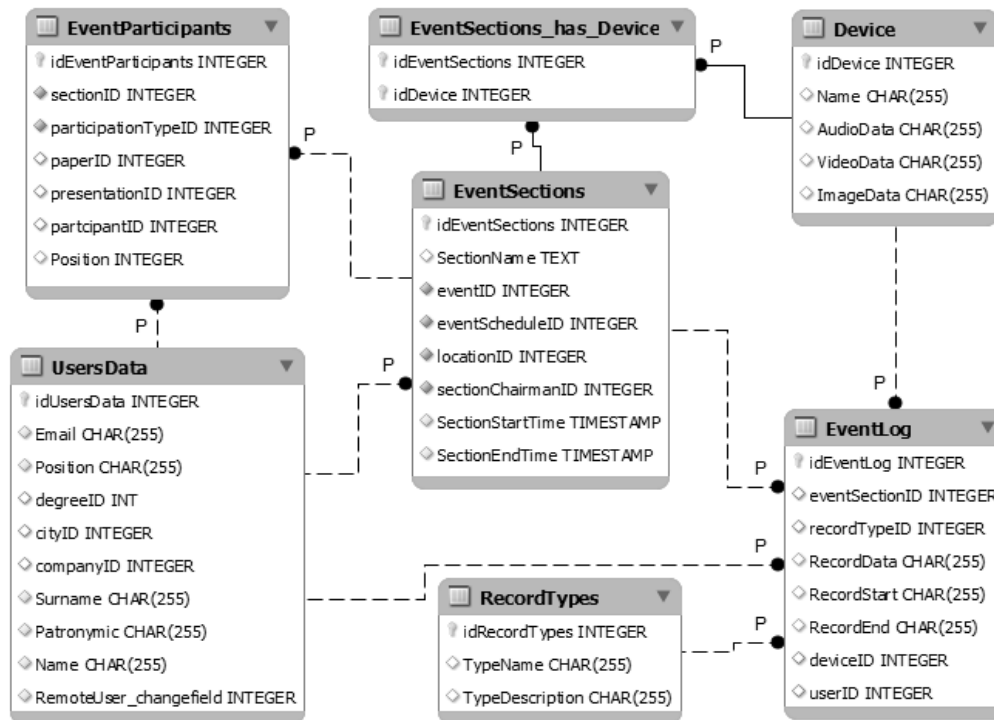


Рис. 4. ER-диаграмма данных мероприятий с персонализированной информацией по участникам

Разработанная структура базы данных используется при занесении данных по предстоящим мероприятиям, а также формировании протоколов и других отчетных материалов по прошедшим совещаниям. Далее в качестве примера рассмотрим один из вариантов подготовки протокола совещания. Ниже приведен фрагмент шаблона, обеспечивающий вывод данных по зарегистрированному участнику в протокол совещания в формате HTML. Данные по участнику включают: фотографию, имя, должность, e-mail, город проживания, организацию.

```

    {% for user in user_data %}
    <div class='event_participant'>
        <img src='img/user/ {{ user.img_name }}.jpg' alt='{{ user.name }}'></img>
        <span><div class='event_participant_data' >
            Имя: {{ user.name }}<br>
            Должность: {{ user.position }}<br>
            E-mail: {{ user.email }}<br>
            Город проживания: {{ user.city }}<br>
            Организация: {{ user.affil }}
        </div>
    </div>
    <br>
    {% endfor %}
    
```


Перечисленная структура данных выводится в протокол по каждому участнику. В начале протокола приводятся данные по мероприятию: организации, время проведения мероприятия, место проведения мероприятия, наименование мероприятия, председатель мероприятия:

```
<div class='event_prot_name'>ПРОТОКОЛ</div>
<div class='event_date'>{{ event.date }}</div>
<div class='aff_event_name'>{{ event.aff }}</div>
<div class='event_place'>{{ event.place }}</div>
<div class='event_name'>{{ event.name }}</div>
<div class='event_chairman'>Председатель: {{ event.chairname }}</div>
```

Кроме приведенных в протоколе по мероприятию данных может быть добавлена и другая информация, содержащаяся в структуре, разработанной базы данных. Редактирование информации по пользователям системы сопровождения распределенных мероприятий, мероприятиям и их участникам производится с помощью разработанного веб-интерфейса. Изменение структуры базы данных производится с использованием приложений, поддерживающих СУБД MySQL. При этом редактирование информации о пользователях может проводиться как самими пользователями, так и администраторами, а редактирование данных по мероприятию может производиться имеющими разрешение пользователями.

3. ЭКСПЕРИМЕНТЫ

Экспериментальные результаты были получены в результате проведения сценария, где несколько людей обсуждают проблемы в умной комнате в СПИИРАНе. Один из участников стоял в презентационной зоне и использовал сенсорную доску и проектор. Остальные участники сидели за конференц-столом. Выступающий начинал речь, когда все участники находились в зале. Каждый из участников мог задавать вопросы после окончания презентации.

Данные по мероприятию, которые влияли на изменение ситуации и графического контента, представлены в табл. 1. Во время регистрации участников весь графический контент жестко задан, и поэтому ошибок в выборе не было. Изменение состояний сенсорной панели и переключение слайдов были определены верно. Большая часть ошибок возникла при определении речевой активности участников, сидящих за конференц-столом. Эти ошибки приводили к неверному выбору изображения с камеры, а также источника звука, что приводило к понижению уровня сигнала.

Таблица 1

Результаты эксперимента

Описания мероприятия	Определено вручную	Определено автоматически		
		Число	FA	MS
Участников	5	5	0	0
Участники, сидящие за конференц-столом	4	4	0	0
Изменений слайда	22	22	0	0
Использований сенсорного экрана	1	1	0	0
Речевая активность главного диктора	10	15	5	0
Речевая активность сидящих участников	9	32	24	1
Временное отсутствие активности в аудитории	2	2	0	0
Переход реплики от выступающего к сидящим участникам	8	13	6	1
Переход реплики от сидящих участников к выступающему	7	13	6	0
Переход реплики между выступающими	1	18	17	0
Всего переходов реплик	16	44	29	2

Результаты эксперимента показывают, что большинство ошибок (ложное срабатывание FA и пропуск сегментов MS) сделаны алгоритмом выявления активного диктора. Такие ошибки происходят, когда участник на конференции задает вопрос, но изображение другого участника, который находится рядом, показывается в диалоговом окне активного диктора. Такие ошибки происходят из-за ошибок, сделанных SSL (Sound Source Localization) по причине высокого уровня реверберации в зале. Кроме того, небольшое расстояние между участниками, сидящими за конференц-столом, увеличивает число ошибок, потому что погрешность в SSL для комнаты – 0.5m. В то же время количество ошибок при переключении между активным участником и ведущим меньше, так как расстояние больше, чем между участниками за конференц-столом.

Кроме того, паузы в речевой деятельности основного докладчика приводят к переключению камеры на аудиторию или другого участника, чья речь была неправильно обнаружена. В общей сложности около 37 % графического содержания были правильно выбраны при анализе текущей ситуации в конференц-зале, но время отображения ложных активных участников составляло около 3 % от всего времени заседания. На данный момент разработанная модель веб-страницы была проверена в режиме, когда удаленные участники являлись только слушателями и не могли влиять на ход мероприятия. Для повышения активности таких участников будут разработаны панели инструментов, позволяющих удаленному слушателю задавать вопросы и участвовать в дискуссии.

ЗАКЛЮЧЕНИЕ

Проанализированные системы сопровождения веб-конференций обладают широкими функциональными возможностями в области телекоммуникаций, однако вопросам автоматической обработки речи, анализа поведения участников во время диалога и другим требованиям, предъявляемым к информационным системам сопровождения распределенных мероприятий, уделено недостаточно внимания. Представлена архитектура разработанной системы сопровождения распределенных мероприятий, формирующая на основе анализа состояния аудио-, видео- и презентационного оборудования мультимедийный контент, необходимый удаленному участнику для восприятия хода дискуссии, и применяющая персонализированную базу данных мероприятий при подготовке отчетных материалов по заданным шаблонам. Описана структура реляционной базы данных для хранения данных, получаемых от различных программно-аппаратных модулей, и информации по проводимым мероприятиям и их участникам. Разработанные средства реализованы в интеллектуальном зале СПИИРАН и проходят апробацию при проведении распределенных совещаний по международным и российским проектам.

СПИСОК ЛИТЕРАТУРЫ

1. *Erol B., Li Y.* An overview of technologies for e-meeting and e-lecture // Proceedings IEEE International Conference on Multimedia and Expo. – 2005. – P. 6–12.
2. *Cicco L., Mascolo S, Palmisano V.* Skype Video congestion control: An experimental investigation // Computer Networks. – 2011. – Vol. 55, iss. 3. – P. 558–571.
3. Разработка веб-системы для предоставления обучающих сервисов удаленным мобильным пользователям / Д. Ганбат, А.Л. Ронжин, Р. Найдандорж, В.Ю. Будков, М.В. Прищепа // Труды СПИИРАН. – 2010. – Вып. 1 (12). – С. 21–34.
4. *Volfin I., Cohen I.* Dominant speaker identification for multipoint videoconferencing // Computer Speech and Language. – 2013. – Vol. 27, iss. 4. – P. 895–910.
5. *Басов О.О., Носов М.В., Шалагинов В.А.* Исследование характеристик джиттера периода основного тона речевого сигнала // Труды СПИИРАН. – 2014. – Вып. 1 (32). – С. 27–44.
6. *Ронжин А.Л., Будков В.Ю.* Технологии поддержки гибридных е-совещаний на основе методов аудиовизуальной обработки // Вестник компьютерных и информационных технологий. – 2011. – № 4. – С. 31–35.
7. Redesigning multi-channel P2P live video systems with View-Upload Decoupling / D. Wu, C. Liang, Y. Liu, K.W. Ross // Computer Networks. – 2010. – Vol. 54, iss. 12. – P. 2007–2018.
8. *Савельев А.И., Прищепа М.В.* Архитектура обмена данными без потерь в пиринговом веб-приложении видеоконференц-связи // Доклады Томского государственного университета систем управления и радиоэлектроники. – 2014. – № 2 (32). – С. 238–245.

9. Савельев А.И. Оптимизация алгоритмов распределения потоков мультимедийных данных между сервером и клиентом в приложениях видеоконференцсвязи // Труды СПИИРАН. – 2013. – Вып. 8 (31). – С. 61–79.
10. Мецераков Р.В. Структура систем синтеза и распознавания речи // Известия Томского политехнического университета. – 2009. – Т. 315, № 5. – С. 127–132.
11. Ронжин А.Л. Топологические особенности морфофонемного способа представления словаря для распознавания русской речи // Вестник компьютерных и информационных технологий. – 2008. – № 9. – С. 12–19.
12. Тиунов С.Д., Мецераков Р.В., Черных Д.В. Оптимизация вычисления одновременной маскировки речевого сигнала // Труды СПИИРАН. – 2014. – Вып. 1 (32). – С. 45–57.
13. Ronzhin A., Budkov V., Karpov A. Multichannel System of Audio-Visual Support of Remote Mobile Participant at E-Meeting // Smart Spaces and Next Generation Wired/Wireless Networking. Third Conference on Smart Spaces, ruSMART 2010, and 10th International Conference, NEW2AN 2010, St. Petersburg, Russia, August 23-25, 2010: Proceedings. – Berlin, Heidelberg, Springer-Verlag, 2010. – P. 62–71. – (Lecture Notes in Computer Science; vol. 6294). – DOI: 10.1007/978-3-642-14891-0_6.

Будков Виктор Юрьевич, кандидат технических наук, научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации Российской академии наук. Имеет более 40 публикаций. E-mail: budkov@iias.spb.su

Прищепа Мария Викторовна, кандидат технических наук, научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации Российской академии наук. Имеет более 40 публикаций. E-mail: prischepa@iias.spb.su

Ронжин Александр Леонидович, кандидат технических наук, научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации Российской академии наук. Имеет более 40 публикаций. E-mail: ronzhinal@iias.spb.su

Савельев Антон Игоревич, младший научный сотрудник лаборатории речевых и многомодальных интерфейсов Санкт-Петербургского института информатики и автоматизации Российской академии наук. Имеет 7 публикаций. E-mail: saveliev@iias.spb.su

*Architecture and software of support system for distributed events**

V.Yu. BUDKOV¹, M.V. PRISHCHEPA², A.L. RONZHIN³, A.I. SAVELIEV⁴

¹ St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14-th Line V.O., St. Petersburg, 199178, Russia, Ph.D., e-mail: budkov@iias.spb.su

² St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14-th Line V.O., St. Petersburg, 199178, Russia, Ph.D., e-mail: prischepa@iias.spb.su

³ St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14-th Line V.O., St. Petersburg, 199178, Russia, Ph.D., e-mail: ronzhinal@iias.spb.su

⁴ St.Petersburg Institute for Informatics and Automation of the Russian Academy of Sciences, 39, 14-th Line V.O., St. Petersburg, 199178, Russia, e-mail: saveliev@iias.spb.su

This paper presents the architecture, data processing modules as well as work algorithms of the developed system of distributed event support. The developed system is based on audio, video and presentation equipment state analysis choices a multimedia context for remote participants for discussion understanding. In addition, the system forms report data of meeting results by defined templates. The paper presents the architecture of control server and developed system, which includes five main modules for audiovisual data processing. The example of modules work mode changing and synchronization of audio and video content for transmitting remote participant, depended on receiving events from sound source localization, videomonitoring and multimedia devices control modules, is described. The organization of interaction between modules for formation of meeting multimedia report and history of participants behavior is considered. The structure of relational database main part, which is used for storage data from soft and hardware modules as well as information about carried out meetings and participants, is described. The structure of a relational database, which is used for store of data from various hardware and software modules, and information on the events and their participants, is considered. The experimental results show that time for displaying of a wrong content was about 3 % at whole time of meeting. The developed system is implemented in a smart room of SPIIRAS and is used during distributed meetings.

Keywords: intelligent environments, distributed event support, videoconferencing, distant learning, computer vision, audiovisual data processing, multimodal interfaces, web applications

* Received 06 May 2013.

The work was supported by the Ministry of education and science of the Russian Federation, project no 7.559.2011, state registration number of scientific research works 01201255056

REFERENCES

1. Erol B., Li Y. An overview of technologies for e-meeting and e-lecture. Proceedings IEEE International Conference on Multimedia and Expo. 2005. P. 6-12.
2. Cicco L., Mascolo S., Palmisano V. Skype Video congestion control: An experimental investigation. Computer Networks, 2011, vol. 55, iss. 3, pp. 558-571.
3. Ganbat D., Ronzhin A.L., Naidandorzh R., Budkov V.Yu., Prishchepa M.V. Razrabotka veb-sistemy dlya predstavleniya obuchayushchikh servisov udalennym mobil'nym pol'zovatelyam [Development of web-system for providing learning services for remote mobile users]. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2010, iss. 1 (12), pp. 21-34.
4. Volfin I., Cohen I. Dominant speaker identification for multipoint videoconferencing. Computer Speech and Language, 2013, vol. 27, iss. 4, pp. 895-910.
5. Basov O.O., Nosov M.V., Shalaginov V.A. Issledovanie kharakteristik dzhittera perioda osnovnogo tona rechevogo signala [Pitch-jitter analysis of the speech signal]. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2014, iss. 1 (32), pp. 27-34.
6. Ronzhin A.L., Budkov V.Yu. Tekhnologii podderzhki gibridnykh e-soveshchani na osnove metodov audiovizual'noi obrabotki [Support Technologies of E-Meetings Based on Methods for Audiovisual Processing]. *Vestnik komp'yuternykh i informatsionnykh tekhnologii – Herald of computer and information technologies*, 2011, no. 4, pp. 31-35.
7. Wu D., Liang C., Liu Y., Ross K.W. Redesigning multi-channel P2P live video systems with View-Upload Decoupling. Computer Networks, 2010, vol. 54, iss. 12, pp. 2007-2018.
8. Savel'ev A.I., Prishchepa M.V. Arkhitektura obmena dannymi bez poter' v piringovom veb-prilozhenii videokonferents-svyazi [Architecture of lossless data exchange in peer-to-peer web application of videoconference]. *Doklady Tomskogo gosudarstvennogo universiteta sistem upravleniya i radioelektroniki – Proceedings of Tomsk State University of Control Systems and Radioelectronics*, 2014, no. 2 (32), pp. 238-245.
9. Savel'ev A.I. Optimizatsiya algoritmov raspredeleniya potokov mul'timediinykh dannykh mezhdru serverom i klientom v prilozheniyakh videokonferents-svyazi [Optimization algorithms distribution streams of multimedia data between server and client in videoconferencing application]. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2013, iss. 8 (31), pp. 61-79.
10. Meshcheryakov R.V. Struktura sistem sinteza i raspoznavaniya rechi [Structure of speech synthesis and recognition systems]. *Izvestiya Tomskogo politekhnicheskogo universiteta – Bulletin of the Tomsk Polytechnic University*, 2009, vol. 315, no. 5, pp. 127-132.
11. Ronzhin A.L. Topologicheskie osobennosti morf fonemnogo sposoba predstavleniya slovary dlya raspoznavaniya russkoi rechi [Topological features of a morfofonemny way of submission of the dictionary for recognition of the Russian speech]. *Vestnik komp'yuternykh i informatsionnykh tekhnologii – Herald of computer and information technologies*, 2008, no. 9, pp. 12-19.
12. Tiunov S.D., Meshcheryakov R.V., Chernykh D.V. Optimizatsiya vychisleniya odnovremennoi maskirovki rechevogo signala [Optimization of calculation of frequency masking of speech signal]. *Trudy SPIIRAN – SPIIRAS Proceedings*, 2014, iss. 1 (32), pp. 45-57.
13. Ronzhin A., Budkov V., Karpov A. Multichannel System of Audio-Visual Support of Remote Mobile Participant at E-Meeting. Smart Spaces and Next Generation Wired/Wireless Networking. Third Conference on Smart Spaces, ruSMART 2010, and 10th International Conference, NEW2AN 2010, St. Petersburg, Russia, August 23-25, 2010. Proceedings. Berlin, Heidelberg, Springer-Verlag, 2010. Lecture Notes in Computer Science, 2010, vol. 6294, pp. 62-71. DOI: 10.1007/978-3-642-14891-0_6