



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21), (22) Заявка: 2006102661/09, 10.05.2005

(24) Дата начала отсчета срока действия патента:
10.05.2005(30) Конвенционный приоритет:
18.05.2004 EP 04291272.5

(43) Дата публикации заявки: 10.08.2007

(45) Опубликовано: 27.05.2010 Бюл. № 15

(56) Список документов, цитированных в отчете о
поиске: WO 02079896 A2, 10.10.2002. RU 2188450
C2, 27.08.2002. EP 1304614 A2, 23.04.2003. US
2003126330 A1, 03.07.2003.(85) Дата перевода заявки РСТ на национальную
фазу: 30.01.2006(86) Заявка РСТ:
EP 2005/052110 (10.05.2005)(87) Публикация РСТ:
WO 2005/112386 (24.11.2005)

Адрес для переписки:
129090, Москва, ул. Б.Спасская, 25, стр.3,
ООО "Юридическая фирма Городиский и
Партнеры", пат.пов. Ю.Д.Кузнецову,
рег.№ 595

(72) Автор(ы):

ЗИНЕЛ Юрген (DE),
РЕССЛЕР Хорст (DE),
НОЙБАУЭР Даниель (DE)

(73) Патентообладатель(и):

АЛЬКАТЕЛЬ (FR)

(54) СПОСОБ И СЕРВЕР ДЛЯ ОБЕСПЕЧЕНИЯ МУЛЬТИМОДАЛЬНОГО ДИАЛОГА

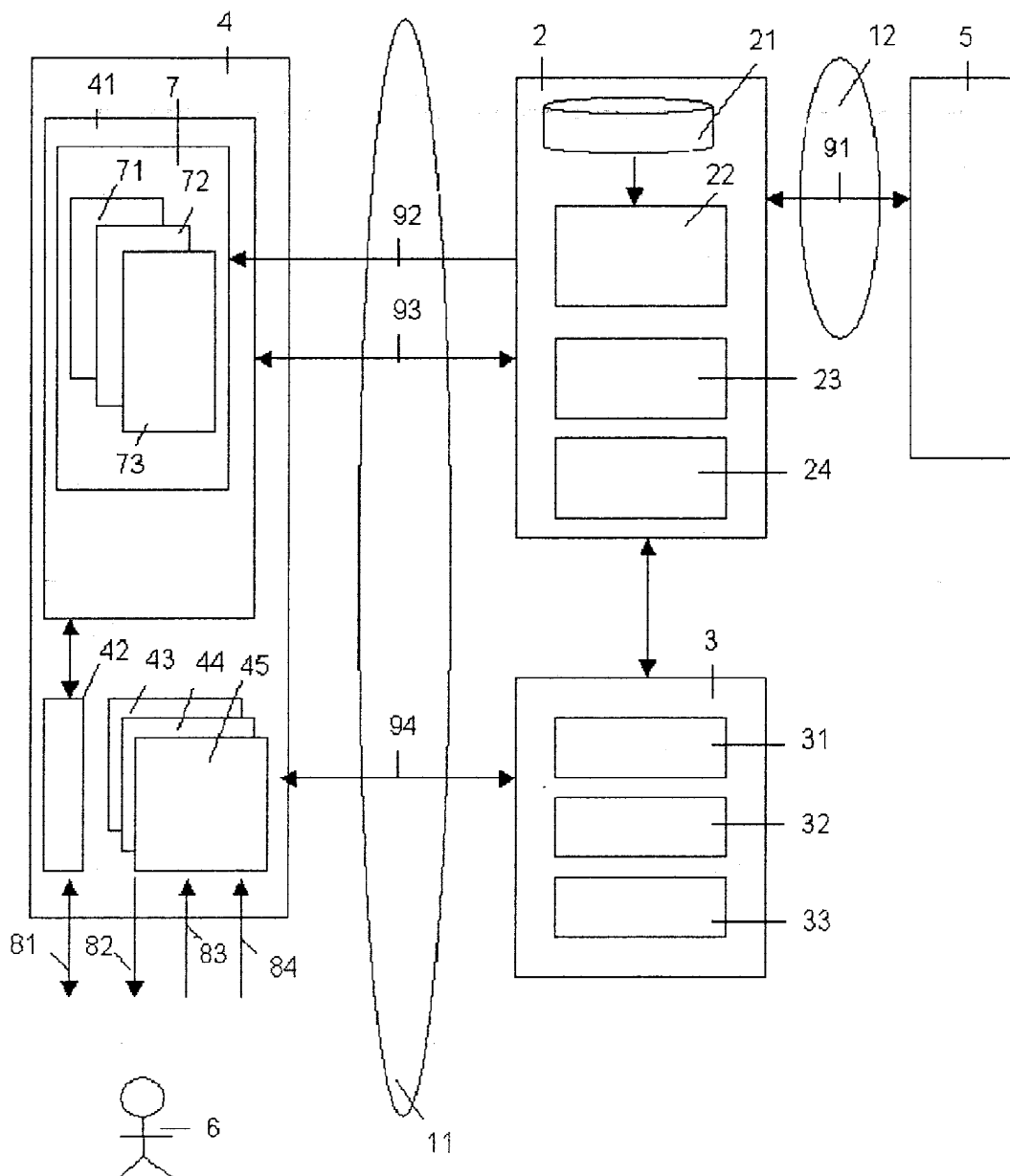
(57) Реферат:

Изобретение относится к системам передачи данных. Технический результат заключается в усовершенствовании установления мультимодального диалога. Изобретение описывает способ обеспечения мультимодального диалога между мультимодальным приложением (5) и пользователем (6), осуществляющим связь с мультимодальным приложением (5) через клиент (4), выполненный с возможностью обмена документами и представления документов (7), закодированных в

стандартном или расширенном языке гипертекстовой разметки. Изобретение дополнительно описывает прокси-сервер (2) для осуществления этого способа. Мультимодальный диалог между мультимодальным приложением (5) и пользователем (6) устанавливается через прокси-сервер (2), взаимодействующий с клиентом (4) посредством обмена информацией, закодированной в стандартном или расширенном языке гипертекстовой разметки. Прокси-сервер (2) извлекает по меньшей мере один дополнительный ресурс

модальности, запрошенный в мультимодальном диалоге. Прокси-сервер (2) компонует мультимодальное взаимодействие с пользователем (6) на основе взаимодействий с клиентом (6), основывающихся на стандартном

или расширенном языке гипертекстовой разметки, и на основе извлеченных дополнительных ресурсов (71, 72, 73, 31, 32, 33). 2 н. и 8 з.п. ф-лы, 2 ил.



ФИГ. 1

RU 2390958 C2

RU 2390958 C2



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY,
PATENTS AND TRADEMARKS

(12) ABSTRACT OF INVENTION

(21), (22) Application: **2006102661/09, 10.05.2005**
 (24) Effective date for property rights:
10.05.2005
 (30) Priority:
18.05.2004 EP 04291272.5
 (43) Application published: **10.08.2007**
 (45) Date of publication: **27.05.2010 Bull. 15**
 (85) Commencement of national phase: **30.01.2006**
 (86) PCT application:
EP 2005/052110 (10.05.2005)
 (87) PCT publication:
WO 2005/112386 (24.11.2005)
 Mail address:
129090, Moskva, ul. B.Spasskaja, 25, str.3, OOO
"Juridicheskaja firma Gorodisskij i Partnery",
pat.pov. Ju.D.Kuznetsovu, reg.№ 595

(72) Inventor(s):
ZINEL Jurgen (DE),
RESSLER Khorst (DE),
NOJBAUEhR Daniel' (DE)
 (73) Proprietor(s):
AL'KATEL' (FR)

(54) METHOD AND SERVER FOR PROVIDING MULTIMODE DIALOGUE

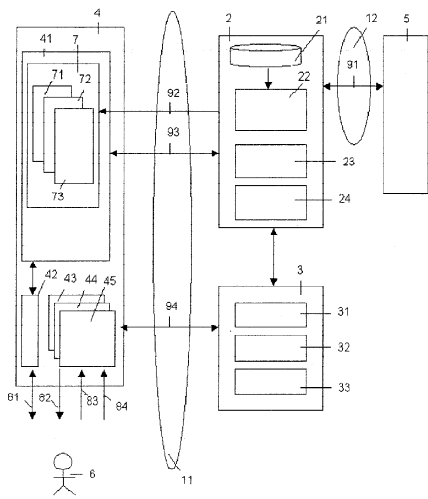
(57) Abstract:
 FIELD: information technology.
 SUBSTANCE: invention describes a method of providing multimode dialogue between a multimode application (5) and a user (6), communicating with the multimode application (5) through a client (4), made with possibility of exchanging documents and presenting documents (7), encoded in standard or extensible hypertext markup language. The invention also describes a proxy server (2) for implementing the method. Multimode dialogue between the multimode application (5) and the user (6) is established through the proxy server (2), which

interacts with the client (4) by exchanging information encoded in standard or extensible hypertext markup language. The proxy server (2) extracts at least one additional modality resource requested in the multimode dialogue. The proxy server (2) arranges multimode interaction with the user (6) based on interactions with the client (6), based on standard or extensible hypertext markup language and based on the extracted additional resources (71, 72, 73, 31, 32, 33).

EFFECT: improved establishment of multimode dialogue.
 10 cl, 2 dwg

RU 2 390 958 C2

RU 2 390 958 C2



ФИГ. 1

Область техники, к которой относится изобретение

Настоящее изобретение относится к способу обеспечения мультимодального диалога между мультимодальным приложением и пользователем, осуществляющим связь с мультимодальным приложением через клиент, выполненный с возможностью обмена документами и представления документов, закодированных в стандартном или расширенном языке гипертекстовой разметки. Более того, изобретение относится к прокси-серверу (специализированному серверу-посреднику) для поддержки мультимодальных диалогов между мультимодальными приложениями и пользователями.

Предшествующий уровень техники

В последние годы компьютеры стали снабжаться множеством устройств ввода разных типов, таких как клавиатура, мышь, сенсорная панель, сканер изображений, видеокамера, перо и микрофон, для обеспечения возможности ввода различных информационных элементов в различных формах. Также предусмотрено множество разных типов устройств вывода, такие как разнообразные устройства отображения и громкоговоритель, для вывода различных информационных элементов в различных формах, таких как разные графические формы или разговорный язык. Дополнительно усовершенствованные терминалы связи оборудованы разными типами устройств ввода и вывода, которые позволяют вводить и выводить информацию в различных формах.

Например, JP 101 07877 А описывает мультимодальный телефонный аппарат, который использует и дисплей, и синтезированный голос для взаимодействия с пользователем.

Дополнительно мультимодальные средства просмотра (броузеры) расширяют функциональные возможности исходных броузеров с помощью дополнительных модулей, таких как мультимодальные подключаемые расширения, например, SALT или X+V (SALT =..., X+V =...). Но подобные мультимодальные броузеры требуют больших вычислений и ресурсов памяти и не выполняются на малых устройствах с ограниченными ресурсами, например мобильных устройствах или PDA (PDA = персональный цифровой ассистент).

Сущность изобретения

Задачей настоящего изобретения является предложение улучшенной мультимодальной системы, делающей возможным мультимодальный диалог между мультимодальным приложением и пользователем.

Задача настоящего изобретения решается посредством способа обеспечения мультимодального диалога между мультимодальным приложением и пользователем, осуществляющим связь с мультимодальным приложением через клиент, выполненный с возможностью обмена документами и представления документов, закодированных в стандартном или расширенном языке гипертекстовой разметки, при этом способ содержит этапы, на которых: устанавливается мультимодальный диалог между мультимодальным приложением и пользователем через прокси-сервер, взаимодействующий с клиентом посредством обмена информацией, закодированной в стандартном или расширенном языке гипертекстовой разметки; извлекают с помощью прокси-сервера по меньшей мере один дополнительный ресурс модальности, запрошенный в мультимодальном диалоге; и komponуют с помощью прокси-сервера мультимодальное взаимодействие с пользователем на основе взаимодействий с клиентом, основывающихся на стандартном или расширенном языке гипертекстовой разметки, и на основе извлеченных дополнительных ресурсов. Задача настоящего

изобретения также решается с помощью прокси-сервера для поддержки мультимодальных диалогов между мультимодальными приложениями и пользователями, осуществляющими связь с мультимодальным приложением через соответствующие клиенты, выполненные с возможностью обмена документами и представления документов, закодированных в стандартном или расширенном языке гипертекстовой разметки, при этом прокси-сервер содержит менеджер (средство управления) диалогов и компоновщик. Менеджер диалогов устанавливает мультимодальный диалог между мультимодальным приложением и пользователем через прокси-сервер, взаимодействует с клиентом посредством обмена информацией, закодированной в стандартном или расширенном языке гипертекстовой разметки, и извлекает по меньшей мере один дополнительный ресурс модальности, запрошенный в мультимодальном диалоге. Компоновщик компоует мультимодальное взаимодействие с пользователем на основе взаимодействий с клиентом, основанных на стандартном или расширенном языке гипертекстовой разметки, и на основе извлеченных дополнительных ресурсов.

Мультимодальный прокси-сервер устанавливает мультимодальные взаимодействия на основе стандартной технологии браузера, например, на многозвенных web-приложениях, основанных на браузере. Компоненты прокси-сервера осуществляют на стороне клиента управление и синхронизацию в отношении мультимодального диалога, которые не могут быть выполнены компонентами клиента. Эта архитектура объединяет высокую скорость выполнения с уменьшением требований к вычислительной мощности и памяти на стороне клиента. Эта архитектура поддерживает использование тонких клиентов. Дополнительно эта архитектура увеличивает гибкость и применимость мультимодальных приложений и функциональных возможностей, так как является достаточным оснастить клиент стандартными компонентами браузера, чтобы использовать этот клиент в контексте мультимодального диалога.

Дополнительно архитектура тонкого клиента превращает настольные/локальные приложения в среду с централизованной обработкой, управлением и поддержкой. Таким образом, изобретение увеличивает масштабируемость и обслуживаемость мультимодальной системы. Изобретение обеспечивает возможность как масштабирования при изменении потребностей бизнеса, так и приспособления к увеличению количества пользователей, объемов транзакций и архивов - и все это без ухудшения производительности. Дополнительно это гарантирует сосуществование с унаследованными приложениями.

Дополнительные преимущества достигаются с помощью вариантов осуществления изобретения, указанных в зависимых пунктах.

В соответствии с предпочтительным вариантом осуществления изобретения мультимодальная система содержит клиента со стандартным Интернет-браузером без дополнений, например, без специфических встраиваемых модулей, мультимодальный прокси-сервер, ответственный за управление мультимодальными сеансом и диалогом, и мультимодальный сервер, который содержит несколько ресурсных модулей, таких как модули распознавания речи, преобразования текста в речь и распознавания рукописного ввода. Прокси-сервер отправляет/получает сообщения к/от мультимодального сервера. Дополнительно он передает HTTP-запросы (HTTP = протокол передачи гипертекста) соответствующим приложениям и web-серверам. Прокси-сервер также генерирует JAVA-скрипты (программы-сценарии на языке JAVA) для встраивания обработки мультимодальных событий в HTML/XHTML-

страницы (HTML = язык гипертекстовой разметки; XHTML = расширенный язык гипертекстовой разметки). Мультимодальный сервер содержит средства для установления RTP-соединений (RTP = протокол реального времени) с устройствами ввода/вывода клиента, например с микрофоном и громкоговорителем клиента.

Мультимодальный диалог компонуется с помощью JAVA-скриптов.

Мультимодальное взаимодействие может быть применено к HTML-элементам, таким как поля ввода текста, списки, кнопки, ссылки и так далее.

Предпочтительно клиент является мобильным устройством, например, мобильным телефоном в соответствии с GSM или UMTS стандартом (GSM = глобальная система мобильной связи; UMTS = универсальная мобильная телекоммуникационная система). Изобретение дает возможность такому мобильному устройству, оборудованному стандартным Интернет-браузером, поддерживать мультимодальное взаимодействие. Это имеет значительные преимущества: пользователь может выбрать в любое время предпочтительную модальность взаимодействия и не является привязанным к последовательности представления конкретного канала. Таким образом, взаимодействие становится персональным и повышает уровень восприятия пользователя.

В соответствии с предпочтительным вариантом осуществления изобретения прокси-сервер компонует упомянутые один или более запрошенных ресурсов в документ, представленный в стандартном или расширенном языке гипертекстовой разметки, и передает документ клиенту. Ресурс, переданный на сторону клиента с помощью такого механизма, поддерживает части мультимодального взаимодействия, которые не поддерживаются функциональными возможностями стандартного web-браузера. Дополнительные ресурсы модальности становятся доступными простым и эффективным способом.

Предпочтительно прокси-сервер создает скрипт, обеспечивающий ресурс модальности, и компонует скрипт в документ. Например, скрипт может инициировать вывод определенного голосового сообщения. Но также является возможным, что прокси-сервер создает скрипт для встраивания обработки мультимодального события в документ и компонует подобный скрипт в документ. Предпочтительно прокси-сервер дополнительно создает соответствующий обработчик ресурса, обеспечивающий для скрипта интерфейс с прокси-сервером. Обмен сообщениями осуществляется в соответствии со стандартным или расширенным языком гипертекстовой разметки между скриптом, расположенным на клиенте, и обработчиком ресурса, расположенным на прокси-сервере. Подобные сообщения используются для управления ресурсами, встроенными в скрипт, и для интерфейса с прокси-сервером, предоставленного для таких ресурсов с целью запрашивания и приема дополнительной информации.

В соответствии с дополнительным вариантом осуществления прокси-сервер извлекает ресурсный модуль в соответствии с запрошенной в мультимодальном диалоге модальностью. Предпочтительно мультимодальный сервер предоставляет набор ресурсных модулей, в частности модуль распознавания речи, модуль перевода текста в речь и модуль распознавания рукописного ввода.

Прокси-сервер инициализирует сеансы между клиентом и извлеченными ресурсными модулями. Дополнительно он компонует мультимодальное взаимодействие согласно инициализированным сеансам. Предпочтительно извлеченный ресурсный модуль осуществляет связь с ресурсами ввода/вывода клиента через одно или более RTP-соединений. Прокси-сервер создает обработчик ресурса для

выбранного ресурсного модуля, который осуществляет связь через TCP/IP соединение с ресурсным модулем (TCP = протокол управления передачей; IP = межсетевой протокол).

В соответствии с предпочтительным вариантом осуществления прокси-сервер создает обработчик ресурса для выбранного ресурсного модуля. Этот обработчик ресурса обеспечивает интерфейс с прокси-сервером для скрипта, расположенного на клиенте, и дает возможность скрипту обмениваться информацией с извлеченным ресурсным модулем. Для скрипта, расположенного на клиенте, становится возможным управлять ресурсным модулем или получать информацию, введенную пользователем и обработанную ресурсным модулем. Обработчики ресурса предоставляют своего рода прикладной мультимодальный интерфейс для скриптов, расположенных на клиенте, что позволяет встраивать обработку мультимодальных событий в документы, кодированные в HTML/XHTML.

Перечень чертежей

Эти и другие признаки и преимущества изобретения могут быть лучше оценены по достоинству после прочтения приведенного ниже подробного описания предпочтительных в настоящее время иллюстративных вариантов осуществления, рассматриваемых совместно с прилагающимися чертежами, на которых:

Фиг.1 - блок-схема, показывающая мультимодальную систему с прокси-сервером в соответствии с изобретением.

Фиг.2 - функциональное представление мультимодальной системы с прокси-сервером в соответствии с изобретением.

Подробное описание изобретения

Фиг.1 показывает сети 11 и 12 связи, клиент 4, сервер 5 приложений, прокси-сервер 2 и мультимодальный сервер 3.

Клиент 4 является «тонким» клиентом, т.е. клиентом с пониженными вычислительными ресурсами и памятью. Предпочтительно клиент 4 является переносным устройством, например PDA (PDA = персональный цифровой ассистент) или интеллектуальным телефоном (смартфоном), например UMTS-телефоном с возможностями мультимодального ввода и вывода. Но также возможно, что клиент является телефоном с возможностью обработки и передачи данных, например сотовым GSM-телефоном с GPRS-возможностями (GPRS = служба пакетной радиосвязи общего назначения), переносным компьютером или обычным персональным компьютером.

Клиент 4 обеспечивает набор из двух или более различных модальностей для взаимодействия с пользователем 6.

Модальность описывает способ, которым информация предоставляется от клиента 4 пользователю 6 или от пользователя 6 клиенту 4. Например, информация может быть представлена в виде голосового сообщения, написанной на экране информации, пиктограммы или графического изображения на экране, нажатия определенной клавиши на клавиатуре, ввода рукописной команды, пером, мышью, голосовой команды, ввода командного слова или касания пиктограммы на сенсорной панели.

В качестве примера Фиг.1 показывает четыре различных типа взаимодействия с 81 по 84 между клиентом 4 и пользователем 6, каждое из которых соответствует отличающейся от других модальности.

Сеть 11 связи соединяет клиента 4 с прокси-сервером 2 и мультимодальным сервером 3. Дополнительно сеть 12 связи соединяет прокси-сервер 2 с сервером 5

приложений.

Сети 11 и 12 связи позволяют осуществлять обмен данными между упомянутыми выше компонентами, преимущественно на основе стека протоколов TCP/IP.

5 Например, прокси-сервер 2, мультимодальный сервер 3 и сервер 5 приложений соединены через IP-сеть, которая связана с сетью мобильной связи, играющей роль сети доступа для клиента 4 для доступа к этой IP-сети (IP = межсетевой протокол). В таком случае сеть 11 связи образована сетью мобильной связи, например сетью GSM или UMTS, и фиксированной сетью передачи данных, т.е. упомянутой выше IP-сетью. 10 IP-сеть может состоять из множества физических сетей связи, например сетей ATM, MPLS или Ethernet (ATM = асинхронный режим передачи; MPLS = многопротокольная коммутация на основе меток), которые соединены через общий IP-протокол третьего уровня. Дополнительно клиент 4 и прокси-сервер 2/мультимодальный сервер 3 могут обмениваться данными через асинхронное соединение (например, через GPRS службу 15 или через синхронное соединение через сеть мобильной связи).

Но также возможно, что клиент 4 является фиксированным терминалом или терминалом, соединенным с прокси-сервером 2 и мультимодальным сервером 3 через WLAN-интерфейс (WLAN = беспроводная локальная сеть). В таком случае и 20 сеть 11 связи, и сеть 12 связи составлены упомянутой выше IP-сетью.

Сервер 5 приложений выполняет одно или более мультимодальных приложений, например, приложений, имеющих мультимодальные возможности. Например, такие приложения могут быть закодированы в мультимодальном языке разметки, например, с помощью HTML+, SALT или X+V. Если к таким приложениям обратился клиент, то 25 они предоставляют мультимодальный пользовательский интерфейс пользователю соответствующего клиента.

Клиент 4 имеет один или несколько процессоров, исполняющих программное обеспечение, и различные устройства ввода/вывода, например устройства 42, 43, 44, 30 и 45 ввода/вывода. Исполнение такого программного обеспечения процессором клиента 4 обеспечивает функциональные возможности стандартного web-браузера 41, выполненного с возможностью обмена документами и представления документов, закодированных в HTML или XHTML. Устройство 42 ввода/вывода является дисплеем, клавиатурой, мышью и сенсорной панелью, устройство 43 ввода/вывода является 35 громкоговорителем, устройство 44 ввода/вывода является микрофоном и устройство 45 ввода/вывода является пером для предоставления возможности рукописного ввода.

Прокси-сервер 2 является IP-сервером, расположенным в IP-сети. Но также 40 возможна реализация функциональных возможностей прокси-сервера 2 в сервере сетевого доступа, управляющем доступом клиента 4 в упомянутую выше IP-сеть.

Прокси-сервер 2 состоит из одного или нескольких соединенных между собой компьютеров, программной платформы и множества прикладных программ, выполняющихся на этой платформе. Функциональные возможности прокси-сервера 2 45 обеспечиваются выполнением этого программного обеспечения на основе системной платформы. С функциональной точки зрения прокси-сервер 2 предоставляет один или более мультимодальных модулей доступа, каждый из которых имеет базу 21 данных, менеджер 22 диалогов, компоновщик 23 и один или более обработчиков 24 ресурса.

50 Мультимодальный сервер 3 является интернет-сервером, который предоставляет один или более ресурсных модулей, поддерживающих мультимодальное взаимодействие между пользователем 6 и клиентом 4. Например, мультимодальный сервер 3 содержит ресурсные модули 31, 32 и 33. Ресурсный модуль 31 является

модулем распознавания речи, ресурсный модуль 32 является модулем преобразования текста в речь и ресурсный модуль 33 является модулем распознавания рукописного ввода. Ресурсные модули с 31 по 33 мультимодального сервера 3 могут совместно использоваться множеством мультимодальных модулей доступа, расположенных на
5 разных прокси-серверах. Но также возможно, что ресурсные модули с 31 по 33 встроены в прокси-сервер 2 и используются исключительно одним или более мультимодальными модулями доступа, предоставляемыми прокси-сервером 2.

Менеджер 22 диалога осуществляет управление мультимодальным диалогом. При
10 приеме HTTP-сообщения от клиента 4, который запрашивает доступ к мультимодальному приложению, он устанавливает мультимодальный диалог между этим мультимодальным приложением и пользователем 6 клиента 4. Он передает HTTP-запросы в соответствующей форме соответствующему мультимодальному серверу приложений, например мультимодальному серверу 5 приложений. При приеме ответа
15 от сервера 5 приложений он взаимодействует с клиентом 4 посредством обмена информацией, закодированной в стандартном или расширенном языке гипертекстовой разметки, и извлекает по меньшей мере один дополнительный ресурс модальности, запрошенный в мультимодальном диалоге.

Например, менеджер 22 диалога создает набор из одного или более скриптов, обеспечивающих ресурс модальности, или встраивает обработку мультимодальных событий посредством базы 21 данных. Он компонует скрипты в HTML- или XHTML-документ, отправляемый клиенту 4 для выполнения браузером 41. Дополнительно менеджер диалога может создавать набор из одного или более обработчиков ресурса
25 для обеспечения интерфейса с прокси-сервером для этих скриптов.

Дополнительно менеджер 22 диалога извлекает - основываясь на ответе или ответах от сервера 5 приложений - один или более ресурсных модулей мультимодального сервера 3 и инициализирует сеанс между компонентами клиента 4 и этими
30 извлеченными ресурсными модулями.

Например, менеджер 22 диалога создает HTML-документ 7 с набором скриптов 71, 72 и 73 и передает этот документ в качестве HTTP-ответа 92 клиенту 4. Скрипт 71 реализует вывод определенного голосового сообщения и добавляет ресурс «определенное голосовое сообщение» в качестве дополнительного ресурса
35 модальности к набору возможностей браузера 41. Скрипт 73 встраивает обработку мультимодальных событий в HTML-документ 7. Например, скрипт связывает обработку мультимодальных событий с HTML-элементами документа 7, такими как поля ввода текста, списки, кнопки, ссылки и прочее. Предпочтительно скрипт 73
40 обменивается через коммуникационное соединение 93 HTML-сообщениями или XHTML-сообщениями с соответствующим обработчиком ресурса, расположенным на прокси-сервере 2, который обеспечивает прикладной интерфейс с прокси-сервером для скрипта 73.

Дополнительно менеджер 22 диалога выбирает ресурсные модули
45 мультимодального сервера 3 в соответствии с определенными модальностями, запрошенными в мультимодальном диалоге, т.е. резервирует ресурсные модули и привязывает их к менеджеру 22 диалога.

Компоновщик 23 компонуется мультимодальное взаимодействие с пользователем 6 на основе взаимодействий с клиентом 4, основывающихся на стандартном или
50 расширенном языке гипертекстовой разметки, и на основе извлеченных дополнительных ресурсов. Например, компоновщик 23 компонуется мультимодальное взаимодействие в контексте мультимодального диалога из взаимодействий с 81 по 84,

где взаимодействие 81 является взаимодействием, основывающимся на стандартном или расширенном языке гипертекстовой разметки, взаимодействие 82 является голосовым сообщением, созданным скриптом 71, взаимодействие 83 является голосовым вводом, обработанным ресурсным модулем 31 под управлением компоновщика 23 и скрипта 73, и взаимодействие 84 является рукописным вводом, обработанным ресурсным модулем 33 под управлением компоновщика 23 и скрипта 73.

Компоновщик 23 компоует мультимодальный диалог с пользователем 6 в соответствии с информацией, принятой от сервера 5 приложений, и событиями, принятыми от пользователя 6, при этом эти события могут быть предварительно обработаны скриптом 73.

Например, мультимодальное приложение запрашивает вывод информации в виде голосового сообщения или пользователь 6 выбирает, что вывод информации должен быть осуществлен в виде голосового сообщения. Система реализует это голосовое сообщение путем создания скрипта 71 и выполнения скрипта 71 браузером 41 или путем извлечения ресурсного модуля 23, установления RTP-соединения и инициализации сеанса между ресурсным модулем 32 и компонентом 43 клиента 4, и инициирования ресурсного модуля 32 через обработчик 24 ресурса для создания голосового сообщения и передачи его компоненту 43.

Например, скрипт 73 обнаруживает событие, означающее запрос от пользователя для выполнения голосового ввода. Скрипт 73 соединяется с прокси-сервером, который извлекает ресурсный модуль 31, иницирует установление RTP-соединения 94 с компонентом 44, иницирует сеанс между этими компонентами и принимает, и обрабатывает результаты выполненной ресурсным модулем 31 обработки.

Фиг.2 показывает некоторые подробные реализации мультимодальной системы, показанной на Фиг.1. Фиг.2 показывает клиент 4, прокси-сервер 2, сервер 5 приложений и ресурсные модули с 31 по 33. Клиент 4 содержит web-браузер 41 и компоненты с 43 по 45.

Скрипты с 71 по 73 выполнены в виде Java-скриптов. Браузер 41 содержит ActiveX-объект 46, используемый Java-скриптом для вызова сервлета (обслуживающей программы на стороне сервера) прокси-сервера 2, предоставленного соответствующим Java-классом. Сервлет обрабатывает информацию и создает соответствующую ответную информацию, которая обрабатывается ActiveX-объектом 46 и может быть использована скриптом без перезагрузки HTML-документа 41.

В данном контексте web-сторона, сформированная документом 41, играет роль клиента, и прокси-сервер 2 с сервлетом играет роль сервера. Формат сообщений, которыми обмениваются клиент и сервер, удовлетворяет XML-стандарту (XML = расширяемый язык разметки).

JAVA-скрипт 73 имеет функцию (запрос), запрашивающую через XML сообщение «XML-запрос» метод сервлета «управление сервлетом». Функция JAVA-скрипта (запрос) передает свой входной параметр «запрос XML-строки» строчного типа методу сервлета «управление сервлетом». В то же время она устанавливает обработчик события для события «изменение состояния готовности (Ready State)» для ActiveX-объекта 46. Если функция 74 Java-скрипта успешно принимает ответ от сервлета, обработчик события вызывает функцию «обработка XML-ответа» и передает DOM-документ ответа (DOM = объектная модель документа).

Java-функция 75, называемая «обработка XML-ответа», ожидает DOM-документ

в качестве ввода, осуществляет поиск во вводе узла с именем "mathResponse" и вставляет это значение в текстовое поле документа 7.

Предпочтительно браузер 41 предоставляет XMLHttpRequest-объект, который может быть использован Java-скриптами с 71 по 73. Java-скрипт выполняется на браузере 41 клиента и отправляет XML-сообщения сервлету с использованием этого объекта. Например, XMLHttpRequest-объект поддерживается Microsoft Internet Explorer версии 5.0 или более поздней. Другие браузеры (например, Mozilla and Netscape версии 6) обеспечивают подобный интерфейс.

Информация (сообщения) может быть отправлена асинхронно по протоколу HTTP от сервера клиенту без перезагрузки страницы. Клиент делает неблокирующий запрос серверу, который принимает решение в отношении времени для ответа. Это имитирует технологию «server-push».

Обобщенное изложение сообщения XMLHttpRequest-объекта:

void	open(String method, String url, Boolean async) инициализирует запрос и задает метод, URL (URL = унифицированный указатель информационного ресурса) и то, будет ли вызов асинхронным или нет. Методами являются "POST" и "GET".
void	send(String message) отправляет запрос.
void	setRequestHeader(String parameter, String value) устанавливает параметры в HTTP-заголовке.
String	responseText() возвращает ответ в виде строки.
XMLDOM	responseXML() возвращает ответ в виде XMLDOM.
EventHandler	onreadystatechange() обработчик события, который будет вызван, если параметр readyState изменится.
int	readyState() возвращает состояние соединения.

На стороне клиента периодически вызывается функция Java-скрипта «request ()». Если нет активного запроса, то она открывает новый HTTP-запрос посредством вызова команды POST XMLHttpRequest-объекта и отправки строки сообщения сервлету.

Сервлет, запущенный на сервере, принимает переданную строку, которую он может прочитать из потока ввода, и взаимодействует с приложением. После обработки потока ввода на основании статуса приложения сервлет формулирует свой ответ.

Параметр «async» может быть установлен в значение TRUE "Истина" в команде POST XMLHttpRequest-объекта для асинхронного выполнения запроса, что означает, что может пройти несколько секунд или минут между передачей и приемом без блокирования HTML-браузера. Таким образом, сервлет может ожидать его ответа, пока событие «push» («проталкивание») не будет вызвано приложением. Эта новая информация передается клиенту.

Когда клиент принимает ответ, событие «onreadystatechange» XMLHttpRequest-объекта вызывается с параметром «readyState», имеющим значение 4. Обработчик события, установленный на это событие, может затем обработать ответ.

Благодаря этому механизму нет необходимости в перезагрузке HTTP-документа, и обмен информацией между скриптами с 71 по 73 и соответствующими обработчиками ресурса является для пользователя невидимым.

Обработчик 24 ресурса предоставляет MMAPI-HTTP-интерфейс (MMAPI = мультимодальный интерфейс прикладного программирования). Он содержит набор из одного или более сервлетов, например, сервлеты 25 и 26, непосредственно взаимодействующие со скриптами с 71 по 73, и поддерживает взаимодействие со

скриптами с 71 по 73 через XML-запросы и XML-ответы. Дополнительно он содержит набор MMAPI-компонентов, служащих в качестве интерфейса для связи с модулями с 31 по 33 ресурсов. Например, Фиг.2 показывает MMAPI-компонент 27, имеющий сокет 274, буфер 271 запроса, буфер 272 ответа и поток 273 наблюдения.

Сокет 274 поддерживает связь между MMAPI-компонентом 27 и соответствующим ресурсным модулем, например, ресурсным модулем 33, через TCP/IP соединение связи. Он анализирует буфер запроса 271 и передает соответствующие запросы через это соединение связи. Поток сервера сканирует информацию, принятую через TCP/IP соединение, и заполняет буфер ответа, если он обнаружил соответствующие ответы.

Сервлеты 25 и 26 передают запросы в буфер запроса и принимают ответы из буфера ответа с помощью соответствующих функций заполнения запроса и получения ответа. Таким образом, система выполняет синхронный обмен данными между сервлетами и MMAPI-компонентами.

Ресурсы с 31 по 33 предоставляют функциональные возможности для связи с клиентом 4 через TCP/IP-соединение и через RTP-соединение. Например, ресурсный модуль 33 содержит модуль 34 связи, поддерживающий TCP/IP-интерфейс, и модуль 35 связи, поддерживающий RTP-интерфейс. Каждый из ресурсных модулей с 31 по 33 может подобным образом поддерживать множество клиентов.

Например, MMAPI-компонент 27 передает запрос на инициализацию сеанса ресурсному модулю 33, который устанавливает соответствующий сеанс с предназначенным для этого компонентом клиента 4 и возвращает результаты этого сеанса.

Формула изобретения

1. Способ обеспечения мультимодального диалога между мультимодальным приложением (5) и пользователем (6), осуществляющим связь с мультимодальным приложением (5) через клиент (4), выполненный с возможностью обмена документами и представления документов, закодированных в стандартном или расширенном языке гипертекстовой разметки, при этом способ содержит этапы, на которых

устанавливают мультимодальный диалог между мультимодальным приложением (5) и пользователем (6) через прокси-сервер (2), взаимодействующий с клиентом (4) посредством обмена информацией, закодированной в стандартном или расширенном языке гипертекстовой разметки;

извлекают с помощью прокси-сервера (2), по меньшей мере один дополнительный ресурс (71, 72, 73, 31, 32, 33) модальности, запрашиваемый в мультимодальном диалоге; и komponуют с помощью прокси-сервера (2) мультимодальное взаимодействие с пользователем (6) на основе взаимодействий (81) с клиентом, основывающихся на стандартном или расширенном языке гипертекстовой разметки, и на основе извлеченных дополнительных ресурсов (71, 72, 73, 31, 32, 33).

2. Способ по п.1, отличающийся тем, что содержит дополнительные этапы, на которых komponуют набор из одного или более запрошенных ресурсов (71, 72, 73) в документ (7), представленный в стандартном или расширенном языке гипертекстовой разметки; передают документ (7) клиенту (4).

3. Способ по п.2, отличающийся тем, что прокси-сервер (2) создает скрипт (71), предоставляющий ресурс модальности, и komponует скрипт (71) в документ (7).

4. Способ по п.2, отличающийся тем, что прокси-сервер (2) создает скрипт (73) для встраивания обработки мультимодального события в документ (7) и komponует скрипт (73) в документ (7).

5. Способ по п.4, отличающийся тем, что содержит дополнительные этапы, на которых создают обработчик (24) ресурса, который предоставляет интерфейс с прокси-сервером для скрипта (73), и обмениваются сообщениями (93) в соответствии со стандартным или расширенным языком гипертекстовой разметки между
5 скриптом (73), расположенным на клиенте (4), и обработчиком (24) ресурса, расположенным на прокси-сервере (2).

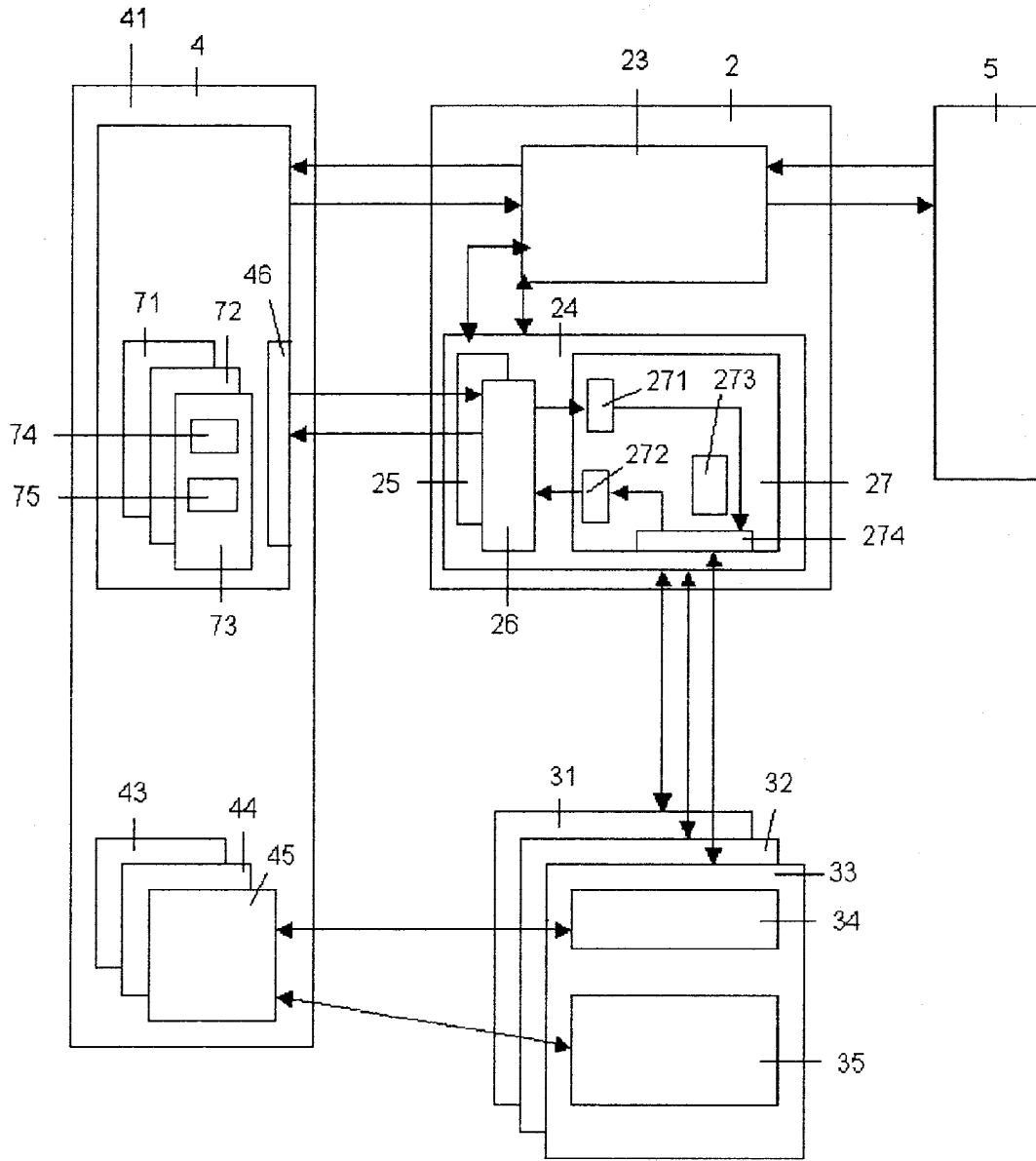
6. Способ по п.1, отличающийся тем, что прокси-сервер (2) извлекает ресурсный модуль (31, 32, 33), в частности модуль распознавания речи, преобразования текста в
10 речь или распознавания рукописного ввода, в соответствии с модальностью, запрошенной в мультимодальном диалоге, инициализирует сеанс (94) между клиентом (4) и извлеченным модулем (31, 32, 33) ресурса и компоует мультимодальное взаимодействие на основе инициализированного сеанса (94).

7. Способ по п.6, отличающийся тем, что содержит дополнительные этапы, на
15 которых создают обработчик (24) ресурса для выбранного ресурсного модуля (31, 32, 33), который предоставляет интерфейс с прокси-сервером для скрипта (73), расположенного на клиенте (4), при этом обработчик (24) ресурса дает возможность скрипту (73) обмениваться информацией с извлеченным ресурсным модулем (31, 32, 33), в частности с модулем распознавания речи, преобразования текста в речь или
20 распознавания рукописного ввода.

8. Способ по п.6, отличающийся тем, что содержит дополнительный этап, на котором создают обработчик (24) ресурса для извлеченного ресурсного модуля (31, 32, 33), причем этот обработчик ресурса осуществляет связь через ТСРЛР-соединение с
25 ресурсным модулем (31, 32, 33).

9. Способ по п.6, отличающийся тем, что содержит дополнительный этап, на котором устанавливают RTP-соединение между компонентами (43, 44, 45) клиента (4) и извлеченным ресурсным модулем (31, 32, 33).

10. Прокси-сервер (2) для поддержки мультимодальных диалогов между
30 мультимодальными приложениями (5) и пользователями (6), осуществляющими связь с мультимодальным приложением с помощью соответствующих клиентов (4), выполненных с возможностью обмена документами и представления документов, закодированных в стандартном или расширенном языке гипертекстовой разметки,
35 при этом прокси-сервер содержит менеджер (22) диалога, выполненный с возможностью установления мультимодального диалога между мультимодальным приложением (5) и пользователем (6) через прокси-сервер (2), для взаимодействия с клиентом (4) посредством обмена информацией, закодированной в стандартном или
40 расширенном языке гипертекстовой разметки, и для извлечения по меньшей мере одного дополнительного ресурса (71, 72, 73, 31, 32, 33) модальности, запрошенного в мультимодальном диалоге, и компоновщик (23) для компонования мультимодального взаимодействия с пользователем на основе взаимодействий с клиентом (4), основывающихся на стандартном или расширенном языке гипертекстовой разметки, и
45 на основе извлеченных дополнительных ресурсов (71, 72, 73, 31, 32, 33).



ФИГ. 2