



(51) МПК
H04N 7/15 (2006.01)
H04L 12/18 (2006.01)

ФЕДЕРАЛЬНАЯ СЛУЖБА
 ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(21)(22) Заявка: 2015123696, 19.12.2013

(24) Дата начала отсчета срока действия патента:
 19.12.2013

Дата регистрации:
 05.10.2017

Приоритет(ы):

(30) Конвенционный приоритет:
 19.12.2012 US 13/719,314

(43) Дата публикации заявки: 10.01.2017 Бюл. № 1

(45) Опубликовано: 05.10.2017 Бюл. № 28

(85) Дата начала рассмотрения заявки РСТ на
 национальной фазе: 18.06.2015

(86) Заявка РСТ:
 US 2013/076671 (19.12.2013)

(87) Публикация заявки РСТ:
 WO 2014/100466 (26.06.2014)

Адрес для переписки:
 129090, Москва, ул. Б. Спасская, 25, стр. 3, ООО
 "Юридическая фирма Городисский и Партнеры"

(72) Автор(ы):

ВЕРТЭЙН Уилльям Джордж (US),
 ЛЕОРИН Саймон (US)

(73) Патентообладатель(и):

МАЙКРОСОФТ ТЕКНОЛОДЖИ
 ЛАЙСЕНСИНГ, ЭлЭлСи (US)

(56) Список документов, цитированных в отчете
 о поиске: US 2009002480 A1, 2009-01-01. US
 2012127259 A1, 2012-05-24. US 7161939 B2 IP,
 2007-01-09. RU 2010133959 A, 2012-02-20. RU
 2009146029 A, 2011-06-20. US 7081915 B1,
 2006-07-25.

C2

9

6

4

2

6

U

R

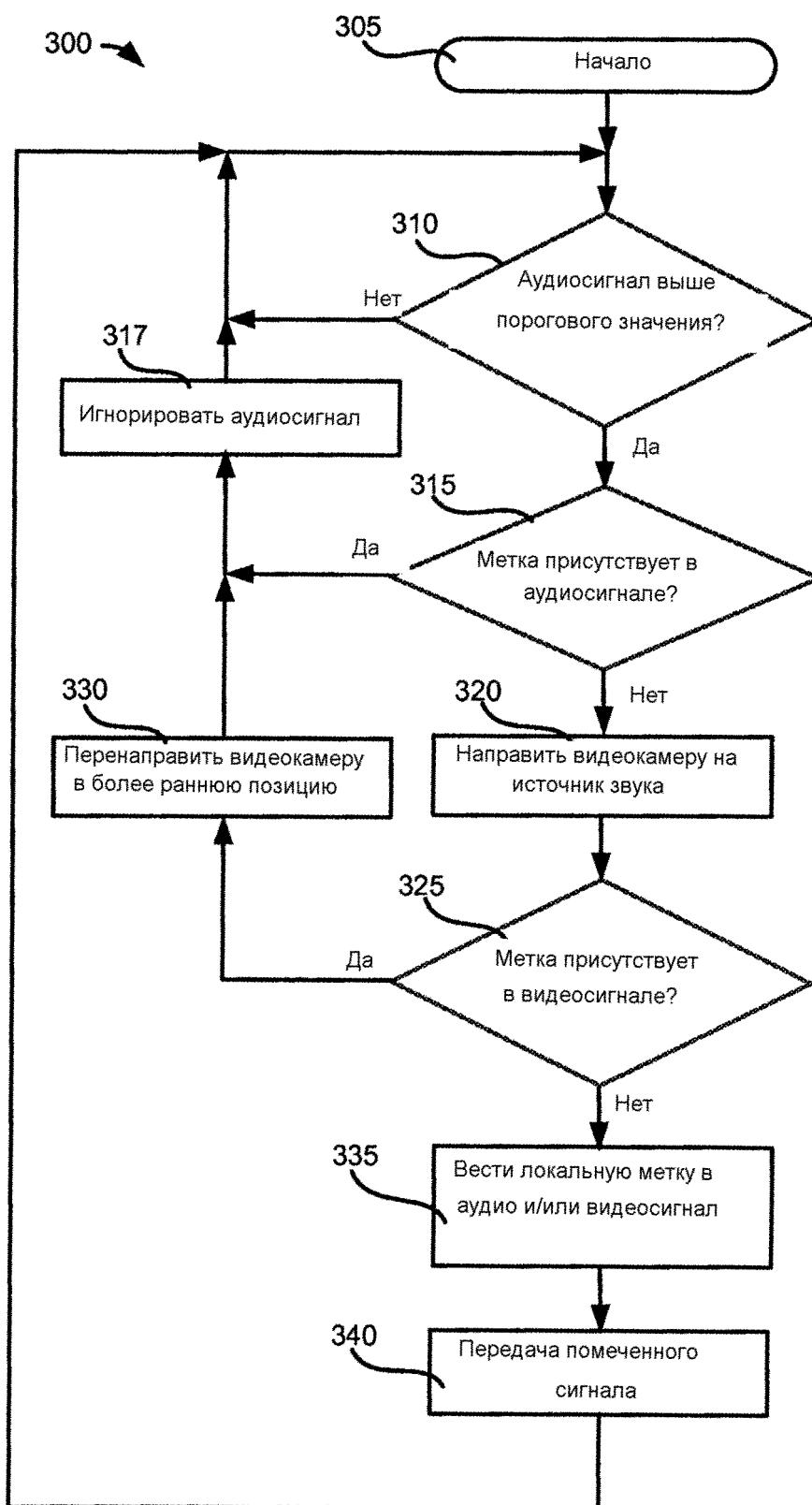
(54) ПРИСВОЕНИЕ ВИДЕО-И АУДИОМЕТОК ДЛЯ ОБНАРУЖЕНИЯ АКТИВНОГО ГОВОРЯЩЕГО УЧАСТНИКА

(57) Реферат:

Изобретение относится к видеоконференцсвязи. Техническим результатом является выбор активного говорящего участника при исключении ошибочного выбора микрофона или видеокамеры, которые захватывают аудио- или видеосигнал из соединенного сигнала из удаленного местоположения. Предложена система видеоконференцсвязи, в которой определяют, находится ли аудиосигнал выше порогового уровня. Если это так, то осуществляют определение того, присутствует ли в этом аудиосигнале метка. Если это так, то сигнал игнорируется. Если нет, то видеокамеру

направляют к источнику звука, идентифицированному посредством аудиосигнала. Определяют, присутствует ли метка в видеосигнале, поступающем от этой видеокамеры. Если это так, то видеокамеру перенаправляют. Если нет, то в аудиосигнал и/или видеосигнал вводят локальную метку (метки). Затем осуществляют передачу помеченного сигнала (сигналов). Таким образом, система будет игнорировать звук или видео, которые имеют встроенную метку от другой системы видеоконференцсвязи. 3 н. и 10 з.п. ф-лы, 5 ил.

R
 U
 2
 6
 3
 2
 4
 6
 9
 C
 2



ФИГ.3

R U 2 6 3 2 4 6 9 C 2

RUSSIAN FEDERATION



(19) RU (11) 2 632 469⁽¹³⁾ C2

(51) Int. Cl.
H04N 7/15 (2006.01)
H04L 12/18 (2006.01)

FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) ABSTRACT OF INVENTION

(21)(22) Application: 2015123696, 19.12.2013

(24) Effective date for property rights:
19.12.2013

Registration date:
05.10.2017

Priority:

(30) Convention priority:
19.12.2012 US 13/719,314

(43) Application published: 10.01.2017 Bull. № 1

(45) Date of publication: 05.10.2017 Bull. № 28

(85) Commencement of national phase: 18.06.2015

(86) PCT application:
US 2013/076671 (19.12.2013)

(87) PCT publication:
WO 2014/100466 (26.06.2014)

Mail address:
129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"

(72) Inventor(s):
VERTEJN Uillyam Dzhordzh (US),
LEORIN Sajmon (US)

(73) Proprietor(s):
MAJKROSOFT TEKNOLODZHI
LAJSENSING, EIEISi (US)

R U 2 6 3 2 4 6 9 C 2

(54) ASSIGNMENT OF VIDEO AND AUDIO TAGS FOR DETECTING ACTIVE SPEAKER

(57) Abstract:

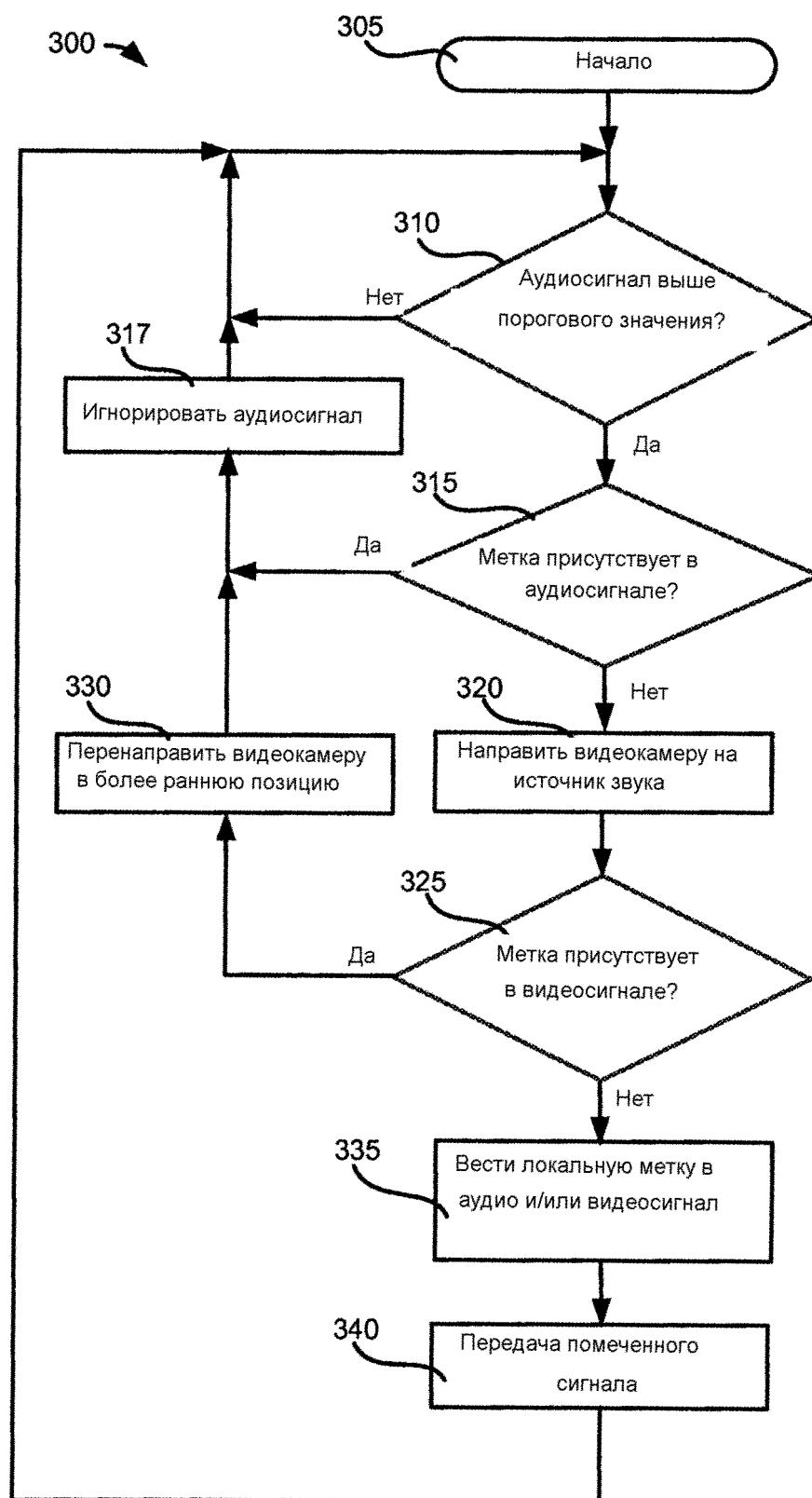
FIELD: physics.

SUBSTANCE: videoconferencing system is proposed in which it is determined whether the audio signal is above a threshold level. If so, a determination is made as to whether there is a mark in this audio signal. If so, the signal is ignored. If not, the video camera is routed to the sound source identified by the audio signal. Determining whether a mark is present in the video signal coming from this video camera. If so, the camera is redirected. If not, a local mark (tags) is

input into the audio signal and/or video signal. Then, the marked signal (s) is transmitted. Thus, the system will ignore sound or video that have built-in tag from another videoconferencing system.

EFFECT: active speaker selection party when you exclude the erroneous choice microphone or video camera, which capture audio or video from the signal from the remote location.

13 cl, 5 dwg



ФИГ.3

УРОВЕНЬ ТЕХНИКИ

Видеоконференцсвязь стала широко распространенной, и во многих учреждениях имеются помещения, конфигурированные главным образом для сеансов видеоконференцсвязи. Такие комнаты обычно содержат аппаратуру для

- 5 видеоконференцсвязи, такую как одна или более подвижных видеокамер и один или более микрофонов, причем микрофоны обычно размещают в местоположениях вокруг столика в комнате для участников. Обнаружение активного говорящего участника (ASD) часто используют для выбора видеокамеры или для перемещения (наведения и/или наклона) видеокамеры для показа человека в помещении, который говорит, и/или
- 10 для выбора микрофона, который будет активным. Когда говорит человек, находящийся в удаленном местоположении, его изображение и/или звук исходит из аудио-видеодисплея, такого как телевизор (ТВ), монитор или другой тип дисплея в комнате. Это может побудить ASD случайно выбрать изображение говорящего человека, находящегося в удаленном местоположении, на ТВ вместо выбора последнего
- 15 говорящего или говорившего человека в данном местоположении.

Также при сеансах видеоконференцсвязи с несколькими местоположениями, при которых во время одного сеанса видеоконференцсвязи имеют место три или более отдельных местоположения, то, как правило, будут отображаться несколько панелей, причем одна панель будет больше, чем другие, и будет показан человек, который

- 20 говорит, а другие панели будут показывать изображение из видеокамеры при других местоположениях. При возникновении случайного ASD, как было упомянуто выше, оборудование в помещении, где человек говорит, будет отправлять сигнал на оборудование в других местоположениях, при условии, что человек в своем местоположении говорит, и, таким образом, основной дисплей должен находиться в
- 25 его видеокамере. Когда это происходит, более крупная панель может переключаться от показа человека, который в данный момент говорит, на показ изображения ТВ-экрана или пустого кресла. Таким образом, проблема с ASD состоит в том, что если звук от удаленной системы видеоконференцсвязи отражается, или он настолько громкий, что он приводит в действие ASD, то звук из удаленного местоположения может быть
- 30 перенаправлен назад в удаленную систему и/или побудить видеокамеру в данном местоположении сфокусироваться на пустом кресле или на экране дисплея, показывающем удаленное местоположение видеоконференцсвязи.

Одна технология, которая была использована для исключения такого ошибочного выбора ASD, состоит в отслеживании линии развертки изображения пятна на экране телевизора для определения того, что звук исходит от телевизора, а не от человека в данном местоположении. Однако телевидение с высокой разрешающей способностью (HDTV), обладает высокими (240 Гц или выше) скоростями поступательной развертки и разрешениями изображения, которые одинаковы для видеокамер, и, таким образом, отслеживание линии развертки изображения имеют ограниченное использование, при

- 35 задействовании HDTV. Кроме того, ASD часто может иметь помехи при отражении звука в окружающем пространстве помещения. Звукоотражающая поверхность, такая как окно или картинка, покрытая стеклом, может отражать звук от телевизора таким образом, чтобы возникал звук, исходящий от человека за столом, расположенного в данном местоположении, даже если в данный момент отсутствует человек, сидящий в
- 40 этом местоположении за столом. Кроме того, если запись осуществляют в ходе видеоконференцсвязи, это зависит от того, что человек, чтобы не забыть, точно помечает запись по меньшей мере например, датой видеоконференцсвязи. Это часто забывают и делают позже, иногда с ошибочной или неполной меткой. Это, применительно к этим
- 45

и другим соображениям, которые описаны в раскрытии, представлено в настоящем документе.

РАСКРЫТИЕ ИЗОБРЕТЕНИЯ

В настоящем документе описаны технологии для системы видеоконференцсвязи,

- 5 которые выбирают активного говорящего участника, при исключении ошибочного выбора микрофона или видеокамеры, которые захватывают аудио- или видеосигнал из соединенного сигнала из удаленного местоположения. В одном варианте выполнения метку добавляют к исходящему аудио- и/или видеосигналу. Если микрофон захватывает звук, который содержит метку от удаленной системы, то звук игнорируется, и ASD не
- 10 реализуется. Если звук не содержит удаленную метку, то просматривают видео, полученное от видеокамеры в данном местоположении. Если он содержит удаленную метку, то ASD не выполняется. Если удаленная метка не присутствует ни в каком сигнале, то ASD выполняется.

Согласно одному варианту выполнения, представленному в настоящем документе,

- 15 передающая система для системы видеоконференцсвязи имеет генератор меток для формирования по меньшей мере одного из аудиометки или видеометки, блок объединения сигналов для по меньшей мере одного из (i) объединения принятого аудиосигнала с аудиометкой для получения помеченного аудиосигнала, или (ii) объединения принятого видеосигнала с видеометкой для получения помеченного
- 20 видеосигнала, и передатчик для передачи (i) помеченного аудиосигнала и принятого видеосигнала, (ii) принятого аудиосигнала и помеченного видеосигнала, или (iii) помеченного аудиосигнала и помеченного видеосигнала. Удаленная система видеоконференцсвязи может затем использовать встроенные метки, позволяющие отличать звуки и изображения из данного местоположения от звуков и изображений из удаленного местоположения.

Способ работы передатчика согласно системе видеоконференцсвязи включает в себя прием аудиосигнала, прием видеосигнала, формирование по меньшей мере одного из аудиометки или видеометки, по меньшей мере одно из (i) объединения аудиосигнала с аудиометкой для получения помеченного аудиосигнала, или (ii) объединения

- 30 видеосигнала с видеометкой для получения помеченного видеосигнала, и передачу (i) помеченного аудиосигнала и видеосигнала, (ii) аудиосигнала и помеченного видеосигнала, или (iii) помеченного аудиосигнала и помеченного видеосигнала.

Компьютерный запоминающий носитель имеет хранящиеся на нем машиноисполнимые команды. Эти команды побуждают компьютер формировать по

- 35 меньшей мере одно из аудиометки или видеометки, выполнять о меньшей мере одно из (i) объединения принятого аудиосигнала с аудиометкой для получения помеченного аудиосигнала или (ii) объединения принятого видеосигнала с видеометкой для получения помеченного видеосигнала, и передавать (i) помеченный аудиосигнал и принятый видеосигнал, (ii) принятый аудиосигнал и помеченный видеосигнал, или (iii) помеченный
- 40 аудиосигнал и помеченный видеосигнал.

Следует учитывать, что вышеописанный предмет изобретения также может быть реализован в виде устройства, управляемого компьютером, компьютерного процесса, вычислительной системы, или в виде готового изделия, такого как машиночитаемый носитель информации. Эти и различные другие признаки станут ясными из прочтения

- 45 следующего раздела «Осуществление изобретения» и обзора соответствующих чертежей.

Данное раскрытие изобретения приведено для ознакомления с набором понятий, приведенных в упрощенной форме, которые дополнительно описаны ниже в разделе «Осуществление изобретения». Данное раскрытие изобретения не предназначено ни

для идентификации ключевых признаков или основных признаков заявленного предмета изобретения, ни для использования данного раскрытия изобретения для ограничения объема заявленного предмета изобретения. Кроме того, заявленный предмет изобретения не ограничен вариантами выполнения, которые устраниют любой или все недостатки, отмеченные в любой части данного раскрытия.

КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ

Фиг. 1 представляет собой примерную конфигурацию передающей системы для системы видеоконференцсвязи.

Фиг. 2 представляет собой иллюстрацию примерной среды системы

10 видеоконференцсвязи.

Фиг. 3 представляет собой блок-схему, показывающую примерное обнаружение метки и технологию управления видеокамерой и микрофоном.

Фиг. 4 представляет собой блок-схему примерной технологии записи информации.

Фиг. 5 представляет собой схему архитектуры компьютера, показывающую

15 иллюстративное аппаратное обеспечение компьютера и архитектуру программного обеспечения для вычислительной системы, пригодной для реализации аспектов вариантов выполнения, представленных в настоящем документе.

ОСУЩЕСТВЛЕНИЕ ИЗОБРЕТЕНИЯ

Нижеследующее подробное описание направлено на создание технологий

20 видеоконференцсвязи, которые могут корректно выбирать активного говорящего участника, при исключении ошибочного выбора микрофона или видеокамеры, которые захватывают аудио- или видеосигнал из соединенного сигнала из удаленного местоположения. В следующем подробном описании будут сделаны ссылки на прилагаемые чертежи, которые составляют его часть, и которые показаны в качестве

25 иллюстрации конкретных вариантов выполнения или примеров. Обратимся теперь к чертежам, на которых одинаковые номера ссылочных позиций обозначают одинаковые элементы среди нескольких чертежей, и будут представлены аспекты вычислительной системы и способы видеоконференцсвязи.

Фиг. 1 представляет собой примерную конфигурацию передающей системы 105 для

30 системы 100 видеоконференцсвязи. Передающая система 105 имеет систему 120 управления и выбора видеокамеры и микрофона, генератор 125 видеометок, блок 130 формирования видеосигналов который обеспечивает выходной видеосигнал 135, генератор 140 аудиометок и блок 145 формирования аудиосигналов, который обеспечивает выходной аудиосигнал 150. Выходные видео и аудиосигналы могут быть

35 переданы путем радиопередачи или переданы передатчиком 155. Система 120 управления также может отправлять сигналы, предназначенные для удаленных систем, при условии, что она имеет активного говорящего участника, которая должна представлять собой более крупную панель, если для отображения нескольких местоположений используется несколько панелей. В передатчике 155 может быть использовано любое удобное средство

40 для отправки выходных видео- и аудиосигналов и любые управляющие сигналы на одну или более системы приемников 160 в удаленных местоположениях. Следует учитывать, что передающая система 105 и приемная система 160 существует в каждом местоположении, и что передающая система 105 и приемная система 160 в местоположении могут быть объединены в одно устройство.

45 Одна или более видеокамер 110 (110A-110N) и один или более микрофонов 115 (115A-115N) передают, соответственно, видеосигналы и аудиосигналы в передающую систему 105, а точнее, в систему 120 управления, которая имеет входы для приема этих сигналов. Система 120 управления и выбора видеокамеры и микрофона может позволить выбрать,

какая видеокамера 110 и какой микрофон 115 будут использованы для формирования изображения и звука из данного местоположения, и если используется более одного из любых устройств, то можно управлять панорамированием, приближением и/или наклоном выбранной видеокамеры 110, если видеокамеру можно регулировать 5 указанным образом, и может позволить регулировать формирование сигнала или других сигналов для передачи на удаленные системы.

Генератор 125 видеометок и генератор 140 аудиометок формирует, соответственно, видео- и аудиометки. Блок 130 формирования видеосигналов управляет или 10 модифицирует видеопиксели в потоке видеоданных, для добавления видеометки и создания помеченного видеосигнала 135. Блок 145 формирования аудиосигналов управляет или модифицирует биты в потоке аудиоданных для получения помеченного аудиосигнала 150. Это можно рассматривать как «присвоение меток» сигналу или добавление метки к сигналу. Генераторы 125 и 140 меток могут быть реализованы в 15 одном устройстве, блоки 130, 145 объединения сигналов могут быть реализованы в одном устройстве, и от одного до всех этих компонентов могут быть реализованы в виде части системы 120 управления.

Предпочтительно, чтобы поток видео- и/или аудиоданных был модифицирован с использованием таких способов, или лишь до таких уровней, которые являются 20 трудноуловимыми и/или не обнаружимыми для людей, но которые могут быть обнаружены путем алгоритмического анализа потока видео- или аудиоданных. Уровень искажения меньший заданного уровня может быть неразличимым для обычного человека-наблюдателя. Например, модификация наименьшего значащего бита в слове данных, даже если метка присутствовала в каждом слове, как правило, не может быть заметным или нежелательным. В качестве другого примера, размещение 25 видеометки в ходе гашения интервала или периода обратного хода, в видеокадре, или помещение видеометки в угол нижней части дисплея не может быть заметным или нежелательным. Даже размещение видеометки в качестве наибольшего значащего бита не может быть заметным или нежелательным, но только в случае если это делать на одном пикселе в ходе создания кадра.

Поток видео и/или аудиоданных может быть модифицирован, например, путем 30 использования наименьшего значащего бита или битов для доставки информации, отличной от исходного аудио- или видеосигнала. Такая модификация может быть сделана с каждым словом данных, с каждым другим словом данных, с каждым N-ным словом данных, через каждые N миллисекунд, перед или после синхронизации слова 35 данных или бита, и т.д. Например, последний бит (биты) подходящего слова данных (слов данных) может всегда иметь одинаковый объем, например, 0, 1, 2, 3, и т.д., может принимать различные значения, может возрастать по значениям, и т.д. Другие технологии также могут быть использованы для идентификации слова данных, или его части, такой как метка, или как идентификационная информация, связанная с меткой 40 или видеоконференцсвязью. В качестве другого примера, для этой цели может быть использовано все слово данных. Например, если аудиосигналы отбирают при скорости 4000 проб/секунда, то использование ограниченного количества этих слов для передачи метки информация не может заметно ухудшить качество аудиосигнала. Видеосигналы обеспечивают возможность отправления даже большего количества информации, без 45 значительного ухудшения качества видеосигнала.

Фиг. 2 представляет собой иллюстрацию примерной среды системы 200 видеоконференцсвязи. Несколько человек 205 (205A-205C) собрано вокруг стола 210, на котором имеется множество микрофонов 115 (115A-115E). Присутствует дисплей

215, который может представлять собой телевизор, показывающий человека 220 в удаленном местоположении. Также показан динамик 225. Имеется передающая система 105, которая соединена с видеокамерами и микрофонами, и приемная система 160, которая соединена с дисплеем и акустической системой. Как было упомянуто,

5 передающая система 105 и приемная система 160 могут быть, и обычно реализуются в виде одного устройства и соединены подходящим средством связи с одной или более удаленными системами видеоконференцсвязи.

Когда говорит человек, расположенный в данном местоположении, такой как человек 205B, система 120 управления обнаруживает сигнал, поступающий от микрофона 115B, 10 переключается на микрофон 115B, переключается на видеокамеру 110B, предварительно наведенную на область, в которой находится человек 115B, или наводит видеокамеру 110B на область, в которой находится человек 115B, а затем передает аудиосигнал от микрофона 115B и видеосигнал от видеокамеры 115B в удаленное местоположение, возможно наряду с сигналом, указывающим на то, что человек 205B должен быть 15 заметно отображен на удаленном экране. Наведением или направлением видеокамеры в контексте настоящего документа, является панорамирование, наклон и/или приближение видеокамеры для достижения желаемого изображения желаемого местоположения.

Рассмотрим теперь ситуацию, при которой присутствует звукоотражающий объект 20 или поверхность 230, такую как зеркало, картина или окно. Говорящий участник 220 в удаленном местоположении говорит, и голос говорящего участника 220 в удаленном местоположении передается по радио в помещение посредством динамика 225. Звук 235 от говорящего участника 220 в удаленном местоположении отражается от отражающей поверхности 230 и попадает в микрофон 115D. Система 120 управления 25 обнаруживает отраженный голос 235 на микрофоне 115D и ошибочно определяет, что существует расположенный в данном местоположении человек, говорящий у микрофона 115D. Система 120 управления затем переключается на микрофон 115D и наводит видеокамеру 110 на пустое пространство вблизи микрофона 115D. Таким образом, отраженные звуки и эхо может создать проблемы в ходе сеансов видеоконференцсвязи. 30 Это может происходить циклически, пока человек 220 в удаленном местоположении не прекратит говорить или кто-нибудь не уменьшит громкость динамика 225.

Для исключения или по меньшей мере ослабления такого ошибочного действия ASD, передающая система 105 водит метку (метки) в аудиосигнал и/или в видеосигнал.

Дисплей 215 и динамик 225 будут затем воспроизводить эту метку (метки) в своих 35 выходных сигналах. Теперь рассмотрим снова ситуацию, при которой говорит говорящий участник 220 в удаленном местоположении, а голос говорящего участника 220 из удаленного местоположения передается по радио в помещение посредством динамика 225. Звук 235 говорящего участника 220 из удаленного местоположения отражается от отражающей поверхности 230 и попадает на микрофон 115D. Система 120 управления обнаруживает отраженный голос 235 на микрофоне 115D, но также 40 обнаруживает и метку в отраженном голосе 235. Система 120 управления затем определяет, что звук исходит от говорящего участника в удаленном местоположении, а не от говорящего участника в данном местоположении, и поэтому не предпринимает никаких действий применительно к отраженному голосу.

45 В качестве другого подхода, когда в микрофоне 115D присутствует отраженный голос 235, система 120 управления может, вместо этого или кроме этого, проверять выходной сигнал видеокамеры. Если видеометка присутствует, то система 120 управления определяет, что звук представляет собой отраженный звук, и поэтому не

предпринимает никаких действий применительно к отраженному голосу.

Однако, когда говорит человек 205B, расположенный в данном местоположении, микрофон 115B обнаруживает голос человека 205B, расположенного в данном местоположении, но аудиометка отсутствует. Система 120 управления затем корректно

- 5 переключается на микрофон 115B и направляет видеокамеру 110 на человека 205B, расположенного в данном местоположении, и видеометка будет отсутствовать. Таким образом, система 120 управления корректно определяет, что человек 205B говорит, и предпринимает соответствующее действие. Следует учитывать, что на микрофонах 115B также может возникать некоторый отраженный звук 235. Однако, громкость
- 10 отраженного звука 235 будет значительно меньше, чем громкость голоса говорящего участника 205B, расположенного в данном местоположении, так что отраженная метка будет находиться на слишком низком уровне, чтобы обнаруживаться системой 120 управления. То есть при оцифровывании звука, исходящего из микрофона, громкость метки будет находиться ниже уровня наименьшего значащего бита (битов). Отраженный
- 15 звук 235 также может быть уловлен другими микрофонами 115, но система 120 управления будет отвергать эти микрофоны, либо потому что их громкость будет меньше, чем громкость у микрофона 115B, либо потому что метка будет легко обнаружимой.

В некоторых ситуациях возможно, чтобы видеокамера 240 находилась на задней 20 стенке комнаты, в дополнение или вместо видеокамер 110. Предположим теперь, что говорит человек 220 в удаленном местоположении, а звук, испускаемый динамиком 225, принимается микрофоном 115A или 115E. Стандартная система может ошибочно принять этот полученный звук за говорящего участника в данном местоположении и переключиться на этот микрофон и направить видеокамеру 240 на это местоположение.

- 25 Вместо этого, с помощью метки, используемой в настоящем документе, система 120 управления будет обнаруживать метку в аудиосигнале, захваченном микрофоном 115A или 115E, определять, что голос не является голосом говорящего участника в данном местоположении, и не переключаться на микрофон 115A или 115E. Также система 120 управления может наводить видеокамеру 240 на дисплей 215, обнаруживать видеометку, 30 используемую дисплеем 215, а затем наводит видеокамеру 240 назад по ее исходному направлению или по направлению, установленному по умолчанию. Таким образом, аудио и видеометки усиливают эффект видеоконференцсвязи путем снижения или исключения ошибочного переключения видеокамеры и/или микрофона, вызванного голосом говорящего участника в удаленном местоположении.

35 Метки также могут быть использованы при необходимости для идентификации видеоконференцсвязи. Например, метки могут содержать информацию относительно наименования компании, времени, даты, местоположения комнаты, передачи используемого оборудования, такую как (но не ограничиваясь) модель, изготовитель, серийный номер, версия программного обеспечения, информация о торговой марке, 40 информация об авторском праве, информация о конфиденциальности, информация о праве собственности, используемый протокол или стандарт, и т.д. Всю эту информацию не требуется передавать, и ничто не делает необходимой передачу всей этой желаемой информации мгновенно, циклически или непрерывно. Скорее, биты, которые идентифицируют метку как таковую, необходимо лишь передавать достаточно часто, 45 чтобы система 120 управления могла распознавать метку как таковую. Таким образом, например, как было упомянуто выше, биты, которые идентифицируют метку, которая может быть передана лишь в виде N слов данных, отличных от числа слов, используемых для передачи информации, упомянутой выше.

Кроме того, не требуется, чтобы информация, содержащаяся в метке (в метках), была получена из изображения, представленного дисплеем 215, или из звука, представленного динамиком 225. Скорее, и предпочтительно, эту информацию получают непосредственно из видео- и/или аудиосигналов, принятых приемной системой 160.

5 Скорость передачи данных может быть достаточно медленной, но является предпочтительным, чтобы идентифицируемая часть метки передавалась бы предпочтительно циклически, с менее чем половинным гистерезисом задержки ASD. Даже более предпочтительным является, чтобы идентифицируемая часть метки передавалась бы более часто, для размещения остатка данных, из-за помех в ходе 10 передачи данных или шума в помещении. Скорость подачи дополнительной информации является менее чувствительной к времени, и поэтому может быть передана в течение более длительного периода времени.

Фиг. 3 представляет собой блок-схему примерного обнаружения метки и технологии управления видеокамерой и микрофоном 300. После запуска 305 делают определение 310 того, будет ли аудиосигнал более высоким, чем пороговый уровень. Если нет, то следует возвратиться к 310. Если это так, то делают определение 315 того, присутствует ли в этом аудиосигнале метка. Если это так, то этот аудиосигнал игнорируется 317 и следует вернуться к 310. Если нет, то затем видеокамеру направляют или наводят 320 на источник звука, идентифицированный аудиосигналом. Например, если аудиосигнал 20 исходит из микрофона 115A, то видеокамера 110 будет направлена на область, обслуживаемую микрофоном 115A, или будет выбрана видеокамера, которая была ранее наведена на эту область.

Затем делают определение 325 того, присутствует ли метка в видеосигнале, исходящем из этой видеокамеры. Если это так, то видеокамеру перенаправляют 330 на ее более 25 раннюю позицию, или выбирают предыдущую видеокамеру. Если нет, то тогда локальную метку (метки) вводят 335 в аудиосигнал и/или в видеосигнал. Затем передают помеченный сигнал (сигналы). Затем следует вернуться к 310.

Таким образом, если микрофон улавливает звук, и существует аудиометка, встроенная в этот звук, или если видеокамера направлена к источнику этого звука, захватывающему 30 видеометку, встроенную в видеосигнал, то система будет игнорировать этот звук и оставлять микрофон и настройки видеокамеры, как они были. Однако, если встроенная метка не обнаружена ни в каком сигнале, то микрофон и/или видеокамера будут 35 выбраны для передачи этого звука и изображения в удаленную видеоконференцию после введения локальной метки по меньшей мере в один из этих сигналов. Таким образом корректно выбирают активного говорящего участника, игнорируя отраженные звуки из удаленных местоположений.

Фиг. 4 представляет собой блок-схему примерной технологии записи информации 400. После запуска 405, делают определение 410 того, записан ли сеанс. Если нет, то процедура заканчивается 415. Если это так, делают определение 420 того, имеются ли 40 метки. Если никаких меток нет, то сеанс записывается 430. Если присутствует по меньшей мере одна метка, то делают определение 425 того, имеется ли в метке (в метках) информация. Если нет, то сеанс записывают 430. Если это так, то сеанс записывают 435 с использованием по меньшей мере некоторой информации. Информация, которую записывают в ходе сеанса, может представлять собой всю информацию, включенную 45 в метку, или может представлять собой лишь предварительно выбранную часть, такую как дата и время.

Следует учитывать, что логические операции, описанные в настоящем документе, реализуются (1) в виде последовательности действий, выполняемых на компьютере,

или программных модулей, работающих на вычислительной системе, и/или (2) в виде соединенных между собой цепей машинной логики или электронных модулей в вычислительной системе. Реализация является предметом выбора, зависящим от рабочих характеристик и других требований к вычислительной системе. Следовательно,

5 логические операции, описанные в настоящем документе, относятся к различным установленным операциям, структурным устройствам, действиям или модулям. Эти операции, структурные устройства, действия и модули могут быть реализованы в программном обеспечении, во встроенных программах, в цифровой логической схеме специального назначения и в любом их сочетании. Также следует учитывать, что может
10 быть выполнено больше или меньше операций, чем показано на чертежах и описано в настоящем документе. Эти операции также могут быть выполнены в порядке, отличном от порядка, описанного в настоящем документе.

Фиг. 5 показывает иллюстративную архитектуру компьютера для компьютера 500, пригодного для исполнения компонентов программы, описанных в настоящем
15 документе для системы видеоконференцсвязи способом, представленным выше. Показанная архитектура компьютера иллюстрирует стандартный настольный персональный компьютер, портативный компьютер, или компьютер-сервер, и может быть применена для исполнения любых аспектов компонентов программ, представленных в настоящем документе, которые описаны как исполняемые на
20 клиентском компьютере 104, компьютерах-серверах 106A-106N для предварительной обработки данных или компьютерах-серверах 108A-108N базы данных. Показанная архитектура компьютера включает в себя центральный блок обработки данных 502 («ЦП»), память 508 системы, включающую в себя оперативное запоминающее устройство 514 («ОЗУ») и постоянное запоминающее устройство («ПЗУ») 516, а также
25 шину 504 системы, которая связывает память с ЦП 502. Базовая система ввода/вывода, содержащая базовые подпрограммы, которые помогают перемещать информацию между элементами в компьютере 500, например, в ходе запуска, хранится в ПЗУ 516. Компьютер 500 дополнительно включает в себя запоминающее устройство 510 большой
30 емкости, для хранения операционной системы 518, прикладных программ и других программных модулей, которые более подробно описаны в настоящем документе.

Запоминающее устройство 510 большой емкости соединено с ЦП 502 через контроллер устройства массовой памяти (не показан), соединенный с шиной 504. Запоминающее устройство 510 большой емкости и связанный с ним машиночитаемый носитель обеспечивают энергонезависимое запоминающее устройство для компьютера
35 500. Хотя описание машиночитаемого носителя, содержащееся в настоящем документе, относится к запоминающему устройству большой емкости, такому как жесткий диск или привод CD-ROM, специалистам в данной области техники следует учитывать, что машиночитаемый носитель может представлять собой любой подходящий
40 компьютерный носитель информации или канал информации, который может быть доступен для архитектуры компьютера 500.

В качестве примера, и без ограничений, компьютерный носитель информации может включать в себя энергозависимый и энергонезависимый, съемный и несъемный носитель, реализуемый посредством любого способа или технологии для хранения информации, такой как машиночитаемые команды, структуры данных, программные модули или
45 другие данные. Например, компьютерный носитель информации включает в себя (но не ограничен) RAM, ROM, EPROM, EEPROM, флэш-память или другую твердотельную технологию памяти, CD-ROM, универсальные цифровые диски (DVD), HD-DVD, BLU-RAY, или другое оптическое устройство хранения данных, магнитные кассеты,

магнитную ленту, запоминающее устройство на магнитном диске или другие магнитные запоминающие устройства, или любой другой носитель, который может быть использован для хранения желаемой информации и который может быть доступен для компьютера 500. В формуле изобретения словосочетание «компьютерный

5 запоминающий носитель» и ее видоизменения не включает себя волн или сигналов самих по себе и/или каналы информации.

Каналы информации включают в себя машиночитаемые команды, структуры данных, программные модули или другие данные в модулированном сигнале данных, таком как несущая волна или другой механизм переноса и включает в себя любые средства

10 доставки. Термин «модулированный сигнал данных» означает сигнал, который имеет один или более своих характеристик, изменяемых или заданных таким образом, чтобы можно было закодировать информацию в сигнале. В качестве примера (и без ограничений), каналы информации включают в себя проводные носители, такие как проводная сеть или однопроводное соединение, и беспроводные носители, такие как

15 акустические, радиочастотные, инфракрасные и другие беспроводные носители. Сочетания любых из приведенных выше, также должны быть включены в объем машиночитаемого носителя.

Согласно различным вариантам выполнения компьютер 500 может функционировать в сетевой среде, с использованием логических соединений для удаленных компьютеров,

20 через сеть, такую как сеть 520. Компьютер 500 может быть соединен с сетью 520 через блок 506 сетевого интерфейса, соединенный с шиной 504. Следует учитывать, что блок 506 сетевого интерфейса также может быть применен для соединения с другими типами сетей и с удаленными компьютерными системами. Компьютер 500 также может включать в себя контроллер вода/вывода 512 для приема и обработки входных данных,

25 поступающих от множества других устройств, включающих в себя клавиатуру, мышь или световое перо. Аналогично, контроллер вода/вывода может подавать выходной сигнал на экран дисплея, принтер или другой тип устройства вывода.

Как было кратко упомянуто выше, множество программных модулей и файлов данных может быть сохранено в запоминающем устройстве 510 большой емкости и 30 ОЗУ 514 компьютера 500, включая операционную систему 518, пригодную для управления функционированием сетевого настольного компьютера, портативного компьютера или компьютера-сервера. Запоминающее устройство 510 большой емкости и ОЗУ 514 также может хранить один или более программных модулей, которые реализуют различные операции, описанные выше. Запоминающее устройство 510 большой емкости и ОЗУ 514 также могут хранить другие типы программных модулей.

При том, что объект изобретения, описанный в настоящем документе, представлен в общем одним или более программными модулями, которые исполняются в сочетании с исполнением операционной системы и прикладных программ на компьютерной системе, специалистам в данной области техники должно быть понятно, что другие 40 варианты выполнения могут быть выполнены в сочетании с другими типами программных модулей. Как правило, программные модули включают в себя подпрограммы, программы, компоненты, структуры данных и другие типы структур, которые выполняют конкретные задачи или реализуют конкретные типы абстрактных данных. Более того, специалисты в данной области техники должны учитывать, что 45 предмет изобретения, описанный в настоящем документе, может быть реализован на практике, при необходимости с другими конфигурациями компьютерных систем, включая портативные устройства, многопроцессорные системы, микропроцессорную или программируемую бытовую электронную технику, миникомпьютеры, компьютеры

обычных размеров и т.п.

Исходя из вышесказанного, следует отметить, что в настоящем документе предложены технологии видеоконференцсвязи. Хотя предмет изобретения, представленный в настоящем документе, был описан в терминологии, характерной для структурных элементов компьютера, методологических и преобразовательных операций, конкретной вычислительной техники и машиночитаемых носителей информации, следует понимать, что изобретение, определенное в прилагаемой формуле изобретения, не обязательно ограничено конкретными элементами, операциями или носителями информации, описанными в настоящей работе. Скорее, конкретные элементы, операции и носители раскрыты в виде примерных форм реализации формулы изобретения.

Объект изобретения, описанный выше, приведен лишь в качестве иллюстрации, и не должен рассматриваться как ограничивающий. Для объекта изобретения, описанного в настоящем документе, могут быть сделаны различные модификации и изменения, помимо следующих проиллюстрированных и описанных примерных вариантов выполнения и применений, и без отступления от истинной сущности и объема представленного изобретения, которое изложено в следующей формуле изобретения.

(57) Формула изобретения

1. Передающая система для системы видеоконференцсвязи, содержащая:

- систему управления, выполненную с возможностью определять, имеется ли аудиометка в принятом аудиосигнале;

- генератор меток для формирования аудиометки;
- блок объединения для объединения упомянутого принятого аудиосигнала с аудиометкой, если определено, что принятый сигнал не имеет аудиометки, для получения помеченного аудиосигнала, так что аудиометка является обнаруживаемой в помеченном аудиосигнале для системы управления другой системы видеоконференцсвязи; и
- передатчик для передачи помеченного аудиосигнала и принятого видеосигнала как части видеоконференции.

2. Передающая система по п. 1, дополнительно содержащая систему управления, причем система управления встраивает дополнительную информацию в аудиометку.

3. Передающая система по любому из предшествующих пунктов, в которой генератор меток дополнительно выполнен с возможностью формировать видеометку, блок объединения дополнительно выполнен с возможностью объединять принятый видео сигнал с видеометкой для получения помеченного видеосигнала и при этом передатчик дополнительно выполнен с возможностью передавать помеченный аудиосигнал и помеченный видеосигнал.

4. Способ работы передающей системы для системы видеоконференцсвязи, причем способ содержит этапы, на которых:

- определяют, имеется ли аудиометка в принятом аудиосигнале;
- принимают видеосигнал;
- формируют аудиометку;
- объединяют аудиосигнал с аудиометкой, если определено, что принятый сигнал не имеет аудиометки, для получения помеченного аудиосигнала, так что аудиометка является обнаруживаемой в помеченном аудиосигнале для системы управления системы видеоконференцсвязи; и
- передают помеченный аудиосигнал и видеосигнал как часть видеоконференции.

5. Способ по п. 4, в котором имеется множество аудиосигналов и который дополнительно содержит этап, на котором выбирают аудиосигнал для передачи,

который не содержит аудиометку от другой системы видеоконференцсвязи.

6. Способ по п. 4 или 5, дополнительно содержащий:

объединение видеосигнала с видеометкой для получения помеченного видеосигнала;

и

5 передачу помеченного аудиосигнала и видеосигнала.

7. Способ по п. 6, в котором имеется множество видеосигналов и который дополнительно содержит этап, на котором выбирают видеосигнал для передачи, который не содержит видеометку от другой системы видеоконференцсвязи.

8. Компьютерный запоминающий носитель, на котором сохранены

10 машиноисполняемые команды, которые при исполнении компьютером побуждают компьютер:

- определять, имеется ли аудиометка в принятом аудиосигнале;

- формировать аудиометку;

- объединять принятый аудиосигнал с аудиометкой, если определено, что принятый

15 сигнал не имеет аудиометки, для получения помеченного аудиосигнала, так что аудиометка является обнаруживаемой в помеченном аудиосигнале для системы управления системы видеоконференцсвязи; и

- передавать помеченный аудиосигнал и принятый видеосигнал как часть видеоконференции.

20 9. Компьютерный запоминающий носитель по п. 8, причем компьютерный запоминающий носитель дополнительно содержит машиноисполняемые команды, сохраненные на нем, которые при исполнении компьютером побуждают компьютер обнаруживать аудиометки от удаленной системы видеоконференцсвязи во множестве принятых аудиосигналов и выбирать аудиосигнал, который не содержит аудиометку, 25 для объединения со сформированной аудиометкой.

10. Компьютерный запоминающий носитель по п. 8, причем компьютерный запоминающий носитель дополнительно содержит машиноисполняемые команды, сохраненные на нем, которые при исполнении компьютером побуждают компьютер: формировать видеометку;

30 объединять видеосигнал с видеометкой для получения помеченного видеосигнала;

и

передавать помеченный аудиосигнал и помеченный видеосигнал.

11. Компьютерный запоминающий носитель по п. 10, причем компьютерный запоминающий носитель дополнительно содержит машиноисполняемые команды, сохраненные на нем, которые при исполнении компьютером побуждают компьютер обнаруживать видеометки от удаленной системы видеоконференцсвязи во множестве принятых видеосигналов и выбирать видеосигнал, который не содержит видеометку, для объединения со сформированной аудиометкой.

12. Компьютерный запоминающий носитель по п. 8, причем компьютерный запоминающий носитель дополнительно содержит машиноисполняемые команды, сохраненные на нем, которые при исполнении компьютером побуждают компьютер встраивать дополнительную информацию в по меньшей мере одно из сформированной видеометки или аудиометки.

13. Компьютерный запоминающий носитель по п. 10, причем компьютерный запоминающий носитель дополнительно содержит машиноисполняемые команды, сохраненные на нем, которые при исполнении компьютером побуждают компьютер объединять принятый аудиосигнал с аудиометкой для получения помеченного аудиосигнала таким образом, чтобы искажение помеченного аудиосигнала было меньше

заданного уровня, или объединять принятый видеосигнал с видеометкой для получения помеченного видеосигнала, в котором искажение помеченного видеосигнала будет меньше заданного уровня.

5

10

15

20

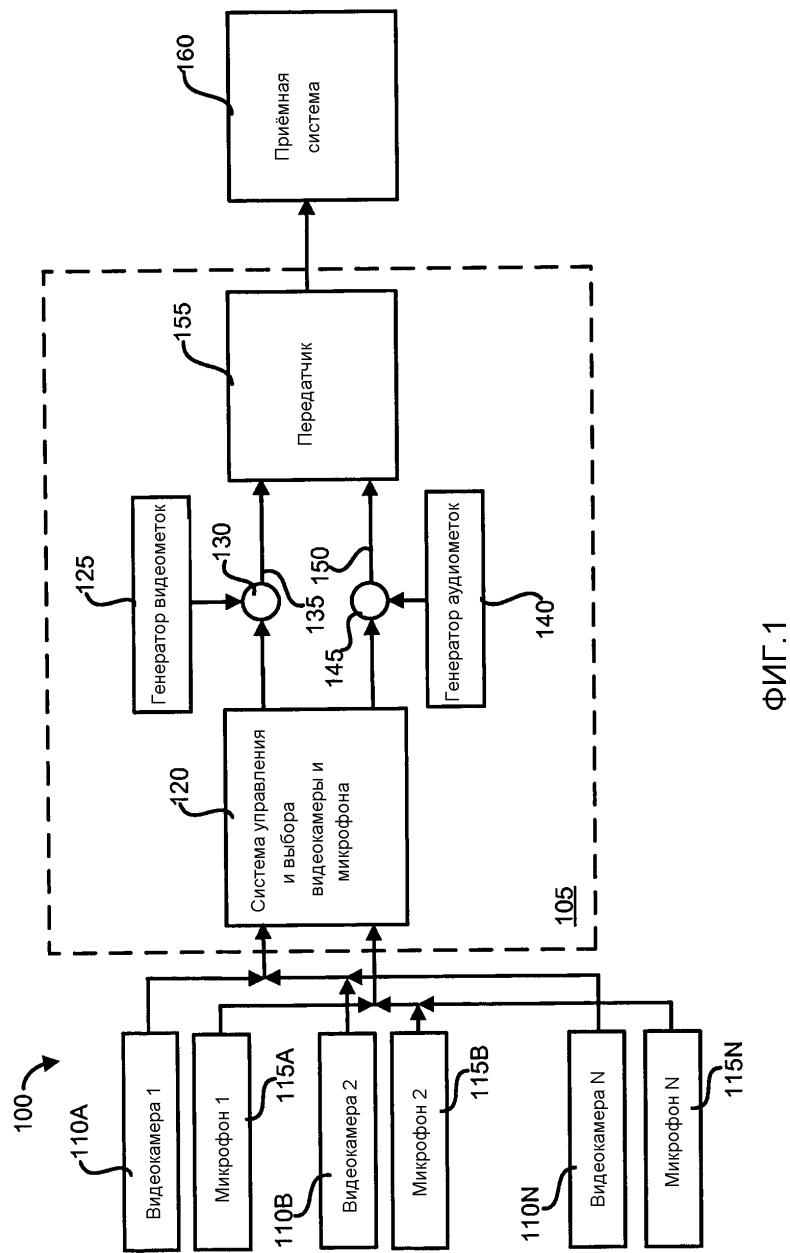
25

30

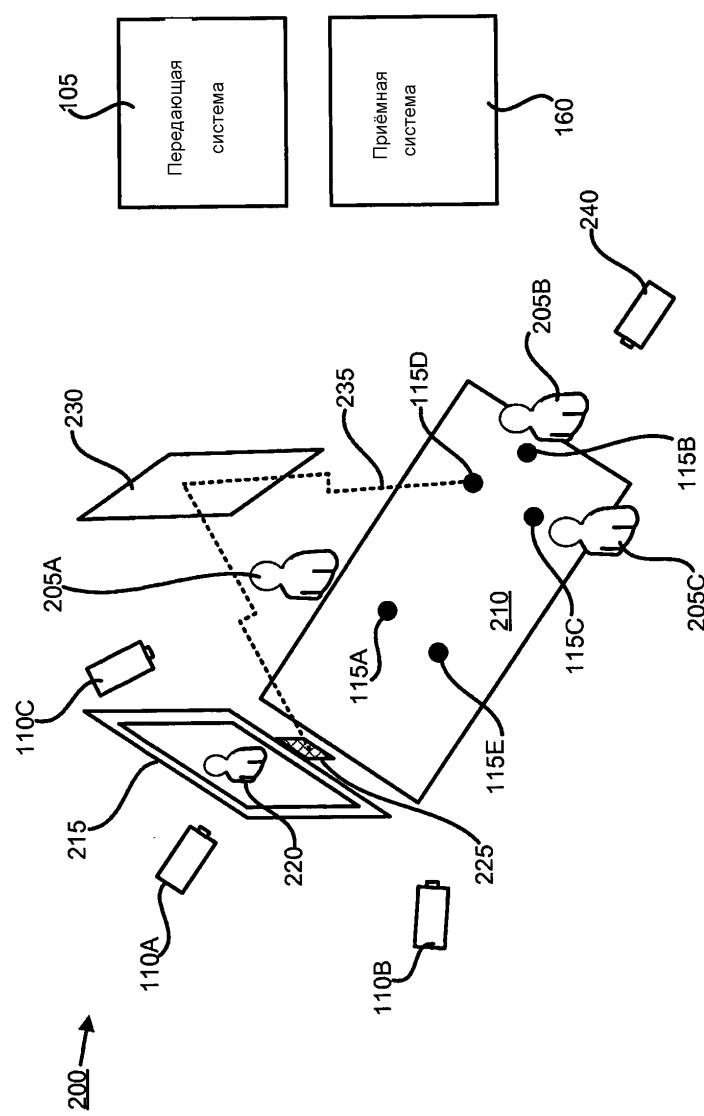
35

40

45

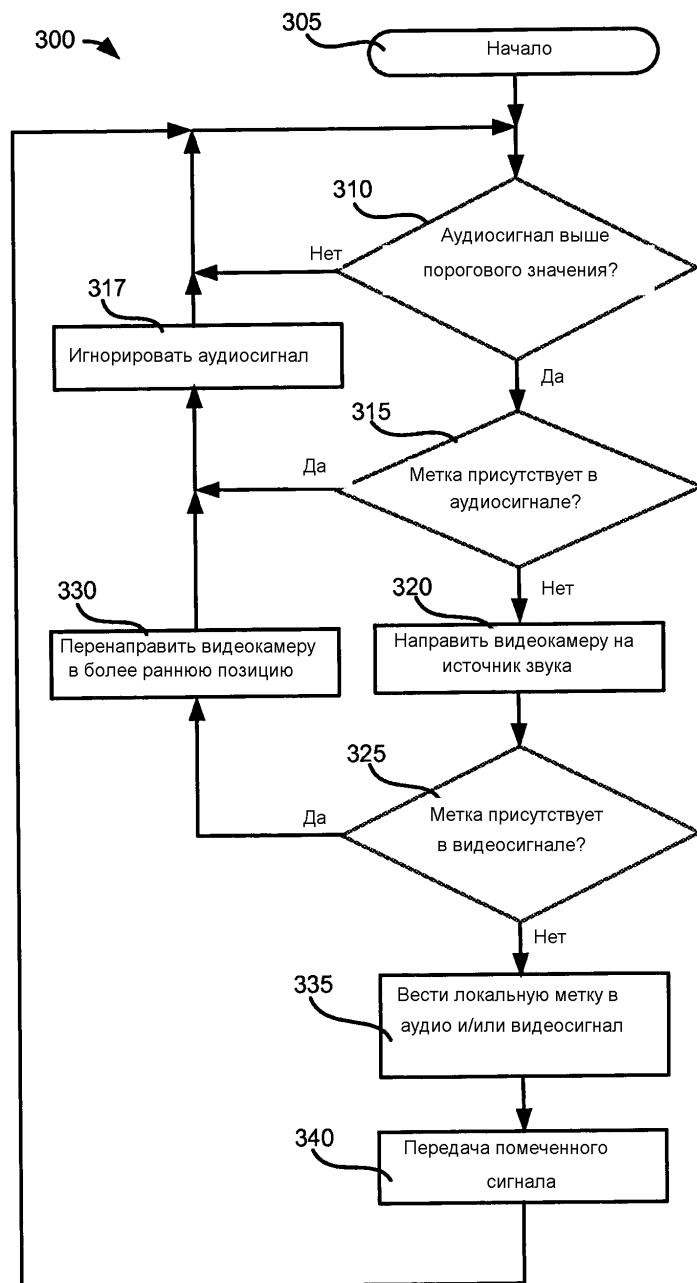


2/5



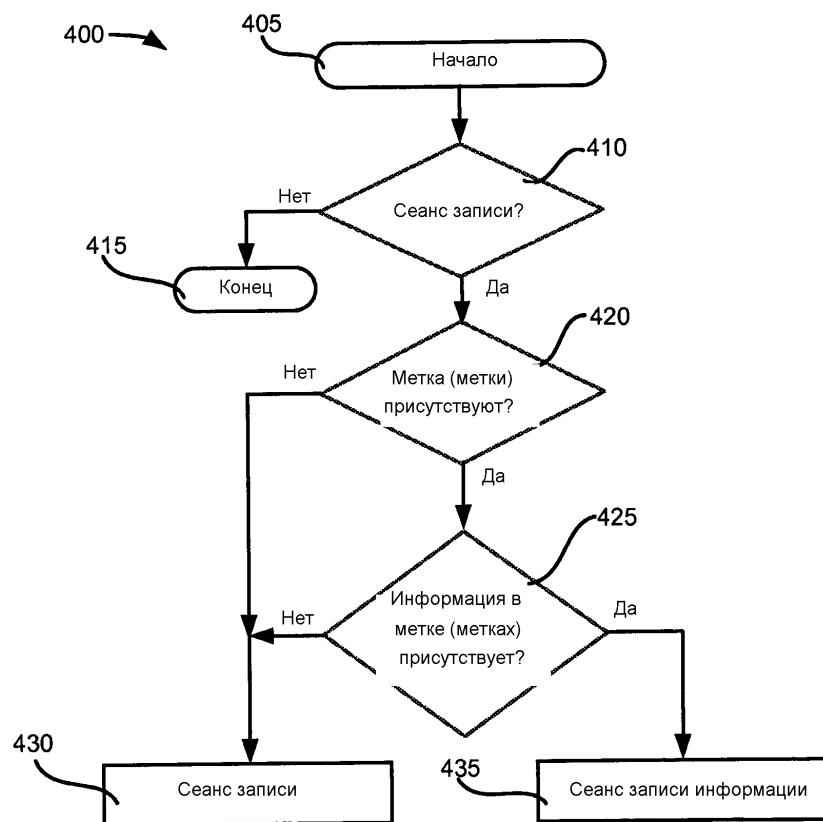
ФИГ.2

3/5



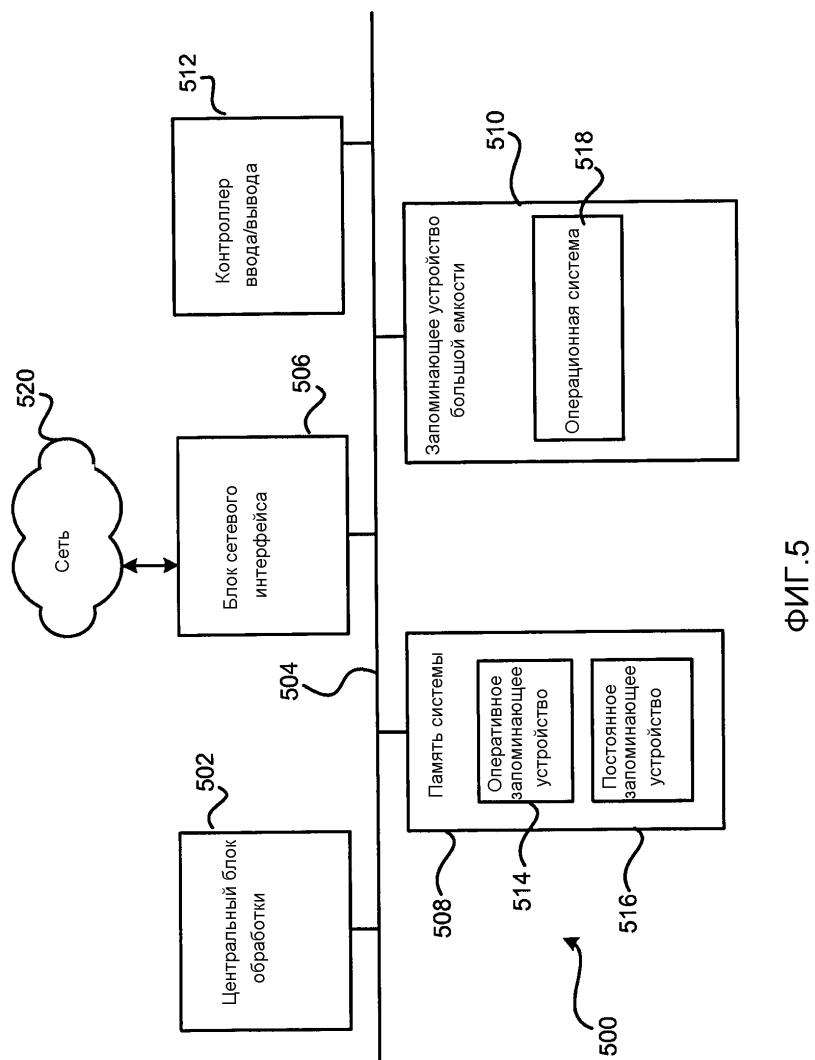
ФИГ.3

4/5



ФИГ.4

5/5



ФИГ.5