



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

H04N 21/2343 (2006.01); H04N 21/4223 (2006.01); H04N 7/141 (2006.01); G06K 9/00369 (2006.01)

(21)(22) Заявка: 2016119109, 17.11.2014

(24) Дата начала отсчета срока действия патента:
17.11.2014Дата регистрации:
10.07.2018

Приоритет(ы):

(30) Конвенционный приоритет:
19.11.2013 US 14/084,090

(43) Дата публикации заявки: 23.11.2017 Бюл. № 33

(45) Опубликовано: 10.07.2018 Бюл. № 19

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 18.05.2016(86) Заявка РСТ:
US 2014/065855 (17.11.2014)(87) Публикация заявки РСТ:
WO 2015/077159 (28.05.2015)

Адрес для переписки:

129090, Москва, ул. Б.Спасская, 25, строение 3,
ООО "Юридическая фирма Городисский и
Партнеры"

(72) Автор(ы):

БЕРРЕТТ Дуглас Д. (US),
ДРАГНЕ Лусиан (US),
КОЛЛЕНКАРК Син И. (US)

(73) Патентообладатель(и):

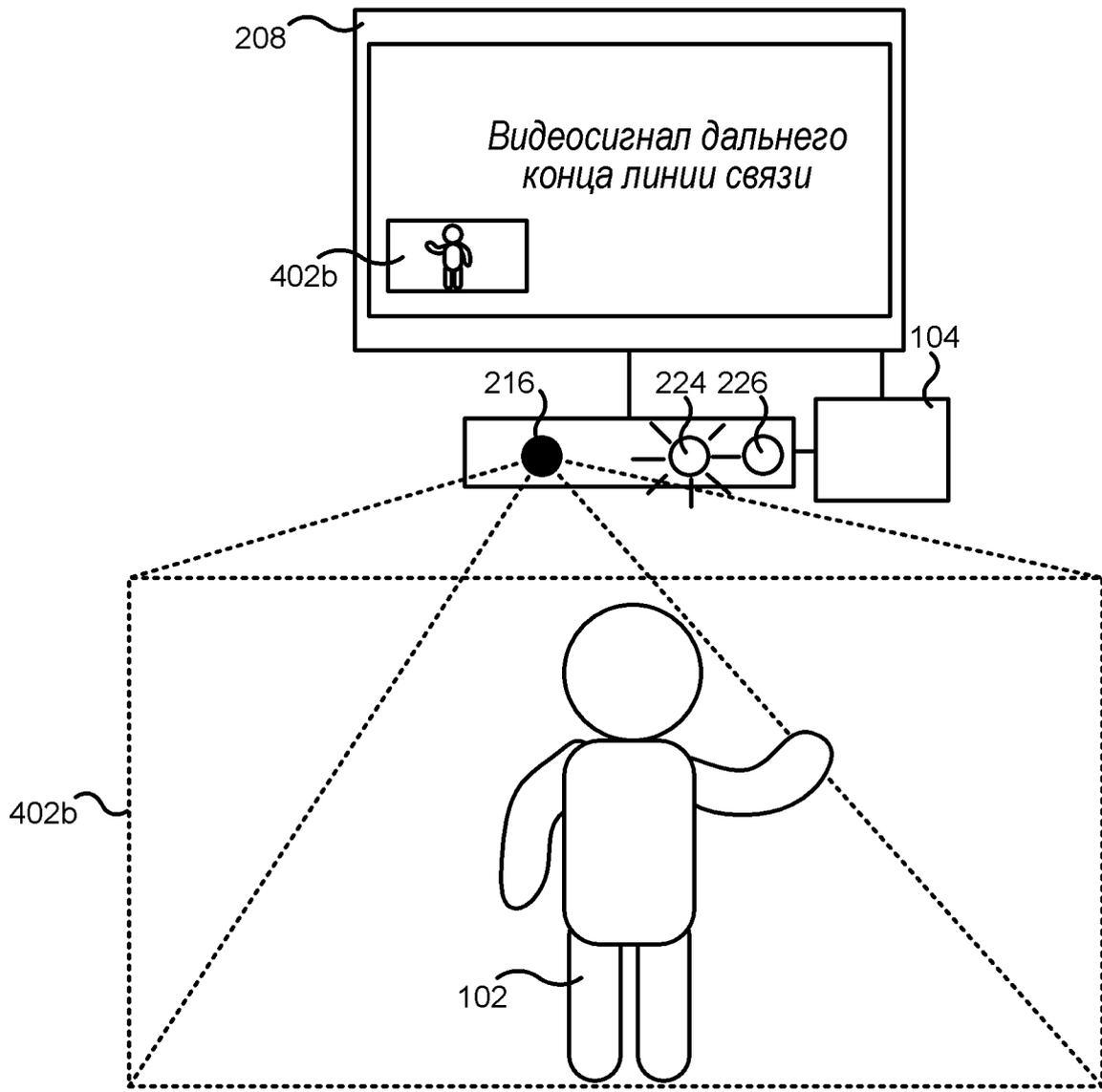
МАЙКРОСОФТ ТЕКНОЛОДЖИ
ЛАЙСЕНСИНГ, ЭлЭлСи (US)(56) Список документов, цитированных в отчете
о поиске: US 2012092445 A1, 2012-04-19.ТАКАYUKI NAKAMURA, Real-time 3-D
object tracking using Kinect sensor,
ROBOTICS AND BIOMIMETICS (ROBIO),
IEEE, 2011. BENABDELKADER C. et al,
Single camera multiplexing for multi-target
tracking, IMAGE ANALYSIS AND
PROCESSING, 1999. PROCEEDINGS.
INTERNATIONAL CONFERENCE ON
VENICE, Italy 27-29 Sept. 1999. US 2010119157
(см. прод.)

(54) ПЕРЕДАЧА ВИДЕОСИГНАЛА

(57) Реферат:

Изобретение относится к области телекоммуникаций, в частности к протоколу передачи голоса или видеосигнала по интернету (VoIP). Технический результат заключается в обеспечении высокого качества кодирования голоса и видеосигнала. Предложен способ передачи видеосигнала по сети, где способ содержит: передачу видеосигнала одного или более пользователей, принятого от устройства захвата изображения, по меньшей мере другому пользовательскому устройству по сети; прием информации о канале связи между

пользовательским устройством и другим пользовательским устройством и/или об одном или более ресурсах пользовательского устройства и/или другого пользовательского устройства; выбор характеристик из множества визуальных пользовательских характеристик на основании принятой информации и управление видеосигналом на основании обнаружения выбранных характеристик для отслеживания выбранных характеристик. 2 н. и 9 з.п. ф-лы, 13 ил.



ФИГ. 4В

(56) (продолжение):
 А1, 2010-05-13. RU 2225682 С2, 2004-03-10.



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(51) Int. Cl.
H04N 21/24 (2011.01)
G06T 7/80 (2017.01)
H04N 7/15 (2006.01)

(12) **ABSTRACT OF INVENTION**

(52) CPC

H04N 21/2343 (2006.01); *H04N 21/4223* (2006.01); *H04N 7/141* (2006.01); *G06K 9/00369* (2006.01)

(21)(22) Application: **2016119109, 17.11.2014**

(24) Effective date for property rights:
17.11.2014

Registration date:
10.07.2018

Priority:

(30) Convention priority:
19.11.2013 US 14/084,090

(43) Application published: **23.11.2017 Bull. № 33**

(45) Date of publication: **10.07.2018 Bull. № 19**

(85) Commencement of national phase: **18.05.2016**

(86) PCT application:
US 2014/065855 (17.11.2014)

(87) PCT publication:
WO 2015/077159 (28.05.2015)

Mail address:
**129090, Moskva, ul. B.Spaskaya, 25, stroenie 3,
OOO "Yuridicheskaya firma Gorodisskiji Partnery"**

(72) Inventor(s):

**BERRETT Douglas D. (US),
DRAGNE Lusian (US),
KOLLENKARK Sin I. (US)**

(73) Proprietor(s):

**MAJKROSOFT TEKNOLODZHI
LAJSENSING, EIEISi (US)**

(54) **VIDEO TRANSMISSION**

(57) Abstract:

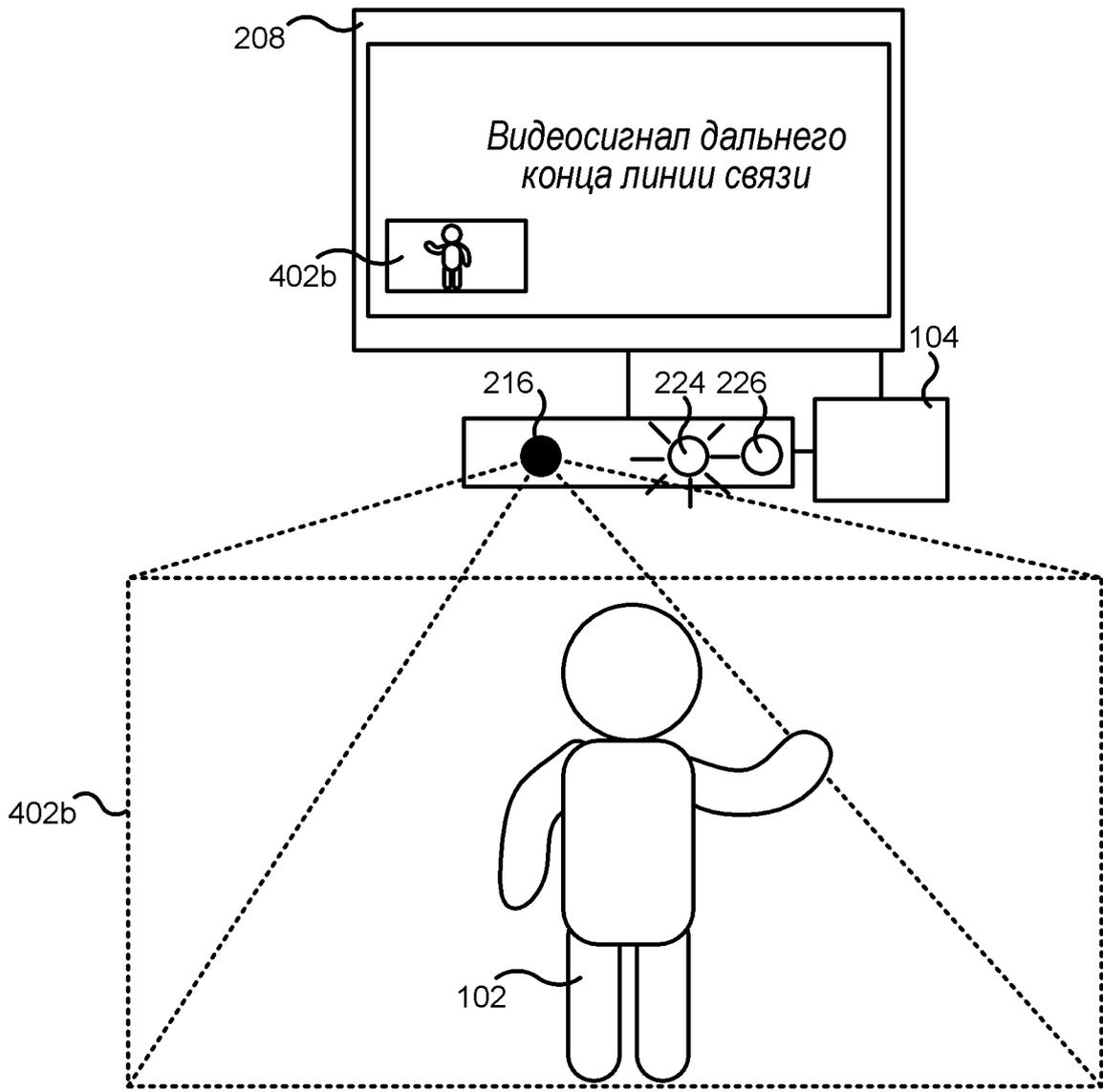
FIELD: electrical communication equipment.

SUBSTANCE: invention relates to the field of telecommunications, in particular to voice or video over Internet protocol (VoIP). Disclosed is a method for transmitting video over a network, the method comprising: transmitting video of one or more users, received from an image capture device, to at least another user device via the network; receiving information about a communication channel between the user device and the other user device and/or about

one or more resources of the user device and/or the other user device; selecting characteristics from a plurality of visual user characteristics based on the received information and controlling the video based on detection of the selected characteristics to track the selected characteristics.

EFFECT: technical result is to provide high quality voice and video coding.

11 cl, 13 dwg



ФИГ. 4В

УРОВЕНЬ ТЕХНИКИ

[1] Традиционные системы связи позволяют пользователю устройства, как, например, персонального компьютера или мобильного устройства, управлять голосовыми или видеозвонками по компьютерной сети на основании пакетной передачи данных, как, например, интернет. Такие системы связи включают в себя системы протокола (VoIP) передачи голоса или видеосигнала по интернету. Эти системы выгодны для пользователя, так как они часто имеют существенно меньшую стоимость, чем традиционные стационарные линии связи или мобильные сотовые сети. Более конкретно, это относится к случаю дальней связи. Для использования системы VoIP, пользователь устанавливает и выполняет клиентское программное обеспечение на своем устройстве. Клиентское программное обеспечение устанавливает соединения VoIP, а также обеспечивает другие функции, как, например, регистрацию и идентификацию пользователя. Дополнительно к голосовой связи, клиент может также устанавливать соединения для других носителей связи, как, например, быстрые сообщения («IM»), сообщения SMS, передача файлов и голосовая почта.

[2] В последнее время, возможности и функциональность интернета были интегрированы в пользовательские устройства, как, например, игровые приставки, выполненные для соединения с телевизионным приемником других (например, большой экран) средств отображения, сами телевизионные приемники (часто называемые «умный телевизор»), телевизионные абонентские приставки, выполненные для соединения с телевизионным приемником, и так далее. Это включает в себя интеграцию клиентского программного обеспечения в игровую приставку, телевизионный приемник, телевизионную абонентскую приставку (или тому подобное) для обеспечения связи по компьютерной сети на основании пакетной передачи данных, как, например, интернет. Эта интеграция клиентского программного обеспечения позволяет использовать большой экран с высоким разрешением для видео звонков путем вывода видеосигналов пользователю ближнего конца линии связи, полученного от пользователя дальнего конца линии связи. Дополнительно, значительная мощность обработки может быть обеспечена пользовательским устройствам, как, например, телевизионным абонентским приставкам, телевизорам и так далее, более конкретно, требования мощности для большого потребительского электронного устройства, заряжающегося от сетевого электричества, менее жесткие, чем, например, для мобильных устройств. Это может позволить дополнительный ряд признаков, включенных во встроеного клиента связи, как, например, высокое качество кодирования голоса и видеосигнала для данных видеосигнала, принятых от камеры, соединенной с пользовательским устройством, или от другого подобного средства ввода изображения.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[3] Это описание сущности изобретения обеспечено для того, чтобы в упрощенной форме представить выбор концепций, которые далее описаны в подробном описании. Это описание сущности изобретения не предназначено для выявления ключевых признаков или существенных признаков заявленного изобретения, а также не предназначено для использования в качестве ограничения объема заявленного изобретения.

[4] Раскрыто пользовательское устройство, содержащее сетевой интерфейс, видеоконтроллер и средство управления ресурсами. Сетевой интерфейс выполнен с возможностью передавать видео, содержащее одного или более пользователей, принятое от устройства захвата изображения, по меньшей мере, одному другому пользовательскому устройству по сети. Видеоконтроллер выполнен с возможностью

выбирать одну или более характеристик из множества визуальных пользовательских характеристик и управлять видеосигналом на основании обнаружения выбранных характеристик для отслеживания выбранных характеристик. Средство управления ресурсами выполнено с возможностью принимать информацию о канале связи между
5 пользовательским устройством и другим пользовательским устройством и/или об одном или более ресурсах пользовательского устройства и/или другого пользовательского устройства и, на основании принятой информации, управлять указанным выбором посредством видеоконтроллера.

[5] Также раскрыто пользовательское устройство, содержащее сетевой интерфейс
10 для соединения с сетью и одно или более устройств обработки. Одно или более устройств обработки выполнены с возможностью передавать видеосигнал одного или более пользователей, принятый от устройства захвата изображения, другому пользовательскому устройству по сети. Одно или более устройств обработки
15 дополнительно выполнены с возможностью выбирать характеристики из множества визуальных пользовательских характеристик и формировать первую и вторую группы данных о границе на основании обнаружения выбранных характеристик в первый и второй моменты времени, соответственно. Одно или более устройств обработки
20 дополнительно выполнены с возможностью формировать данные о переходе на основании первой и второй групп данных о границе с использованием динамической модели и управлять видеосигналом на основании данных о переходе для отслеживания выбранных характеристик.

[6] Также раскрыты соответствующие способы и соответственные компьютерные программные продукты, содержащие выполняемый код, выполненные для каждого из этих способов.

25 **КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ**

[7] Для понимания настоящего изобретения и для иллюстрации того, как его можно реализовать, будет сделана ссылка путем примера на следующие чертежи, на которых:

[8] Фиг. 1 является схематической иллюстрацией системы связи;

[9] Фиг. 2 является схематической блочной диаграммой пользовательского
30 устройства;

[10] Фиг. 3 является функциональной диаграммой способа управления видеосигналом;

[11] Фиг. 4А и 4В являются схематическими иллюстрациями пользователя, взаимодействующего с пользовательским устройством;

[12] Фиг. 5 является схематической иллюстрацией диаграммы направленности
35 излучения;

[13] Фиг. 6А, 6В и 6С являются схематическими передним, горизонтальным и боковым видами пользователя, соответственно;

[14] Фиг. 7А является схематической иллюстрацией обнаруженных скелетных точек
40 пользователя;

[15] Фиг. 7В является схематической иллюстрацией экстраполированных частей тела
пользователя;

[16] Фиг. 8 является блок-схемой способа управления видеосигналом;

[17] Фиг. 9 является блок-схемой способа управления видеосигналом во времени.

ПОДРОБНОЕ ОПИСАНИЕ ОСУЩЕСТВЛЕНИЯ ИЗОБРЕТЕНИЯ

[18] В данной заявке раскрыт способ, с помощью которого во время видео звонка
45 (один-на-один или конференц-вызов) местоположение пользователя в комнате обнаруживают с использованием средства обнаружения глубины (устройство обнаружения глубины или подобное), и видеосигналом звонка, передаваемого другому

пользователю(ям), управляют для отслеживания пользователя на основании этого обнаружения. Более конкретно, соответственные визуальные характеристики этого пользователя (как, например, части тела этого пользователя) также обнаруживают, и видеосигналом управляют для отслеживания выбранных, необязательно всех, из этих характеристик (например, частей тела). Отслеживают или нет конкретную визуальную характеристику, зависит от условий канала связи (например, ширина полосы пропускания канала) и/или ресурсов устройства ближнего конца линии связи и/или дальнего конца линии связи - например, меньше характеристик (например, частей тела) можно отследить для более низкой ширины полосы пропускания канала и/или меньшего размера экрана устройства дальнего конца линии связи, и больше характеристик (например, частей тела) можно отследить для более высокой полосы пропускания канала и/или большего размера экрана устройства дальнего конца линии связи.

[19] Фиг. 1 изображает систему 100 связи, содержащую первого пользователя 102a («Пользователь А»), который связан с первым пользовательским устройством 104, второго пользователя 108 («Пользователь В»), который связан со вторым пользовательским устройством 110, третьего пользователя 112 («Пользователь С»), который связан с третьим пользовательским устройством 114, и четвертого пользователя 102b («Пользователь D») поблизости с первым пользователем. В других осуществлениях, система 100 связи может содержать любое число пользователей и связанных пользовательских устройств. Пользовательские устройства 104, 110 и 114 могут связываться по сети 106 в системе 100 связи, тем самым, позволяя пользователям 102a, 102b, 108 и 112 связываться друг с другом по сети 106. Система 100 связи, изображенная на Фиг. 1, является системой связи на основании пакетной передачи данных, но можно использовать и другие типы систем связи. Сеть 106, например, может быть интернетом. Каждое из пользовательских устройств 104 и 110, например, может быть мобильным телефоном, планшетом, ноутбуком, персональным компьютером («ПК») (включая, например, ПК Windows™, Mac OS™ and Linux™), игровой приставкой или другим игровым устройством, телевизором, персональным цифровым помощником («PDA») или другим встроенным устройством, способным соединиться с сетью 106. Пользовательское устройство 104 приспособлено для приема информации от и вывода информации пользователям 102a, 102b. Пользовательское устройство 104 содержит средство вывода, как, например, устройство отображения и динамики. Пользовательское устройство 104 также содержит средство ввода, как, например, кнопочная панель, сенсорный экран, микрофон, для приема аудио сигналов и/или камеру для захвата изображения видеосигнала, содержащего последовательность видеокадров (изображений). Пользовательское устройство 104 соединено с сетью 106.

[20] Пользовательское устройство 104 выполняет экземпляр клиента связи, обеспеченный провайдером программного обеспечения, связанного с системой 100 связи. Клиент связи является программой, выполняемой на локальном устройстве обработки в пользовательском устройстве 104. Клиент выполняет обработку, требуемую в пользовательском устройстве 104, чтобы пользовательское устройство 104 передавало и принимало данные по системе 100 связи.

[21] Пользовательское устройство 110 соответствует пользовательскому устройству 104 и выполняет, на локальном устройстве обработки, клиента связи, который соответствует клиенту связи, выполняемому в пользовательском устройстве 104. Клиент в пользовательском устройстве 110 выполняет обработку, требуемую, чтобы позволить пользователю 108 связываться по сети 106 тем же образом, что и клиент в пользовательском устройстве 104 выполняет обработку, требуемую, чтобы позволить

пользователям 102a и 102b связываться по сети 106. Пользовательское устройство 114 соответствует пользовательскому устройству 104 и выполняет, на локальном устройстве обработки, клиента связи, который соответствует клиенту связи, выполняемому в пользовательском устройстве 104. Клиент в пользовательском устройстве 110 выполняет
5 обработку, требуемую, чтобы позволить пользователю 108 связываться по сети 106 тем же образом, что и клиент в пользовательском устройстве 104 выполняет обработку, требуемую, чтобы позволить пользователям 102a и 102b связываться по сети 106. Пользовательские устройства 104, 110 и 114 являются конечными точками в системе 100 связи. Фиг. 1 изображает только четырех пользователей (102a, 102b, 108 и 112) и
10 три пользовательских устройства (104 и 110) для ясности, но больше пользователей и пользовательских устройств может быть включено в систему 100 связи и может связываться по системе 100 связи с использованием соответственных клиентов связи, выполняемых на соответственных пользовательских устройствах.

[22] Фиг. 2 изображает детальный вид пользовательского устройства 104, на котором
15 выполняется экземпляр 206 клиента связи для связи по системе 100 связи.

Пользовательское устройство 104 содержит устройство обработки или «модуль 202 обработки», причем устройство обработки содержит один или более центральных блоков («CPU») обработки, с которым соединены: устройства вывода, как, например, устройство 208 отображения, которое может быть реализовано в качестве сенсорного
20 экрана, динамик (или «громкоговоритель») 210 для вывода аудио сигналов и проектор 224 излучения для вывода невидимого излучения, как, например, инфракрасный (IR) проектор; устройства ввода, как, например, микрофон 212 для приема аудио сигналов, камера 216 для приема данных изображения, кнопочная панель 218 и решетка 226 датчика излучения для приема невидимых изображений сигнала невидимого излучения,
25 например, IR камера для приема изображений инфракрасного сигнала; запоминающее устройство 214 для хранения данных; и сетевой интерфейс 220, как, например, модем, для связи с сетью 106. Динамик 210 соединен с устройством 202 обработки по модулю 209 обработки аудио сигналов. Пользовательское устройство 104 может содержать другие элементы, отличные от изображенных на Фиг. 2. Устройство 208 отображения,
30 динамик 210, микрофон 212, запоминающее устройство 214, камера 216, кнопочная панель 218 и сетевой интерфейс 220 могут быть интегрированы в пользовательское устройство 104. Альтернативно, одно или более устройств 208 отображения, динамик 210, микрофон 212, запоминающее устройство 214, камера 216, кнопочная панель 218 и сетевой интерфейс 220 могут не быть интегрированы в пользовательское устройство
35 104 и могут быть соединены с устройством 202 обработки по соответствующим интерфейсам. Одним примером такого интерфейса является интерфейс универсальной последовательной шины (USB). Например, для пользовательских устройств в форме игровых приставок, камера 216, проектор 224, датчик 226, устройство 208 отображения, микрофон 212, кнопочная панель 218 (например, игровое средство управления) могут
40 не быть интегрированы в игровую приставку. Если соединение пользовательского устройства 104 с сетью 106 по сетевому интерфейсу 220 является беспроводным соединением, то сетевой интерфейс 220 может включать в себя антенну для беспроводной передачи сигналов сети 106 и беспроводного приема сигналов от сети 106.

[23] Проектор 224 и датчик 226 составляют устройство 222 обнаружения глубины
45 для захвата данных невидимого излучения в трех измерениях («3D»). В этом осуществлении, проектор 224 проецирует диаграмму направленности излучения, направляет датчику 226, который обнаруживается датчиком 226; данные датчика от датчика 226 используют для построения 3D изображения на основании искажений в

обнаруженной диаграмме направленности излучения (как более подробно объясняется ниже со ссылкой на Фиг. 5 и 6А-6С). Устройство 222 обнаружения глубины и камера 216 могут быть размещены вместе в одном блоке, внешнем к пользовательскому устройству 104, возможно, имеющему источник питания, отдельный от
5 пользовательского устройства 104, соединенный с устройством 202 обработки по подходящему интерфейсу (например, USB или основанному на USB). Примером такого блока является Microsoft Kinect Sensor™.

[24] Фиг. 2 также изображает операционную систему («ОС») 204, выполняемую на устройстве 202 обработки. На ОС 204 функционирует программное обеспечение
10 экземпляра 206 клиента системы 100 связи. Операционная система 204 управляет ресурсами технических средств пользовательского устройства 104 и обрабатывает данные, передаваемые от и к сети 106 по сетевому интерфейсу 220. Клиент 206 связывается с операционной системой 204 и управляет соединениями по системе связи. Клиент 206 имеет пользовательский интерфейс клиента, который используют для
15 представления информации пользователю 102 и приема информации от пользователя 102. Таким образом, клиент 206 выполняет обработку, требуемую, чтобы позволить пользователю 102 связываться по системе 100 связи.

[25] Со ссылкой на Фиг. 3, теперь будет описан способ передачи видеосигнала по сети. Фиг. 3 является функциональной диаграммой части пользовательского устройства
20 104.

[26] Как изображено на Фиг. 3, пользовательское устройство 104 содержит устройство 308 кодирования (например, устройство H264 кодирования). Оно может быть устройством кодирования технических средств, которому поставляют видеоданные с помощью устройства 202 обработки для кодирования, до передачи по сети 106,
25 устройством кодирования программного обеспечения, реализованном в программном обеспечении, выполняемом (например) на устройстве 202 обработки, или их комбинацией. Пользовательское устройство дополнительно содержит блок 302 управления (реализованный, например, как часть программного обеспечения клиента 106), камеру 216, устройство 222 обнаружения глубины и устройство 310 обработки
30 глубины. Блок 302 управления содержит средство 306 управления ресурсами и устройство 304 обработки видеосигнала. Камера 216 захватывает изображения видеосигнала пользователей 102a, 102b (обозначенные как «видеосигнал ближнего конца линии связи» на Фиг. 3) в видимом спектре (то есть, видимом человеческому глазу) и поставляет эти захваченные изображения первому вводу устройства обработки
35 видеосигнала. Датчик 226 захватывает изображения в невидимом спектре (то есть, невидимом человеческому глазу) и поставляет эти захваченные изображения вводу устройства 310 обработки глубины. Устройство 310 обработки глубины имеет вывод, соединенный со вторым вводом устройства обработки видеосигнала. Проектор 224 проецирует невидимое излучение, идущее от датчика 226, по направлению к
40 пользователям 102a, 102b, которое обнаруживается датчиком 226. Блок 300 управления, устройство 308 кодирования и устройство 310 обработки датчика составляют систему 300 обработки видеосигнала.

[27] Средство 306 управления ресурсами имеет первый, второй и третий входы и вывод. Первый вход средства 306 управления ресурсами выполнен с возможностью
45 принимать информацию об одном или более каналах связи между пользовательским устройством 104 и одним или несколькими другими пользовательскими устройствами (например, 108, 112) сети 106. Второй вход средства 306 управления ресурсами выполнен с возможностью принимать информацию о ресурсах пользовательского устройства

104 и/или одного или более других пользовательских устройствах (например, 110, 114). Третий ввод устройства управления ресурсами соединен с выводом устройства 304 обработки видеосигнала.

5 [28] Вывод устройства 302 управления ресурсами соединен с третьим вводом устройства 304 обработки видеосигнала. Устройство 304 обработки видеосигнала имеет вывод, соединенный с вводом устройства 308 кодирования. Устройство кодирования выполнено с возможностью поставлять закодированные видеоданные, которые могут подвергаться дополнительной обработке в устройстве 104 (например, формирование пакетов) до их поставки в сетевой интерфейс 220 для передачи по сети 106, по меньшей мере, одному из второго пользовательского устройства 110 и третьего пользовательского устройства 114.

15 [29] Устройство 310 обработки датчика функционирует для обработки данных датчика, захваченных и принятых от датчика 226, для обнаружения соответственного множества визуальных пользовательских характеристик для каждого пользователя в поле зрения датчика 226.

[30] Устройство 302 обработки видеосигнала выполнено с возможностью выбирать характеристики из множества визуальных пользовательских характеристик на основании информации, выведенной от средства 306 управления ресурсами, и управлять видеосигналом, поставляемым устройству 308 кодирования, на основании обнаружения 20 выбранных характеристик устройством 310 обработки датчика, причем видеосигнал такой, что блок управления отслеживает выбранные характеристики, то есть чтобы видеосигнал отслеживал выбранные характеристики во времени, в конечном итоге, делая обнаруженные характеристики видимыми в видеосигнале.

25 [31] Далее ссылка будет сделана на Фиг. 4А и 4В. Фиг. 4А и 4В изображают пользователя 102 (например, 102а, 102b) поблизости с пользовательским устройством 104 (которое изображено как игровая приставка в данном осуществлении). Пользовательское устройство соединено с устройством 208 отображения, камерой 216, проектором 224 и датчиком 226, все из которых являются внешними по отношению к 30 пользовательскому устройству 104 в данном осуществлении. На Фиг. 4А, видеосигнал первого видимого участка 402а, ограниченный тем, что включает только верхние части (например, голову, плечи и так далее) пользователя 102, но не нижние части (например, ступни, ноги и так далее) пользователя 102 - захваченный камерой 216, поставляют в устройство 308 кодирования для кодирования и конечной передачи другому 35 пользователю (например, 112, 108) по сети 106. Как изображено на Фиг. 4, версию видеосигнала участка 408а передают другому пользователю, также опционально отображаемому на устройстве 208 отображения, с покрытием видеосигнала, принятого от этого другого пользователя. На Фиг. 4В, видеосигнал второго видимого участка 402b, -который включает в себя как верхние части пользователя 102, так и нижние части пользователя 102 - захваченный камерой 216, поставляют в устройство 308 кодирования 40 для кодирования и конечной передачи другому пользователю по сети 106. Как изображено на Фиг. 4, версию видеосигнала второго участка 408b передают другому пользователю, также опционально отображаемому на устройстве 208 отображения, с покрытием видеосигнала, от этого другого пользователя.

45 [32] Первый и второй видимые участки 402а, 402b имеют соответственный размер и местоположение, определенные модулем управления видеосигнала на основании данных, принятых от средства 306 управления ресурсами, и данных, принятых от устройства 310 обработки датчика, как объясняется ниже.

[33] Фиг. 5 показывает диаграмму направленности излучения, испускаемого

проектором 224. Как изображено на Фиг. 5, диаграмма направленности излучения является систематически неоднородной, содержащей множество систематически расположенных участков меняющейся интенсивности. Более конкретно, диаграмма направленности излучения на Фиг. 5 содержит существенно равноамплитудную эквидистантную антенну решетки точек излучения. Диаграмма направленности излучения является инфракрасной (IR) диаграммой направленности излучения в данном осуществлении и дополнительно обнаруживается датчиком 226. Диаграмма направленности излучения на Фиг. 5 является примером, и предусмотрено использование других диаграмм направленности излучения.

[34] Эта диаграмма 500 направленности излучения проецируется по направлению к датчику 224 с помощью проектора 226. Датчик 224 захватывает изображения диаграммы направленности невидимого излучения, как проецируется в этом поле зрения. Эти изображения обрабатывают с помощью устройства 310 обработки датчика, чтобы вычислить глубины пользователей в поле зрения датчика 226 (эффективно построить трехмерное представление пользователя), тем самым, позволяя распознавать различных пользователей и различные соответственные части их тела.

[35] Фиг. 6В изображает горизонтальный вид пользователя 102 поблизости пользовательского устройства 104, причем Фиг. 6А является соответствующим передним видом (как видно камере 216 и датчику 226 устройства 222 обнаружения), а Фиг. 6С является соответствующим боковым видом. Как изображено, пользователь 102 принял позу, в которой его левая рука протянута по направлению к устройству 222 обнаружения (хотя пользователь, конечно, может принимать другие позы). Голова пользователя выступает за его туловище, причем туловище находится перед его правой рукой. Диаграмма 500 направленности излучения проецируется на пользователя с помощью проектора 224.

[36] Как изображено на Фиг. 6А, пользователь 102, таким образом, имеет форму, которая выступает для искажения проецируемой диаграммы направленности излучения, обнаруженной датчиком 224, где части диаграммы направленности излучения, проецируемые на части пользователя, расположенные дальше от проектора, эффективно растянуты (то есть, в данном случае, так, что точки диаграммы направленности излучения более разделены) по отношению к частям излучения, проецируемого на части пользователя, которые ближе к проектору (то есть, в данном случае, так, чтобы точки диаграммы направленности излучения меньше разделены), причем количество растяжения увеличивается по мере удаления от проектора, а части излучения, проецируемые на объекты, значительно позади пользователя, более эффективно невидимы датчику 226. Так как диаграмма 500 направленности излучения систематически неоднородна, то ее искажения формой пользователя можно использовать для улавливания этой формы (например, для идентификации головы пользователя, левой руки, правой руки, туловища и так далее) устройством 310 обработки датчика, обрабатывающего изображения искаженной диаграммы направленности излучения, захваченной датчиком 226. Например, разделение области пользователя от датчика можно определить путем измерения разделения точек обнаруженной диаграммы направленности излучения внутри этой области пользователя.

[37] Хотя на Фиг. 5 и 6А диаграмма направленности излучения изображена как видимая человеческому глазу, это только для помощи в понимании, и диаграмма направленности излучения, проецируемая на пользователя 102, на самом деле не видна человеческому глазу.

[38] Теперь, как будет описано со ссылкой на Фиг. 7А и 7В, изображения (данные

датчика) диаграммы 500 направленности невидимого излучения, захваченные датчиком 226, обрабатывают устройством 310 обработки датчика. Эта обработка содержит обработку обнаружения скелета. Обработка обнаружения скелета известна в текущем уровне техники и в настоящее время реализуется, например, в Microsoft Xbox 360™ (иногда используется совместно с Microsoft Kinect Sensor™), результаты которой доступны посредством интерфейса (API) программирования приложений для использования разработчиками программного обеспечения.

[39] Устройство 310 обнаружения датчика принимает данные датчика от датчика 226 и обрабатывает их для определения числа пользователей (например, 102a, 102b) в поле зрения датчика 226 и для идентификации соответственного множества скелетных точек для каждого пользователя, использующего способы скелетного обнаружения, которые известны в текущем уровне техники. Каждая скелетная точка представляет собой приблизительное местоположение соответственного сочленения человека в видеосигнале.

[40] Более конкретно, в данном осуществлении, устройство 310 обнаружения датчика обнаруживает двадцать соответственных скелетных точек для каждого пользователя в поле зрения датчика 226. Каждая скелетная точка соответствует одному из двадцати распознанных сочленений человека, причем каждое меняется в пространстве и времени, по мере того, как пользователь (или пользователи) движется внутри поля зрения датчика 226. Местоположение этих сочленений в любой момент времени вычисляют на основании трехмерной формы пользователя (или соответственных пользователей), обнаруженной датчиком 226. Скелетная точка также имеет состояние отслеживания: оно может быть «отслеженное» как явно видимое сочленение, «предполагаемое», когда сочленение не явно видимо, но устройство 310 обработки датчика предполагает его местоположение, или «не отслеженное», например, для нижнего сочленения при отслеживании в режиме сидения (то есть, когда обнаружено, что пользователь сидит, и нижние сочленения не отслежены устройством 310 обнаружения датчика).

[41] Каждая скелетная точка может быть обеспечена соответственной величиной уровня достоверности, указывающей вероятность соответствующего корректно обнаруженного сочленения. Точки с величинами уровня достоверности ниже конкретного порога могут быть исключены из обработки модулем 304 видеосигнала.

[42] Эти двадцать скелетных точек изображены на Фиг. 7А с соответствующим сочленением человека, изображенным в Таблице 1.

Таблица 1		
Скелетные точки		
Имя скелетной точки	Соответствующие сочленения человека	Обозначено как
AnkleLeft	Левая лодыжка	722b
AnkleRight	Правая лодыжка	722a
ElbowLeft	Левый локоть	706b
ElbowRight	Правый локоть	706a
FootLeft	Левая ступня	724b
FootRight	Правая ступня	724a
HandLeft	Левая рука	702b
HandRight	Правая рука	702a
Head	Голова	710
HipCenter	Центр между бедрами	716
HipLeft	Левое бедро	718b
HipRight	Правое бедро	718a
KneeLeft	Левое колено	720b
KneeRight	Правое колено	720a

ShoulderCenter	Центр между плечами	712
ShoulderLeft	Левое плечо	708b
ShoulderRight	Правое плечо	708a
Spine	Спина	714
WristLeft	Левое запястье	704b
WristRight	Правое запястье	704a

[43] Скелетные точки и видеосигнал от камеры 216 коррелируют так, что местоположение скелетной точки, зарегистрированное устройством обработки датчика в конкретный момент времени, соответствует местоположению соответствующего сочленения человека внутри видеокадра (изображения) в этот момент времени.

Устройство 310 обнаружения датчика поставляет эти обнаруженные скелетные точки в качестве данных скелетных точек видеоконтроллеру 304 для использования. Для каждого кадра данных видеосигнала, данные скелетных точек, поставленные устройством 310 обработки датчика содержат местоположения скелетных точек внутри кадра, например, выраженные как декартовы прямоугольные координаты (x, y) системы координат, ограниченные по отношению к размеру видеокадра.

[44] Видеоконтроллер 304 принимает обнаруженные скелетные точки для одного или более пользователей (102a, 102b) и выполнен с возможностью определять на их основании множество визуальных характеристик этого пользователя (или соответствующее множество визуальных характеристик этих пользователей). В этом осуществлении, визуальные пользовательские характеристики принимают форму частей тела человека. Части тела обнаруживаются видеоконтроллером, причем каждую обнаруживают путем исследования из одной или более скелетных точек, обеспеченных устройством 310 обработки видеосигнала и соответствующих участку внутри соответствующего видеокадра видеосигнала от камеры 216 (то есть, определенный как участок внутри вышеупомянутой системы координат). Фиг. 7В изображает обнаруженные части тела, которые были обнаружены на основании скелетных точек на Фиг. 7А.

Части тела (Визуальные характеристики)	
Название части тела (визуальной характеристики)	Эта характеристика, обнаруженная на Фиг. 7В
Голова	750
Плечи	752
Средняя часть спины (верхняя часть спины)	756
Нижняя часть спины (нижняя часть спины)	758
Вся спина (спина полностью)	760
Бедра	762
Локти	754a, 754b
Лодыжки	764a, 764b
Ступни	766a, 766b

[45] Необходимо отметить, что эти визуальные характеристики являются визуальными в том смысле, что они представляют собой признаки тела пользователя, которые в действительности можно увидеть и различить; однако, в этом осуществлении, они не «видимы» видеоконтроллеру (например, они не обнаружены в видеоданных, захваченных камерой 216); скорее напротив, видеоконтроллер экстраполирует (приблизительно определяет) относительное местоположение, форму и размер этих признаков внутри кадра видеосигнала от камеры 216 из положения двадцати скелетных точек, обеспеченных устройством 310 обработки датчика (а не на основании, например, обработки этого кадра) - например, путем приблизительного определения каждой части

тела как прямоугольника (или подобного), имеющего местоположение и размер (и, опционально, ориентацию), вычисленного из обнаруженных положений скелетных точек, относящихся к этой части тела.

5 [46] Способ 800 управления видеосигналом, подлежащим передаче по сети, на основании обнаруженных пользовательских характеристик (части тела в этом осуществлении) будет описан со ссылкой на Фиг. 8. Фиг. 8 изображает блок-схему способа 800 слева и соответствующее визуальное представление каждого этапа способа справа.

10 [47] В этом осуществлении, способ реализуют алгоритмически как часть программного обеспечения клиента 106 - более конкретно, блоком 302 управления.

[48] Способ 800 будет описан в контексте видео звонка в реальном времени, управляемом с использованием устройства 104 первого пользователя (ближнего конца линии связи) и устройства 110 второго пользователя (дальнего конца линии связи) между их пользователями.

15 [49] Способ рассматривает следующие участки тела:

- Участок 1: Голова, Плечи
- Участок 2: Голова, Плечи, Средняя часть спины, Локти
- Участок 3: Голова, Плечи, Средняя часть спины, Локти, Нижняя часть спины, Бедро
- Участок 4: Голова, Плечи, Вся спина, Локти, Бедро, Лодыжки, Ступни

20 [50] В начале видео звонка, средство 306 управления ресурсами клиента 106 устройства 104 ближнего конца линии связи определяет разрешение видео (которое будет использовано для передачи видеосигнала устройству 110 дальнего конца линии связи) на основании информации, принятой им. Например, эта информация может включать в себя информацию об одном или более из:

25 - сетевых ресурсах, как, например, ширина полосы пропускания канала связи между устройствами ближнего конца линии связи и дальнего конца линии связи (канал, по которому передают видеосигнал от устройства ближнего конца линии связи к устройству дальнего конца линии связи);

30 - ресурсах устройства 110 дальнего конца линии связи, принятых от устройства дальнего конца линии связи, например, размер экрана, тем самым, экран с большим количеством пикселей имеет больше ресурсов экрана, чем экран с меньшим количеством пикселей; соотношение сторон (например, 4:3, 16:9 и так далее) экрана устройства дальнего конца линии связи, являющегося соотношением горизонтальной длины экрана и вертикальной длины экрана; ресурсах обработки одного или более устройств

35 обработки устройства дальнего конца линии связи (например, доступное устройство обработки циклически повторяет в мегагерцах, гигагерцах и так далее); ресурсах памяти устройства дальнего конца линии связи (например, доступное запоминающее устройство в мегабайтах, гигабайтах и так далее); возможностях декодирования видеосигнала;

40 - ресурсах устройства ближнего конца линии связи, например, ресурсах обработки одного или более устройств обработки устройства ближнего конца линии связи (например, доступное устройство обработки циклически повторяет в мегагерцах, гигагерцах и так далее); ресурсах памяти устройства дальнего конца линии связи (например, доступное запоминающее устройство в мегабайтах, гигабайтах и так далее); возможностях кодирования видеосигнала.

45 [51] Разрешение видео можно определить, по меньшей мере, частично, путем переговоров с устройством 110 дальнего конца линии связи (например, тем самым средство 306 управления ресурсами ближнего конца линии связи запрашивает информацию о ресурсах от устройства дальнего конца линии связи).

[52] Во время звонка, средство 306 управления ресурсами устройства 104 ближнего конца линии связи наблюдает за доступными ресурсами (например, доступной шириной полосы пропускания), принимает решения для увеличения или уменьшения разрешения для видеосигнала, который отправляют устройству 108 дальнего конца линии связи, и передает эти решения видеоконтроллеру 304. Разрешение видео, таким образом, может изменяться динамически во время звонка, например, вследствие нерегулярной ширины полосы пропускания канала, возникающей, например, в одном или обоих устройствах ближнего конца и дальнего конца линии связи, соединенных с сетью 106 по ненадежному беспроводному, например, Wi-Fi, соединению.

[53] Выбор характеристик содержит выбор большего числа характеристик, если принятая информация указывает лучшее качество канала и/или больше ресурсов устройства, и выбор меньшего числа характеристик, если принятая информация указывает худшее качество канала и/или меньше ресурсов устройства. Например, для более низкой ширины полосы пропускания и/или меньшего размера экрана, средство 306 управления ресурсами определяет меньшее разрешение (что ведет к выбору участка с меньшим числом частей тела); для более высокой ширины полосы пропускания и/или большего размера экрана, средство 306 управления ресурсами определяет разрешение больше (что ведет к выбору участка с большим числом частей тела), чем оно определяет для указанной более низкой ширины полосы пропускания и/или указанного меньшего размера экрана.

[54] В этом осуществлении, видеосигнал захватывают от камеры 216 при фиксированном разрешении 1920x1080 пикселей. Однако, как описано ниже, захваченный видеосигнал может быть обрезан (обрезка является удалением внешних частей изображений видеосигнала) до передачи. То есть только выборочные части захваченных видеоданных - захваченных из выбранного видимого участка - поставляют в устройство 308 кодирования для кодирования и последующей передачи в устройство дальнего конца линии связи. Поставка обрезанных видеоданных, таким образом, означает поставку меньшего числа видеоданных устройству 308 кодирования, так как видеоданные вне определенного участка не поставляют устройству кодирования.

[55] Данные скелетных точек принимают видеоконтроллером от устройства 310 обработки датчика на этапе S802. В примере иллюстрации с правой стороны Фиг. 8, изображены скелетные данные для совместно расположенных пользователей 102a и 102b (оба в поле зрения датчика 226). Однако, как будет очевидно, способ 800 можно применить к принятым скелетным данным для любого числа пользователей (одного или более).

[56] На этапе S804, видеоконтроллер 304 выбирает соответственные участки тела для каждого пользователя в поле зрения датчика 216 на основании определенного разрешения, принятого от средства 306 управления ресурсами. В зависимости от это разрешения видео, видеоконтроллер выбирает участки тела из возможного множества участков тела таблицы 2, как указано далее:

- Разрешение 1920x1080: Участок 4
- Разрешение равно или больше, чем 1280x1024: Участок 3
- Разрешение равно или больше, чем 640x480: Участок 2
- Другие разрешения: Участок 1

[57] Или альтернативно, как указано далее:

- Ширина разрешения 1920: Участок 4
- Ширина разрешения равна или больше, чем 1280: Участок 3
- Ширина разрешения равна или больше, чем 640: Участок 2

- Другие разрешения: Участок 1

[58] Выбор характеристик содержит выбор большего числа характеристик для больших разрешений (условия канала лучше и/или больше ресурсов устройства) и выбор меньшего числа характеристик для меньших разрешений (условия канала хуже и/или меньше ресурсов устройства).

[59] Для каждого пользователя 102a, 102b в поле зрения датчика 226, выбранный участок обнаруживают (то есть, обнаруживают соответственные части тела выбранного участка) на основании принятых скелетных данных. Затем видеосигнал обрезают (как описано ниже), до его кодирования и передачи, на основании обнаруженных частей тела (например, для участка 1, видеосигнал обрезают на основании обнаружения головы пользователя 102a, головы пользователя 102b, плеч пользователя 102a и плеч пользователя 102b, но не других частей тела этих пользователей, так как только голова и плечи включены в участок 1). Таким образом, не все скелетные точки обязательно участвуют в обрезке (так как некоторые из них служат только для определения частей тела, которые не являются частью выбранного участка). То есть, управление видеосигналом может быть основано на выборе скелетных точек из множества обнаруженных скелетных точек, но не всех из множества обнаруженных скелетных точек.

[60] Например, на иллюстрации с правой стороны Фиг. 8, выбран участок 2 (голова, плечи, средняя часть спины, локти), и управление видеосигналом, таким образом, основано на характеристике обнаруженного участка 2 как для пользователя 102a (550a на Фиг. 8), так и для пользователя 102b (550b на Фиг. 8).

[61] В качестве части управления видеосигналом (который в этом осуществлении содержит выборочно обрезку видеосигнала в соответствии с определенным разрешением), на этапе S806 видеоконтроллер формирует данные о границе на основании обнаружения частей тела выбранного участка. Данные о границе определяют прямоугольную границу 552 (прямоугольник обрезки). Прямоугольник обрезки формируют из объединения соответственных обнаруженных участков для всех отслеженных пользователей.

[62] На этапе S808, видеоконтроллер 304 модифицирует сформированные данные о границе на основании заранее заданного соотношения сторон (например, соотношение сторон устройства отображения устройства 110 дальнего конца линии связи, как, например, 4:3, 16:9 и так далее), тем самым, прямоугольник обрезки регулируют для заранее определенного соотношения сторон. Отрегулированный прямоугольник обрезки (граничный прямоугольник) изображен как 554 на правой стороне Фиг. 8 (и имеет соотношение сторон 16:9 в этом осуществлении).

[63] На этапе S810, видеосигнал обрезают на основании отрегулированного прямоугольника, как объяснено более подробно ниже. Обрезанный видеосигнал затем масштабируют до разрешения вывода (например, его соответствия экрану устройства дальнего конца линии связи, которому его передают), до его кодирования устройством 308 кодирования.

[64] Способ 800 выполняют для каждого кадра видеоданных. Так как граничные прямоугольники отслеживают пользовательские характеристики - которые могут двигаться по мере того, как пользователь движется, или разделяться, например, когда два пользователя удаляются - между кадрами, то прямоугольники обрезки также перемещаются.

[65] Однако, вместо того, чтобы просто обрезать каждый кадр данных видеосигнала до отрегулированного прямоугольника, определенного для этого кадра (то есть, путем

поставки только данных видеосигнала из части этого кадра, определенных отрегулированным прямоугольником), формируют данные о переходе на основании соответственных отрегулированных прямоугольников, вычисленных для одного или более более ранних кадров, и отрегулированного прямоугольника, вычисленного для

5 текущего кадра. Данные о переходе формируют на основании модели упругой пружины.

[66] В осуществлениях, модель упругой пружины может быть определена следующим образом:

$$10 \quad m * \frac{d^2x}{dt^2} = -k * x - D * \frac{dx}{dt}$$

Где m («масса»), k («жесткость») и D («затухание») являются конфигурируемыми постоянными, а x (перемещение) и t (время) - переменными. То есть, модель ускорения перехода пропорциональна взвешенной сумме перемещения и скорости этого перехода.

[67] Прямоугольник обрезки перемещается в соответствии с моделью упругой пружины. Это сглаживает переходы между кадрами и предотвращает колеблющееся видео. Это также повышает эффективность кодирования по следующей причине. Так как модель упругой пружины эффективно «гасит» движение прямоугольников обрезки, то она сокращает разницу между смежными кадрами, что, как будет очевидно, приводит к более эффективному дифференциальному кодированию.

[68] Теперь описание будет со ссылкой на Фиг. 9. Фиг. 9 изображает блок-схему способа 900 управления видеосигналом, подлежащим передаче пользователю дальнего конца линии связи, во времени, и пример переднего представления способа 900, с правой стороны.

[69] На этапе S902, для текущего видеокadra 956 в последовательности видеокadров захваченного видеосигнала, видеоконтроллер формирует первую группу и одну или более вторых групп данных о границе (например, путем вычисления граничных прямоугольников, имеющих заранее определенные соотношения сторон, как описано выше) на основании обнаружения выбранных пользовательских характеристик (части тела в этом осуществлении) в первый момент времени и один или более вторых моментов времени, соответственно, причем первый момент времени является моментом времени текущего кадра 956, а второй момент(ы) времени является моментом(ами) времени предыдущих кадров. Данные о границе для более ранних кадров могут быть сформированы до данных о границе для более поздних кадров (например, на покадровой основе). Следовательно, S902 может иметь место в периоде, охватывающем несколько видеокadров.

[70] Два примера граничных прямоугольников 950, 952 в соответственные первый момент $t1$ времени и второй момент $t2$ времени изображены с правой стороны Фиг. 9. Прямоугольник для $t2$ изображен больше, чем прямоугольник для $t1$, что может быть, например, вследствие того, что два пользователя (102a, 102b) отслеживаются, и эти два пользователя движутся дополнительно друг от друга, или вследствие перемены в доступных ресурсах (например, увеличение ширины полосы пропускания канала), что побуждает средство 306 управления ресурсами увеличивать определенное соотношение сторон, приводя к участку, в котором выбрано больше частей тела (например, переключение от участка 1 к участку 2), и к тому, что граничный прямоугольник автоматически регулируется видеоконтроллером 304, соответственно.

[71] На этапе S904, модуль 304 видеосигнала формирует данные о переходе на основании первой и второй групп данных о границе, используя динамическую модель, описанную выше. Сформированные данные о переходе эффективно определяют

переходный граничный прямоугольник 954 (изображен на Фиг. 9), представляющий собой граничный прямоугольник текущего кадра 956 из граничных прямоугольников предыдущих кадров.

[72] Например, в соответствии с вышеуказанной моделью упругой пружины, граничные прямоугольники могут быть параметризованы одной или несколькими точками в различных местоположениях (одна форма данных о границе). Точка может иметь второе положение («desiredPosition» [желаемое положение]) в момент t_2 времени, являющийся параметром прямоугольника 952, и первое положение «currentPosition» [текущее положение] в момент t_1 времени, являющийся параметром прямоугольника 950. В таком случае, данные о переходе могут быть сформированы путем обновления «currentPosition», как указано далее, причем обновленное «currentPosition» является параметром переходного граничного прямоугольника 954:

```

velocity=0
previousTime=0
15 currentPosition=<some_constant_initial_value>
UpdatePosition (desiredPosition, time)
{
x=currentPosition - desiredPosition;
force=- stiffness * x - damping * m_velocity;
20 acceleration=force/mass;
dt=time - previousTime;
velocity += acceleration * dt;
currentPosition += velocity * dt;
previousTime=time;
25 }

```

[73] На этапе S908, видеоконтроллер управляет видеосигналом на основании данных о переходе для отслеживания выбранных характеристик. Здесь, это вовлекает обрезку видеосигнала так, что только данные изображения кадра 956, которые находятся внутри граничного прямоугольника 954, поставляют устройству 308 кодирования для кодирования и передачи пользователю 108 дальнего конца линии связи (а не данные изображения кадра 956, которые находятся вне граничного прямоугольника 954).

[74] Устройство 222 обнаружения и устройство 310 обработки датчика выполнены с возможностью так, чтобы пользователи, удаленные больше, чем на 1 метр, от устройства 222 обнаружения, не отслеживались (то есть вне поля зрения датчика), так что они не влияют на прямоугольник обрезки. Таким образом, например, если два пользователя первоначально отслеживались, и один перемещается вне поля зрения, в соответствии со способами 800 и 900, то прямоугольники обрезки, вычисленные далее, будут основаны только на характеристиках оставшегося пользователя и, таким образом, будут автоматически «приближать» оставшегося пользователя (причем модель упругой пружины обеспечивает плавный переход для этого приближения).

[75] В осуществлениях, видеоконтроллер 304 может давать обратную связь средству 306 управления ресурсами, так чтобы, если средство управления ресурсами считает, что ширины полосы пропускания достаточно для видеосигнала 1920x1080, но есть, по меньшей мере, одно тело, которое надо отследить, то разрешение передачи видеосигнала переключают на 1280x720, чтобы позволить динамическое отслеживание (так, для разрешения переданного видеосигнала, равного тому, при котором оно захвачено, нет «места для маневра», так как для каждого кадра прямоугольник обрезки эффективно вмещает весь тот кадр - это верно, например, когда повышение масштабирования не

применяется).

[76] Как указано, устройство 310 обработки датчика поставляет не только информацию, идентифицирующую каждую обнаруженную скелетную точку, но также идентифицирующую, какому из одного или более пользователей в поле зрения датчика 226 какие скелетные точки соответствуют. Таким образом, описанный способ 800 может быть реализован с любым числом пользователей (до тех пор, пока устройство 226 обнаружения датчика способно различать их) и, для каждой характеристики в выбранном участке, будет отслеживать эти характеристики для каждого участка. Таким образом, способ 800 настраивается автоматически по мере того, как множество пользователей входят и выходят из кадра, причем граничные прямоугольники автоматически переходят плавным образом (вследствие модели упругой пружины), чтобы расположить новых пользователей, когда они входят в кадр (путем эффективного удаления для включения любых частей тела этих пользователей для выбранного участка в видеосигнале), и чтобы регулировать, когда пользователи выходят из кадра (путем эффективного приближения для исключения любых участков, ранее занятых частями тела этого пользователя для выбранного участка так, чтобы только части тела для выбранного участка оставшихся пользователей остались в видеосигнале).

[77] Способы 800, 900, реализованные блоком 302 управления, можно реализовать во время конференц-вызова, управляемого по сети 106, с использованием, например, пользовательских устройств 104, 110 и 114 (между пользователями 102a, 102b, 108 и 112), причем индивидуальные граничные прямоугольники определяются пользователем 104 (первое устройство) ближнего конца линии связи для каждого устройства 110 (второе устройство) и 114 (третье устройство) дальнего конца линии связи.

[78] То есть, дополнительно к вышеуказанному, средство управления ресурсами может принимать дополнительную информацию, по меньшей мере, об одном из: канал связи между пользовательским устройством 104 и третьим пользовательским устройством 114, и ресурсы дополнительного пользовательского устройства (дополнительно к приему информации, описанной выше, как подобная информация для второго устройства 110). Средство управления ресурсами затем выбирает дополнительные характеристики, выбранные из множества визуальных пользовательских характеристик (например, частей тела), на основании принятой дополнительной информации для управления видеосигналом, который подлежит передаче третьему пользовательскому устройству 114 (дополнительно к выбору вышеупомянутых характеристик для управления видеосигналом, подлежащим передаче второму пользовательскому устройству 110).

[79] Видеоконтроллер 304 затем управляет видеосигналом, подлежащим передаче третьему пользовательскому устройству 114, на основании обнаружения выбранных дополнительных характеристик, выбранных для третьего пользовательского устройства 114 (чтобы отслеживать выбранные дополнительные характеристики в видеосигнале третьего устройства), одновременно управляя видеосигналом, подлежащим передаче второму пользовательскому устройству 110, на основании обнаружения характеристик, выбранных для второго устройства 110.

[80] Дополнительный выбор характеристик для третьего пользовательского устройства может не зависеть и отличаться от выбора характеристик для второго пользовательского устройства. Таким образом, дополнительный видеосигнал, передаваемый второму пользовательскому устройству, может отличаться от видеосигнала, передаваемого второму пользовательскому устройству, причем дополнительный видеосигнал, передаваемый третьему пользовательскому устройству,

отслеживает больше или меньше пользовательских характеристик, чем видеосигнал, передаваемый второму пользовательскому устройству.

5 [81] Выбор характеристик для видеосигнала второго пользовательского устройства не зависит от выбора характеристик для видеосигнала третьего пользователя. В то время как условие (как, например, первое пользовательское устройство 104 соединено с сетью 106 по медленному соединению) может привести к подобной обрезке для обоих видеосигналов, другие условия (как, например, одно из второго и третьего устройств соединено с сетью 106 по медленному соединению, или одно из этих устройств имеет ограниченные ресурсы) могут привести к другой обрезке.

10 [82] Например, третье пользовательское устройство 114 может иметь маленький экран (например, экран смартфона) и/или быть соединено с сетью 106 по медленному соединению; напротив, второе пользовательское устройство 110 может иметь большой экран (например, быть соединенным с ТВ-экраном) и/или быть соединено с сетью 106 по быстрому соединению. В таком случае, видеосигнал, переданный второму
15 пользовательскому устройству, может подвергнуться обрезке «участок 4» (см. таблицу 2), так что пользователь 108 принимает видеосигнал, изображающий пользователей 102a и 102b с головы до пят. Напротив, видеосигнал, переданный пользовательскому устройству 114, может подвергнуться обрезке «участок 1» (см. таблицу 2), так что пользователь 112 принимает видеосигнал, изображающий только соответственные
20 головы и плечи пользователей 102a и 102b.

[83] В то время как выше указано, что видеосигналом, подлежащим передаче по сети, управляют путем обработки видеосигнала, альтернативно или дополнительно видеосигналом может управлять видеоконтроллер 304 (или блок 302 управления) путем манипуляции камерой на основании обнаружения выбранных признаков, например,
25 манипуляции, содержащей механизм камеры для выполнения, по меньшей мере, одного из: операции наклона, операции приближения и операции поворота. Например, для конференц-звонка, оптическое приближение и цифровое приближение (обрезка) можно использовать совместно, например, с механическим приближением, используемым для захвата видеосигнала, показывающим наибольшее выбранное число пользовательских
30 характеристик (например, механическое приближение можно использовать для выбора участка, который будет показан второму пользовательскому устройству 110 в вышеуказанном примере), причем обрезку этого видеосигнала используют для управления видеосигналом для пользователей, которым отправят видеосигнал с меньшим числом пользовательских характеристик (например, третье устройство 114
35 в вышеуказанном примере).

[84] Дополнительно, в то время как выше использована модель упругой пружины, в качестве альтернативы любую динамическую модель (например, на основании одного или более дифференциальных уравнений во времени) можно использовать для формирования данных о переходе.

40 [85] Дополнительно, в то время как выше указанные выбранные визуальные характеристики (ноги, руки и так далее) обнаружены на основании данных датчика, поставленного устройством обнаружения глубины, которое проецирует диаграмму направленности невидимого излучения от датчика, выполненного с возможностью обнаруживать эту диаграмму, предусмотрены альтернативные обнаружения. Например, обнаружение глубины может быть временем прохождения на основании обнаружения, в котором моменты распространения излучения используют для измерения глубины. Альтернативно, решетка камер, имеющих различные наклоны (как, например, пленоптическая камера, используемая совместно с алгоритмом распознавания 3D

изображения), может быть использована для построения 3D изображения (из множества двумерных изображений), причем визуальные характеристики обнаруживают из 3D изображения.

[86] В общем, любые функции, описанные в настоящей заявке (например, функциональные модули, изображенные на Фиг. 3, и функциональные этапы, изображенные на Фиг. 8 и 9), можно реализовать с использованием программного обеспечения, программно-аппаратных средств, технических средств (например, фиксированная логическая схема) или комбинации этих реализаций. Модули (система 300 обработки видеосигнала, блок 302 управления, устройство 304 обработки видеосигнала, средство 306 управления ресурсами, устройство 308 кодирования и так далее), изображенные отдельно на Фиг. 3, и этапы, изображенные отдельно на Фиг. 8 и 9, можно реализовать или нет в качестве отдельных модулей или этапов. Термины «модуль», «функциональность», «компонент» и «логика», используемые в настоящей заявке, в общем представляют собой программное обеспечение, программно-аппаратные средства, технические средства или их комбинацию. В случае реализации в программном обеспечении, модуль, функциональность или логика представляют собой программный код, который выполняет конкретные задачи при выполнении в устройстве (например, CPU) обработки. Программный код может храниться в одном или более считываемых компьютером запоминающих устройствах. Признаки способов, описанных в настоящей заявке, не зависят от платформы, что означает, что способы можно реализовать на разных коммерческих вычислительных платформах, имеющих разные устройства обработки. Например, пользовательские устройства также могут включать в себя сущность (например, программное обеспечение), которая побуждает технические средства пользовательских устройств выполнять операции, например, функциональные блоки устройств обработки, и так далее. Например, пользовательские устройства могут включать в себя считываемый компьютером носитель, который может быть выполнен с возможностью поддерживать команды, которые побуждают пользовательские устройства и, более конкретно, операционную систему и связанные технические средства пользовательских устройств, выполнять операции. Например, некоторые или все модули на Фиг. 3 можно реализовать в программном обеспечении клиентского приложения, выполняемого одним или несколькими устройствами обработки. Таким образом, команды функционируют для конфигурации операционной системы и связанных технических средств для выполнения операций и, таким образом, приводят к преобразованию операционной системы и связанных технических средств для выполнения функций. Команды могут быть обеспечены считываемым компьютером носителем данных пользовательским устройствам по разным конфигурациям.

[87] Одной такой конфигурацией считываемого компьютером носителя является носитель, несущий сигнал, и, таким образом, выполненный с возможностью передавать команды (например, несущая волна) вычислительному устройству, например, по сети. Считываемый компьютером носитель также может быть выполнен в качестве считываемого компьютером носителя данных и, таким образом, не является носителем, несущим сигнал. Примеры считываемого компьютером носителя данных включают в себя оперативное запоминающее устройство (RAM), постоянное запоминающее устройство (ROM), оптический диск, устройство флэш-памяти, запоминающее устройство на жестком диске и другие запоминающие устройства, которые могут использовать магнитные, оптические и другие способы хранения команд и других данных.

[88] Хотя изобретение описано на языке, специфичном для структурных признаков и/или методологических действий, необходимо понимать, что изобретение, определенное

в приложенной формуле изобретения, не обязательно ограничено конкретными признаками или действиями, описанными выше. Напротив, конкретные признаки или действия, описанные выше, раскрыты как примеры форм осуществления формулы изобретения.

5

(57) Формула изобретения

1. Пользовательское устройство (104), содержащее:

кодер (308), выполненный с возможностью приема видео, содержащего одного или более пользователей (102a, 102b), от устройства захвата изображения и кодирования этого видео для передачи по меньшей мере другому пользовательскому устройству (108) по каналу связи сети (106), причем канал связи обеспечен между пользовательским устройством (104) и другим пользовательским устройством (108);

сетевой интерфейс (220), выполненный с возможностью передавать кодированное видео по каналу связи упомянутому другому пользовательскому устройству (108);

средство (306) управления ресурсами, выполненное с возможностью принимать информацию о канале связи между пользовательским устройством (104) и упомянутым другим пользовательским устройством (108), при этом принимаемая информация содержит информацию о качестве канала, и, на основании принятой информации о качестве канала, определять разрешение видео, подлежащее использованию для видео, основываясь на принятой информации о качестве канала, при этом средство (306) управления ресурсами выполнено с возможностью определения более высокого разрешения видео, если принятая информация о качестве канала указывает лучшее качество канала, и более низкого разрешения видео, если принятая информация о качестве канала указывает худшее качество канала, при этом кодер (308) выполнен с возможностью кодирования видео с определенным разрешением видео; и

видеоконтроллер (304), выполненный с возможностью выбора для пользователя в видео одной или более частей тела пользователя из множества частей тела пользователя на основе разрешения видео, определенного средством (306) управления ресурсами, и управления видео на основе обнаружения выбранных частей тела пользователя, чтобы отслеживать выбранные части тела пользователя для пользователя посредством: обработки видео перед его кодированием, чтобы обрезать упомянутое видео на основе обнаруженных частей тела пользователя, или управления устройством захвата изображения, чтобы выполнить операцию масштабирования на основе обнаруженных частей тела, при этом видеоконтроллер (304) выполнен с возможностью выбора и отслеживания большего количества частей тела пользователя для пользователя при более высоком разрешении видео и выбора и отслеживания меньшего количества частей тела пользователя для пользователя при более низком разрешении видео.

2. Пользовательское устройство (104) по п.1, в котором средство (306) управления ресурсами выполнено с возможностью принимать информацию об одном или более ресурсах пользовательского устройства (104) и/или другого пользовательского устройства (108).

3. Пользовательское устройство (104) по п.1, в котором множество частей тела содержит: голову (750), и/или плечи (752), и/или среднюю часть (756) спины, и/или нижнюю часть (758) спины, и/или всю спину (760), и/или локти (754a, 754b), и/или бедра (762), и/или ноги (764a, 764b), и/или ступни (766a, 766b).

4. Пользовательское устройство (104) по п.3, в котором выбор упомянутых частей тела пользователя является выбором одного участка тела из четырех участков тела, причем первый из упомянутых четырех участков тела содержит голову (750) и плечи

(752); второй из упомянутых четырех участков тела содержит голову (750), плечи (752), среднюю часть (756) спины и локти (754a, 754b); третий из упомянутых четырех участков тела содержит голову (750), среднюю часть (756) спины, локти (754a, 754b), нижнюю часть (758) спины и бедра (762); четвертый из упомянутых четырех участков тела
5 содержит голову (750), плечи (752), всю спину (760), локти (754a, 754b), бедра (762), ноги (764a, 764b) и ступни (766a, 766b).

5. Пользовательское устройство (104) по п.1 или 2, в котором видео содержит множество пользователей (102a, 102b), причем видеоконтроллер (304) дополнительно выполнен с возможностью управлять видео на основании обнаружения выбранных
10 частей тела пользователя для каждого из пользователей (102a, 102b) для отслеживания выбранных частей тела пользователя для каждого из пользователей (102a, 102b).

6. Пользовательское устройство (104) по п.1 или 2, в котором сетевой интерфейс (220) дополнительно выполнен с возможностью передавать дополнительное видео, содержащее одного или более пользователей (102a, 102b), принятое от устройства
15 захвата изображения, дополнительному пользовательскому устройству (114) по упомянутой сети;

причем видеоконтроллер (304) дополнительно выполнен с возможностью выполнять дополнительный выбор одной или более частей тела пользователя из множества частей тела пользователя и управлять дополнительным видео на основании обнаружения
20 выбранных дополнительных частей тела пользователя для отслеживания выбранных дополнительных частей тела пользователя; и

при этом средство (306) управления ресурсами дополнительно выполнено с возможностью принимать дополнительную информацию о канале связи между пользовательским устройством (104) и дополнительным пользовательским устройством
25 (114) и/или одним или несколькими ресурсами дополнительного пользовательского устройства (114) и, на основании принятой дополнительной информации, управлять упомянутым выбором с помощью видеоконтроллера (304).

7. Пользовательское устройство (104) по п.1 или 2, в котором обнаружение выбранных частей тела пользователя содержит прием, для по меньшей мере одного из пользователей
30 (102a, 102b), множества обнаруженных скелетных точек этого пользователя, причем части тела пользователя определяются исходя из этих скелетных точек.

8. Пользовательское устройство (104) по п.1 или 2, в котором упомянутое управление содержит формирование данных о границе на основании обнаружения выбранных частей тела пользователя, и по меньшей мере одно из: обработки видеоданных, принятых
35 от устройства захвата изображения, на основании сформированных данных о границе, и манипуляции механизмами устройства захвата изображения на основании сформированных данных о границе.

9. Пользовательское устройство (104) по п.1, в котором видеоконтроллер (304) дополнительно выполнен с возможностью:

40 формировать первую и вторую группы данных о границе, задающих границу соответствующего прямоугольника, отслеживающего части тела пользователя на основании обнаружения выбранных частей тела пользователя в первый и второй моменты времени соответственно;

формировать переход, представляющий переход от границы первого прямоугольника
45 к границе второго прямоугольника с использованием динамической модели; и
управлять видео на основании данных о переходе для отслеживания выбранных частей тела пользователя посредством обрезки видео по границе прямоугольника.

10. Считываемый компьютером носитель, хранящий код, который, при выполнении

на одном или более устройствах обработки пользовательского устройства (104), выполнен с возможностью реализовывать способ передачи видео по сети (106), содержащий:

5 прием видео, содержащего одного или более пользователей (102a, 102b), от устройства захвата изображения;

прием информации о канале связи между пользовательским устройством (104) и другим пользовательским устройством (110);

10 определение разрешения видео, подлежащего использованию для видео, основываясь на принятой информации о качестве канала, при этом более высокое разрешение видео определяется, если принятая информация о качестве канала указывает лучшее качество канала, и более низкое разрешение видео определяется, если принятая информация о качестве канала указывает худшее качество канала, при этом видео кодируется с определенным разрешением видео;

15 выбор частей тела пользователя для пользователя в видео из множества частей тела пользователя на основе разрешения видео, определенного средством (306) управления ресурсами, при этом большее количество частей тела пользователя выбирается и отслеживается для пользователя при более высоком разрешении видео и меньшее количество частей тела пользователя выбирается и отслеживается для пользователя при более низком разрешении видео;

20 управление видео на основе обнаружения выбранных частей тела пользователя, чтобы отслеживать выбранные части тела пользователя для пользователя посредством: обработки видео перед его кодированием, чтобы обрезать упомянутое видео на основе обнаруженных частей тела пользователя, или управления устройством захвата изображения, чтобы выполнить операцию масштабирования на основе обнаруженных

25 частей тела;

кодирование видео с определенным разрешением для передачи через канал связи сети, причем канал связи обеспечен между пользовательским устройством (104) и другим пользовательским устройством (108); и

30 передачи кодированного видео по каналу связи другому пользовательскому устройству (108).

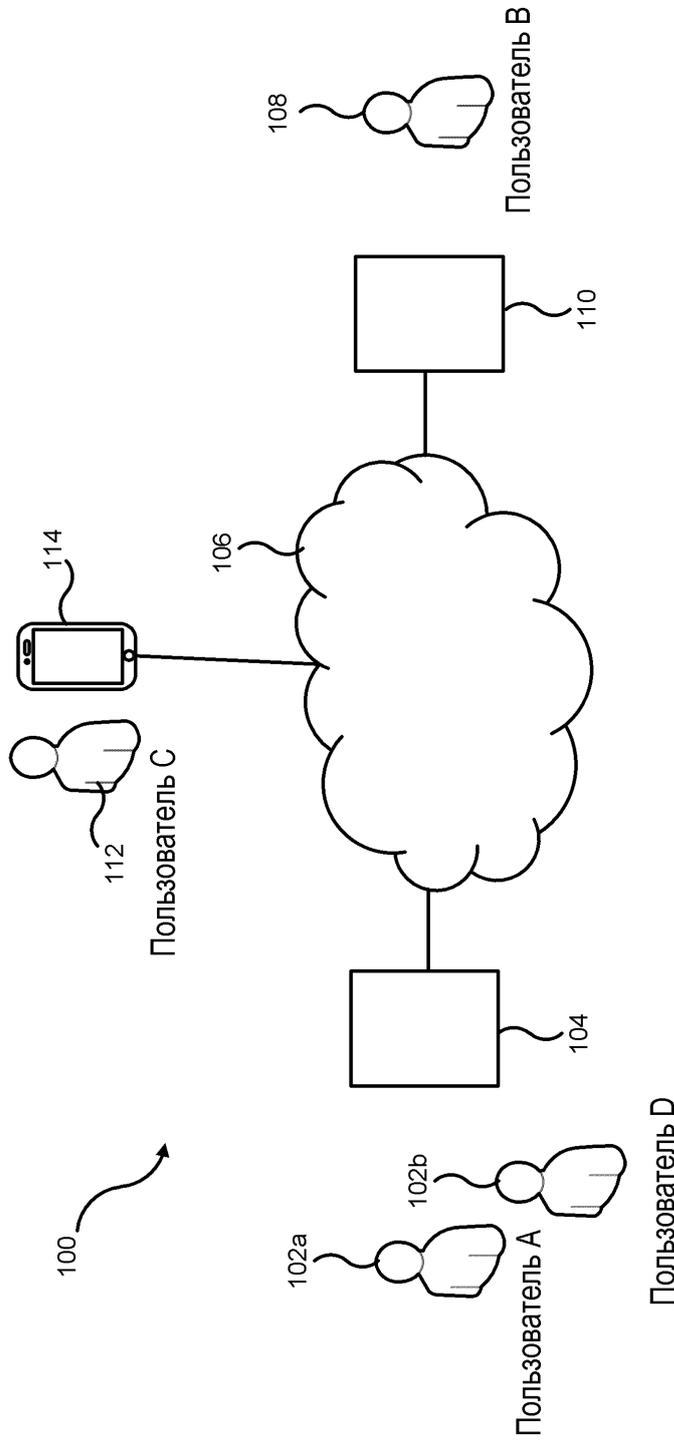
11. Считываемый компьютером носитель, хранящий код, по п.10, при этом средство (306) управления ресурсами выполнено с возможностью принимать информацию об одном или более ресурсах пользовательского устройства (104) и/или другого пользовательского устройства (110).

35

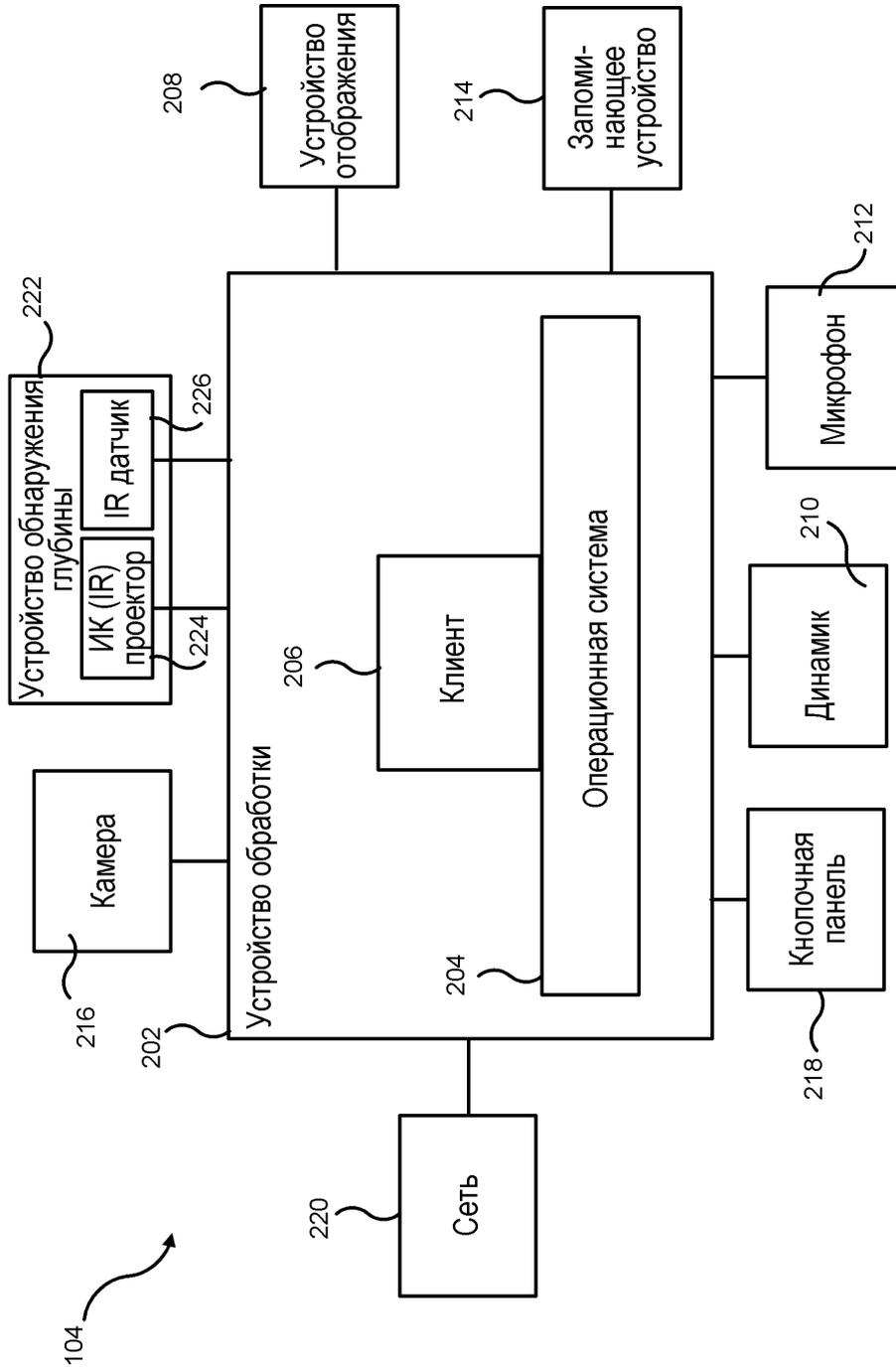
40

45

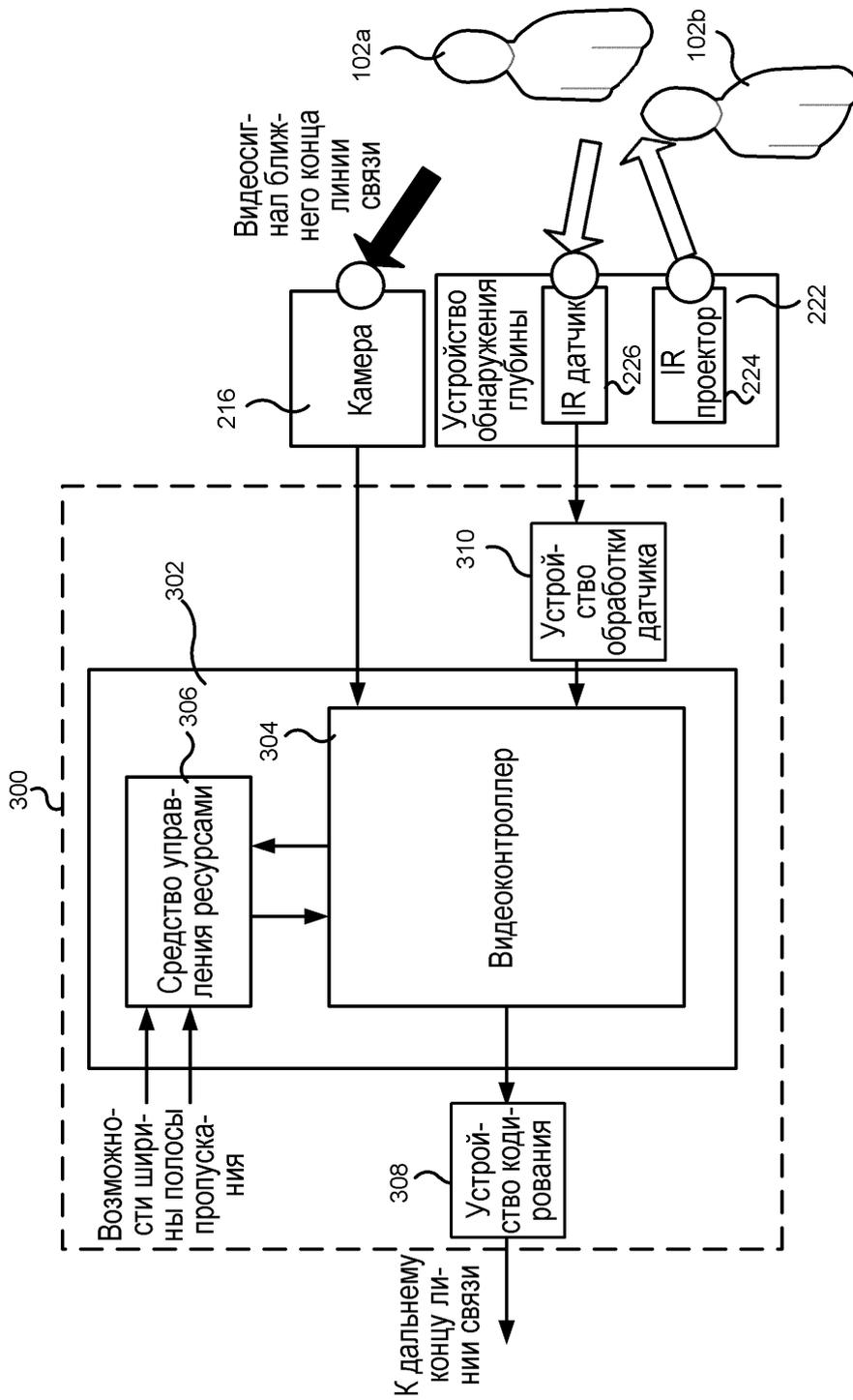
1/9



ФИГ. 1

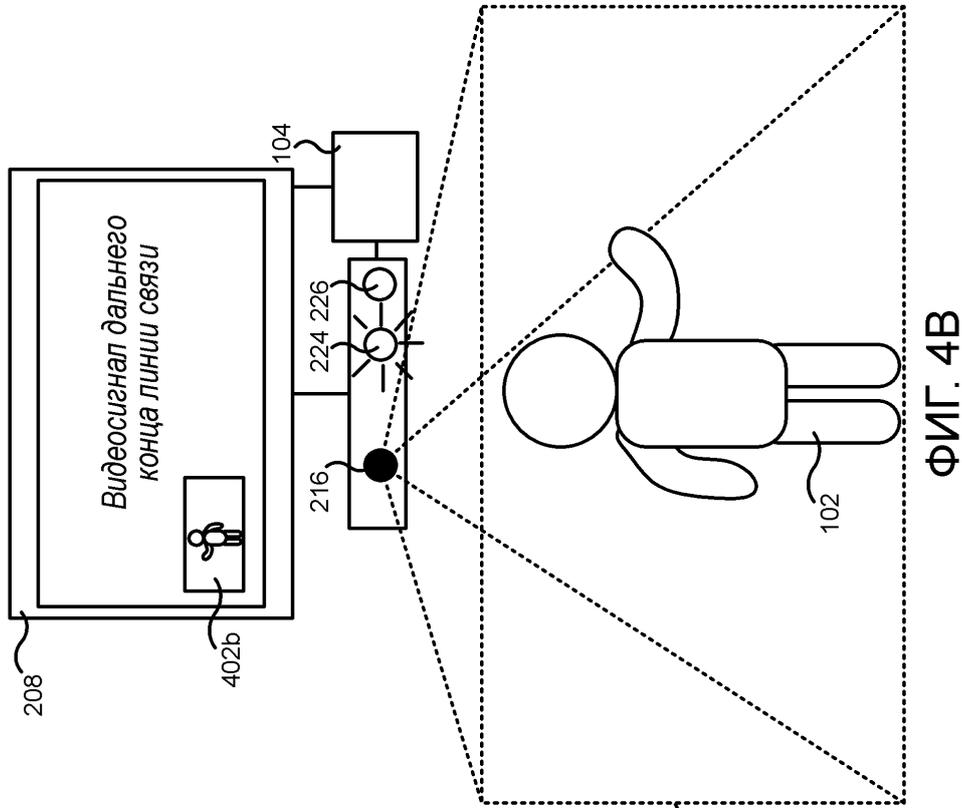


ФИГ. 2

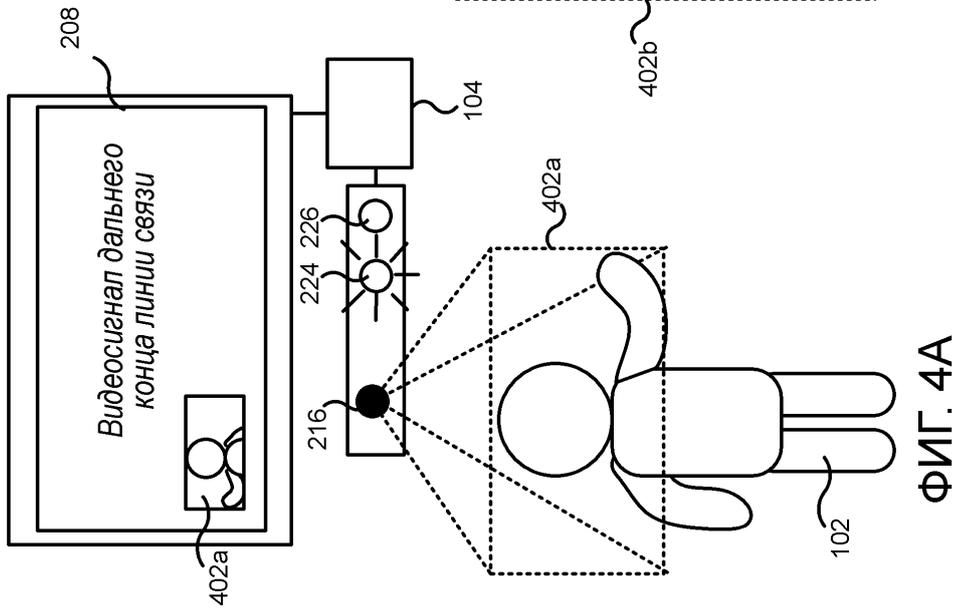


ФИГ. 3

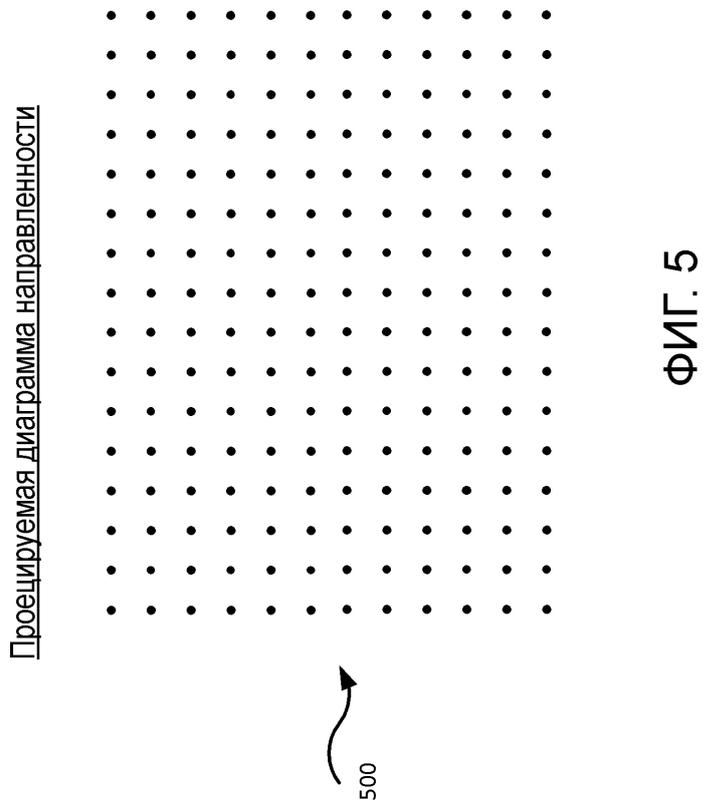
4/9



ФИГ. 4В



ФИГ. 4А



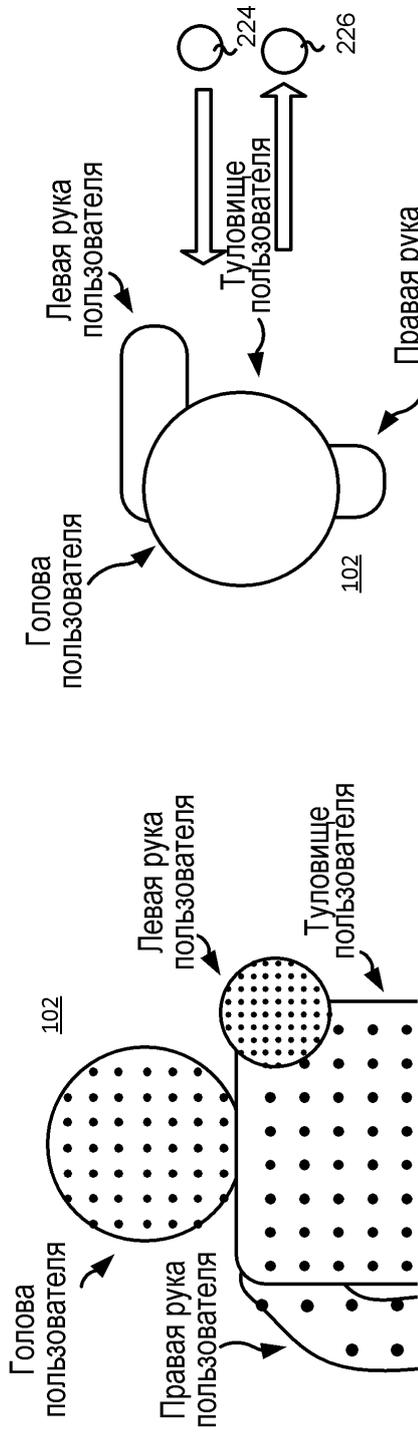


FIG. 6B
Горизонтальный вид

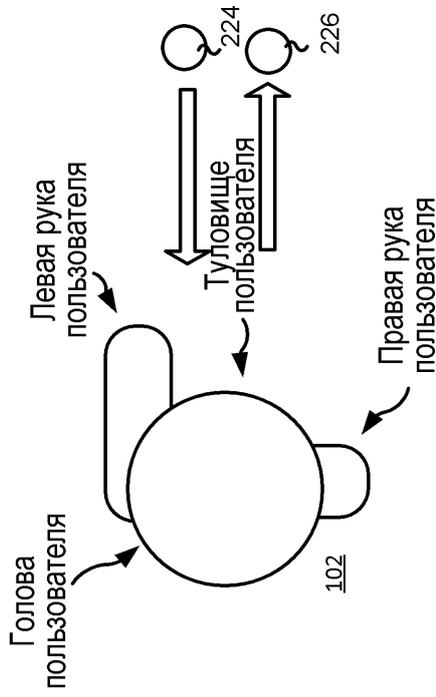
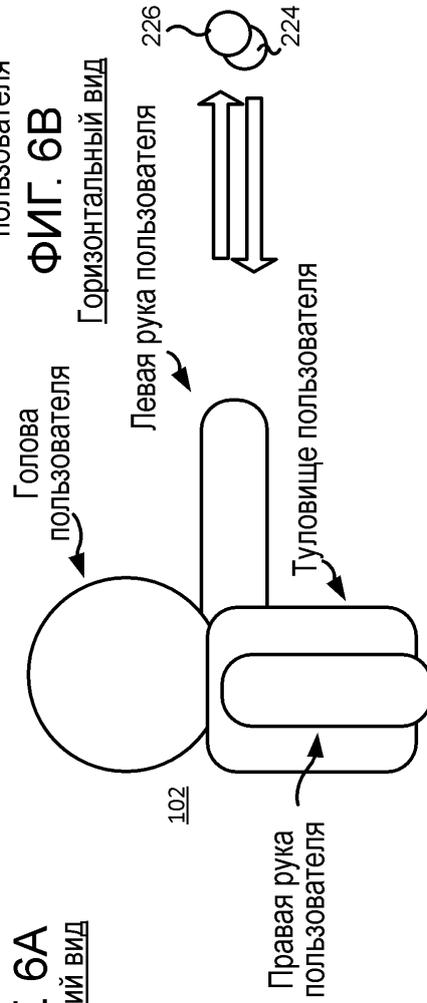
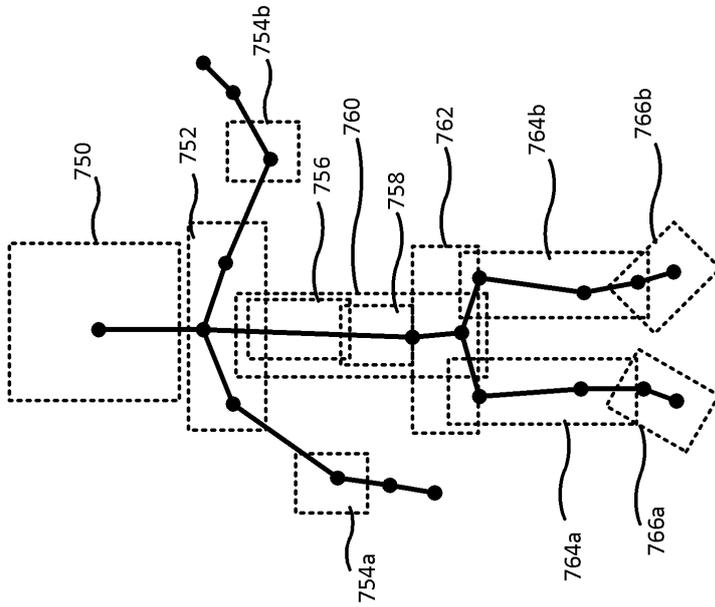
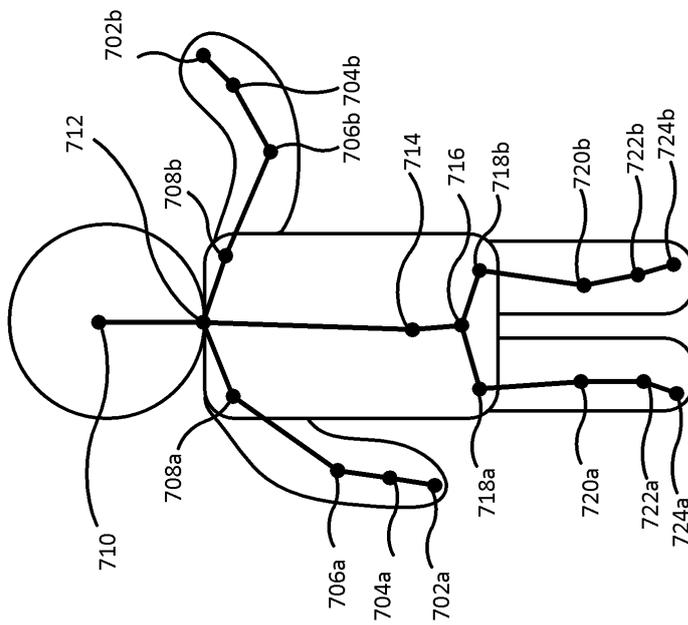


FIG. 6C
Боковой вид

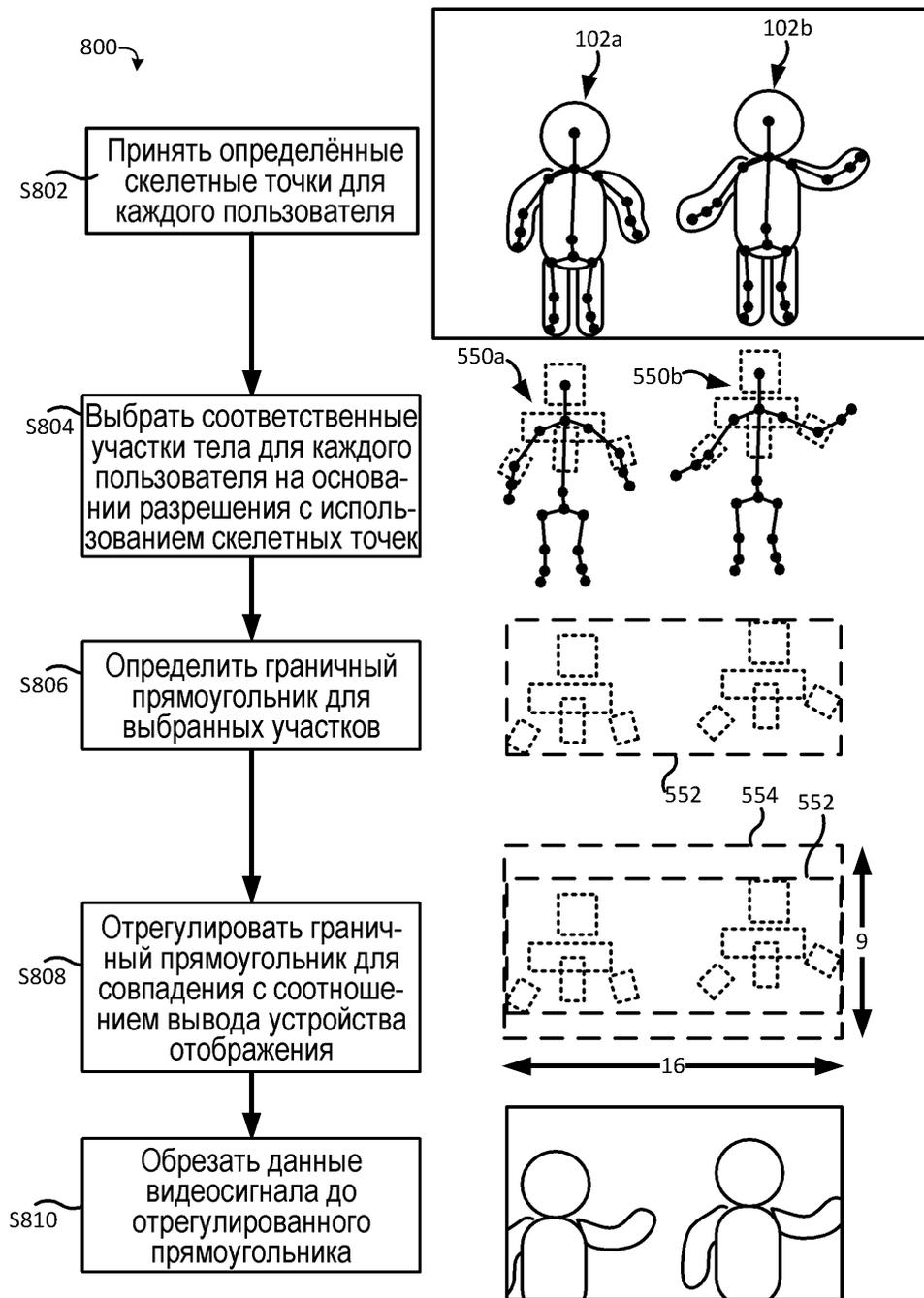




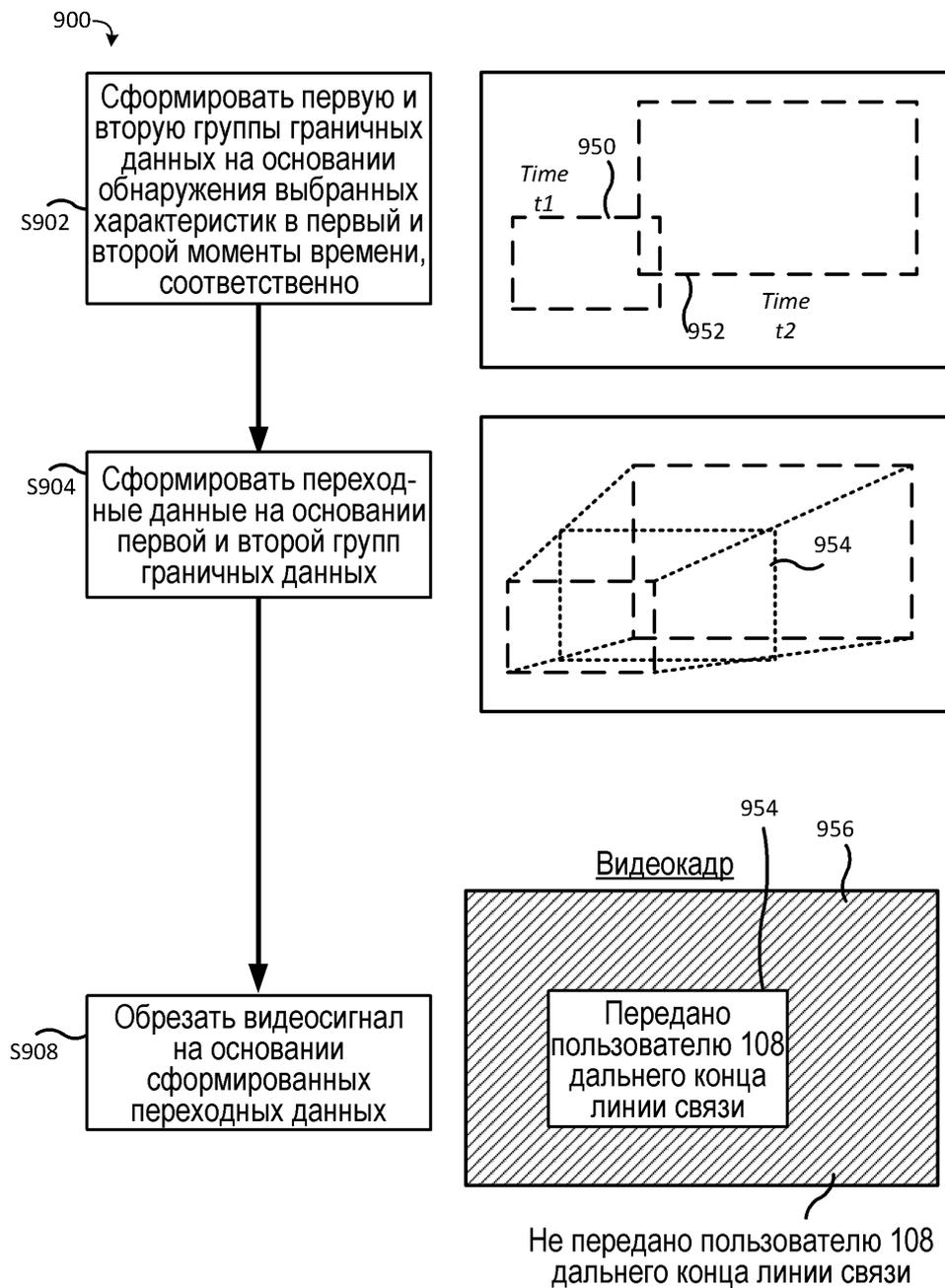
ФИГ. 7В



ФИГ. 7А



ФИГ. 8



ФИГ. 9