

(19)



**Евразийское  
патентное  
ведомство**

(21) **202291893** (13) **A1**

(12) **ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ЕВРАЗИЙСКОЙ ЗАЯВКЕ**

(43) Дата публикации заявки  
2023.12.29

(51) Int. Cl. *H04N 21/4788* (2023.01)  
*H04L 65/403* (2023.01)

(22) Дата подачи заявки  
2022.07.06

---

(54) **СИСТЕМА И СПОСОБ ДЛЯ ВИДЕОКОНФЕРЕНЦСВЯЗИ**

---

(31) 2022114549

(32) 2022.05.30

(33) RU

(71) Заявитель:  
**ПУБЛИЧНОЕ АКЦИОНЕРНОЕ  
ОБЩЕСТВО "СБЕРБАНК  
РОССИИ" (ПАО СБЕРБАНК) (RU)**

(72) Изобретатель:

**Егоров Иван Евгеньевич, Насыров  
Михаил Миннимухаметович,  
Большаков Александр Николаевич,  
Сергеев Даниэл Игоревич, Фандюшин  
Михаил Викторович, Новиков  
Герман Эдуардович, Балиев Дмитрий  
Леонидович (RU)**

(74) Представитель:  
**Герасин Б.В. (RU)**

---

(57) Изобретение относится к области телекоммуникаций и компьютерных технологий, в частности, к телекоммуникационной технологии взаимодействия двух и более удалённых пользователей. Техническим результатом является обеспечение функциональной возможности параллельного вывода потока данных по меньшей мере одному клиентскому устройству и распознавания этого потока данных для формирования сообщения для вывода в единое информационное пространство. Заявленный технический результат достигается за счет реализации системы для видеоконференцсвязи, содержащей по меньшей мере два клиентских устройства, выполненных с возможностью подключаться к модулю обмена сообщениями; и сервер видеоконференцсвязи, содержащий по меньшей мере модуль обмена сообщениями, модуль пересылки потоков и модуль распознавания потоков, причем модуль обмена сообщениями выполнен с возможностью регистрировать подключения клиентских устройств к конференции, отправлять в единое информационное пространство по меньшей мере одно сообщение, полученное по меньшей мере от одного клиентского устройства, модуль пересылки потоков выполнен с возможностью устанавливать защищенное соединение с клиентскими устройствами, принимать поток данных по меньшей мере от одного клиентского устройства и передавать поток данных в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков, принимать от модуля распознавания потоков по меньшей мере одно сгенерированное сообщение и отправлять его в единое информационное пространство, модуль распознавания потоков выполнен с возможностью обрабатывать поток данных и генерировать по меньшей мере одно сообщение на основе обработанного потока данных, причем обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных модулем пересылки потоков по меньшей мере одному другому клиентскому устройству.

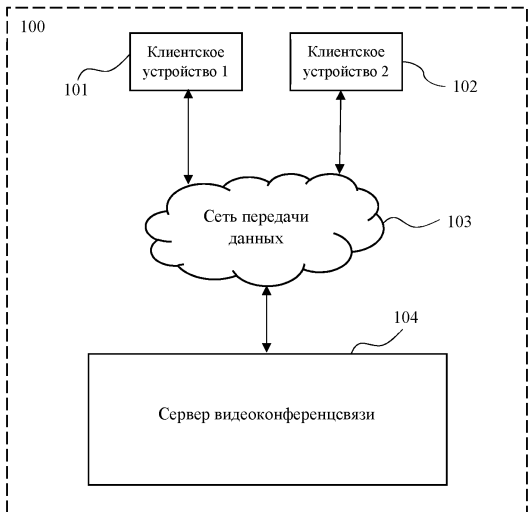
---

**A1**

**202291893**

**202291893**

**A1**



## СИСТЕМА И СПОСОБ ДЛЯ ВИДЕОКОНФЕРЕНЦСВЯЗИ

### ОБЛАСТЬ ТЕХНИКИ

[0001] Заявленное решение относится к области телекоммуникаций и компьютерных технологий, в частности, к телекоммуникационной технологии взаимодействия двух и более удалённых пользователей, при которой между ними возможен обмен аудио- и видеoinформацией в реальном времени.

### УРОВЕНЬ ТЕХНИКИ

[0002] Из уровня техники известно множество различных систем видеоконференцсвязи (ВКС), как от российских, так и от зарубежных разработчиков.

[0003] Наиболее известными из них являются:

- Zoom
- Cisco Webex
- MS Teams
- Google Meet
- TrueConf
- IVA
- Videomost.

[0004] К основным недостаткам большинства решений относится отсутствие возможности одновременной работы с речью и текстовой коммуникацией во время проведения видеоконференции, в частности, отсутствие безбарьерной среды для одновременного общения во время видеоконференции с участниками с нарушением слуха.

[0005] В патентной заявке US 20120306993 A1 «System and Method for Teleconferencing» (правообладатель: Visionary Mobile Corporation, опубликована 06.12.2012) раскрывается система проведения видеоконференций по сети для множества клиентов с использованием сервера в качестве посредника. На сервер поступает видеопоток в реальном времени от каждого клиента. Этот видеопоток в реальном времени может быть высокочастотным потоком, представляющим видео с камеры, связанной с конечным пользователем. Сервер предоставляет любому клиенту видеопоток в реальном времени, выбранный данным клиентом из числа видеопотоков, полученных на сервере.

[0006] Недостатком данного решения является отсутствие возможности распознавания видеопотока для формирования сообщения, предоставляемого пользователю. Это в свою очередь обеспечивает сложность с погружением в контекст обсуждения для участников

видеоконференции, которые вынуждены были на какое-то время отвлечься в процессе обсуждения.

[0007] Кроме того, в известном уровне техники существует сложность с погружением в контекст обсуждения для участников видеоконференции, которые вынуждены были на какое-то время отвлечься в процессе обсуждения.

## **СУЩНОСТЬ ИЗОБРЕТЕНИЯ**

[0008] Заявленное изобретение позволяет решить техническую проблему, заключающуюся в обеспечении возможности извлечения информации из потока данных с её последующим преобразованием в сообщения и их интеграцию в единое информационное пространство.

[0009] Техническим результатом является обеспечение функциональной возможности параллельного вывода потока данных по меньшей мере одному клиентскому устройству и распознавания этого потока данных для формирования сообщения для вывода в единое информационное пространство.

[0010] Заявленный технический результат достигается за счет реализации системы для видеоконференцсвязи, содержащей:

- по меньшей мере два клиентских устройства, выполненных с возможностью подключаться к модулю обмена сообщениями; и
- сервер видеоконференцсвязи, содержащий по меньшей мере модуль обмена сообщениями, модуль пересылки потоков и модуль распознавания потоков, причем модуль обмена сообщениями выполнен с возможностью
  - регистрировать подключения клиентских устройств к конференции,
  - отправлять в единое информационное пространство по меньшей мере одно сообщение, полученное от по меньшей мере одного клиентского устройства,модуль пересылки потоков выполнен с возможностью
  - устанавливать защищенное соединение с клиентскими устройствами,
  - принимать поток данных от по меньшей мере одного клиентского устройства,
  - передавать поток данных в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков,
  - принимать от модуля распознавания потоков по меньшей мере одно сгенерированное сообщение и отправлять его в единое информационное пространство,

модуль распознавания потоков выполнен с возможностью обрабатывать поток данных и генерировать по меньшей мере одно сообщение на основе обработанного потока

данных, причем обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных модулем пересылки потоков по меньшей мере одному другому клиентскому устройству.

[0011] В одном из частных примеров реализации принятый от по меньшей мере одного клиентского устройства поток данных содержит по меньшей мере аудио и/или видео поток.

[0012] В другом частном примере реализации обработка аудио потока включает преобразование речи в текст (Automatic Speech Recognition).

[0013] В другом частном примере реализации обработка видео потока включает преобразование жестов в текстовую и/или графическую информацию (Gesture Recognition).

[0014] В другом частном примере реализации обработка видео потока включает распознавание языка жестов (Sign Language Recognition).

[0015] В другом частном примере реализации модуль распознавания потоков дополнительно выполнен с возможностью определять тип потока данных и осуществлять соответствующую обработку потока данных в зависимости от типа потока данных.

[0016] В другом частном примере реализации сервер видеоконференцсвязи дополнительно содержит модуль фильтрации аудио потока, выполненный с возможностью подавления шума.

[0017] В другом частном примере реализации сообщения, сгенерированные на основе обработанного потока данных, содержат по меньшей мере текстовую и/или графическую информацию.

[0018] В другом частном примере реализации сервер видеоконференцсвязи дополнительно содержит модуль модификации видео потока, выполненный с возможностью

- замены фона,
- применения фильтров, улучшающих визуальные характеристики изображения пользователя,
- применения фильтров дополненной реальности (AR-фильтров).

[0019] В другом частном примере реализации модуль пересылки потоков дополнительно выполнен с возможностью принимать от по меньшей мере одного клиентского устройства команду на распознавание аудио и/или видео потока; и передавать аудио и/или видео поток модулю распознавания потоков после получения упомянутой команды на распознавание.

[0020] В другом частном примере реализации модуль распознавания потоков дополнительно выполнен с возможностью

- осуществлять разбиение аудио потока на фрагменты в процессе преобразования речи в текст,
- обрабатывать фрагменты, полученные в результате разбиения аудио потока,
- генерировать по меньшей мере одно сообщение на основе обработанных фрагментов;

сервер видеоконференцсвязи дополнительно содержит модуль согласования, выполненный с возможностью

- определять в режиме реального времени, осуществляет ли модуль распознавания потоков обработку аудио потока,
- устанавливать значение флага, идентифицирующего состояние текущего голосового ввода пользователя клиентского устройства,
- получать от модуля обмена сообщениями запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства,
- отправлять модулю обмена сообщениями разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства на основании значения упомянутого флага, при этом

- если значение флага идентифицирует отсутствие текущего голосового ввода пользователя клиентского устройства, модуль согласования отправляет модулю обмена сообщениями разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства;

- если значение флага идентифицирует наличие текущего голосового ввода пользователя клиентского устройства, модуль согласования выполнен с возможностью

- отправлять модулю пересылки потоков запрос на отправку в единое информационное пространство сгенерированного сообщения, содержащего уже обработанные фрагменты,

- после отправки модулем пересылки потоков сгенерированного сообщения, содержащего уже обработанные фрагменты, в единое информационное пространство, отправлять модулю обмена сообщениями разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства;

модуль обмена сообщениями дополнительно выполнен с возможностью

- отправлять модулю согласования запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства,
- отправлять в единое информационное пространство сообщения от другого клиентского устройства в ответ на прием разрешения отправки от модуля согласования.

[0021] В другом частном примере реализации сервер видеоконференцсвязи дополнительно содержит модуль управления единым информационным пространством, выполненный с возможностью определять, что по меньшей мере одно сгенерированное сообщение получено от модуля пересылки потоков; и отображать при выводе по меньшей мере одного сгенерированного сообщения метку, идентифицирующую формирование сгенерированного сообщения из аудио потока.

[0022] В другом частном примере реализации модуль управления единым информационным пространством дополнительно выполнен с возможностью

- сохранять сообщения, размещенные в упомянутом пространстве,
- обеспечивать доступ подключенным клиентским устройствам к сохраненным сообщениям.

[0023] В другом частном примере реализации модуль управления единым информационным пространством дополнительно выполнен с возможностью осуществлять

- поиск по содержимому сообщений, размещенных в упомянутом пространстве,
- скачивание содержимого сообщений, размещенных в упомянутом пространстве,
- управление отображением сообщений, размещенных в упомянутом пространстве.

[0024] В другом частном примере реализации сервер видеоконференцсвязи дополнительно содержит модуль перевода, выполненный с возможностью перевода по меньшей мере одного сообщения, полученного от по меньшей мере одного клиентского устройства и/или сообщения, сгенерированного на основе обработанного потока данных.

[0025] Заявленный технический результат достигается также за счет выполнения способа осуществления видеоконференцсвязи, реализуемого с помощью сервера видеоконференцсвязи и содержащего этапы, на которых:

- подключают по меньшей мере два клиентских устройства к серверу видеоконференцсвязи,
- регистрируют подключения клиентских устройств к конференции,
- устанавливают защищенное соединение с клиентскими устройствами,
- принимают поток данных от по меньшей мере одного клиентского устройства,

- передают поток данных в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков,
- в модуле распознавания потоков обрабатывают поток данных и генерируют по меньшей мере одно сообщение на основе обработанного потока данных, причем обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных по меньшей мере одному другому клиентскому устройству,
- принимают по меньшей мере одно сгенерированное сообщение и отправляют его в единое информационное пространство.

## **КРАТКОЕ ОПИСАНИЕ ЧЕРТЕЖЕЙ**

[0026] На Фиг. 1 представлена концептуальная схема заявленного решения.

[0027] На Фиг. 2 представлена схема сервера видеоконференцсвязи.

[0028] На Фиг. 3 представлена блок-схема способа осуществления видеоконференцсвязи.

[0029] На Фиг. 4 представлена общая схема вычислительного устройства.

## **ОСУЩЕСТВЛЕНИЕ ИЗОБРЕТЕНИЯ**

[0030] Ниже будут описаны понятия и термины, необходимые для понимания настоящего изобретения.

[0031] Видеоконференцсвязь (ВКС) – это телекоммуникационная технология взаимодействия двух и более удалённых абонентов, при которой между ними возможен обмен аудио- и видеоинформацией в реальном времени, с учётом передачи управляющих данных.

[0032] XMPP (eXtensible Messaging and Presence Protocol, Расширяемый протокол обмена сообщениями и информацией о присутствии) — открытый, основанный на XML (eXtensible Markup Language, Расширяемый язык разметки), свободный для использования протокол для мгновенного обмена сообщениями и информацией о присутствии в режиме, близком к режиму реального времени.

[0033] RTP (Real-time Transport Protocol) — это сетевой протокол для доставки аудио и видео по IP-сетям.

[0034] SRTP (Secure Real-time Transport Protocol, Безопасный протокол передачи данных в реальном времени) — расширение к протоколу RTP, предназначенное для шифрования, установления подлинности сообщения, целостности, защиты от подмены данных RTP.



[0035] gRPC (Remote Procedure Calls) — это система удалённого вызова процедур (RPC) с открытым исходным кодом, предоставляющая такие функции как аутентификация, двунаправленная потоковая передача и управление потоком, блокирующие или неблокирующие привязки, а также отмена и тайм-ауты.

[0036] ASR (Automatic Speech Recognition) — компьютерная технология автоматического распознавания речи в текст.

[0037] Gesture Recognition – объединение компьютерных и языковых технологий, целью которых является интерпретация человеческих жестов с помощью математических алгоритмов.

[0038] SLR (Sign Language Recognition) — компьютерная технология по распознаванию действий на языке жестов.

[0039] AR (Augmented Reality, Дополненная реальность) — результат введения в зрительное поле любых сенсорных данных с целью дополнения сведений об окружении и изменения восприятия окружающей среды.

[0040] WebRTC (Web Real Time Communication, Веб коммуникации в режиме реального времени) – технология, обеспечивающая организацию передачи потоковых данных между браузерами или мобильными приложениями используя прямую одноранговую связь.

[0041] BLE (Bluetooth Low Energy, Беспроводная технология Bluetooth с низким энергопотреблением) — технология беспроводной связи с низким энергопотреблением для обмена данными на небольших расстояниях в течение длительного периода времени.

[0042] На Фиг. 1 представлена общая схема заявленной системы (100) для видеоконференцсвязи. Система (100) может представлять собой программно-аппаратный комплекс, в котором каждый из ее элементов расположен на отдельной вычислительной машине, связанный в рамках единого функционального обеспечения с другими элементами посредством сети передачи данных (103).

[0043] Система (100) состоит из по меньшей мере двух клиентских устройств (101, 102) и сервера видеоконференцсвязи (104) соединенных посредством сети передачи данных (103).

[0044] Клиентские устройства (101, 102) могут быть выполнены с возможностью генерации по меньшей мере одного сообщения посредством устройств ввода информации, записи аудио и/или видеосигналов посредством таких технических средств, как микрофоны и камеры; преобразования упомянутых сигналов в поток данных; соединения с сервером видеоконференцсвязи (104); передачи серверу видеоконференцсвязи (104) упомянутых по меньшей мере одного сообщения и потока данных; получения от сервера

видеоконференцсвязи (104) по меньшей мере одного сообщения и потока данных от других клиентских устройств; отображения по меньшей мере одного сообщения в едином информационном пространстве и потока данных посредством устройств вывода информации.

[0045] Клиентские устройства (101, 102) могут быть реализованы на базе вычислительных устройств, модифицированных в программно-аппаратной части таким образом, чтобы выполнять указанные выше функции клиентских устройств, и обладающих техническими средствами для ввода-вывода информации, а также техническими средствами для коммуникации с сервером видеоконференцсвязи (104) посредством сети передачи данных (103). Примерами таких вычислительных устройств являются, но не ограничиваются: персональный компьютер, смартфон, устройства SberBox Top, SberPortal и т.п. Более подробное описание вычислительного устройства раскрыто далее со ссылкой на Фиг. 4.

[0046] В качестве сети передачи данных (103) может выступать, но не ограничиваться указанными примерами: внутренняя или внешняя вычислительная сеть, например, Интранет, Интернет и т.п.

[0047] Сервер видеоконференцсвязи (104) содержит совокупность модулей, которые будут подробно описаны далее со ссылкой на Фиг. 2. Модули могут выполняться конструктивно в виде программно-аппаратных решений (например, система на чипе, микроконтроллеры и т.п.). Кроме того, указанный сервер (104) может быть реализован на базе вычислительного устройства, более подробное описание которого представлено далее со ссылкой на Фиг. 4. Причем такое вычислительное устройство модифицировано в своей программно-аппаратной части таким образом, чтобы реализовывать функции упомянутых модулей.

[0048] На Фиг. 2 представлена подробная схема реализации сервера видеоконференцсвязи (104). Сервер видеоконференцсвязи (104) содержит следующие модули: модуль обмена сообщениями (201), модуль пересылки потоков (202), модуль распознавания потоков (203), модуль фильтрации аудио потока (204), модуль модификации видео потока (205), модуль согласования (206), модуль управления единым информационным пространством (207), модуль перевода (208). Взаимодействие между модулями реализовано посредством шины данных.

[0049] Модуль обмена сообщениями (201) регистрирует подключения клиентских устройств к конференции. После регистрации модуль обмена сообщениями (201) выполнен с возможностью принимать по меньшей мере одно сообщение, сгенерированное на по

меньшей мере одном клиентском устройстве, от по меньшей мере одного клиентского устройства и отправлять упомянутое сообщение в единое информационное пространство.

[0050] В частном примере реализации в качестве единого информационного пространства может выступать чат, доступный всем клиентским устройствам, которые прошли процесс регистрации подключения к конференции посредством модуля обмена сообщениями (201). Сообщения, отправляемые в единое информационное пространство, отображаются на устройствах вывода клиентских устройствах, например, но не ограничиваясь, в виде текстовых сообщений в чате. В одном из частных примеров реализации отправка по меньшей мере одного сообщения в единое информационное пространство включает отправку упомянутого сообщения модулем обмена сообщениями (201) всем клиентским устройствам, которые прошли процесс регистрации подключения к конференции.

[0051] Необходимо отметить, что в одном из альтернативных вариантов воплощения заявленного решения модуль обмена сообщениями (201) выполнен с возможностью принимать от модуля пересылки потоков (202) по меньшей мере одно сообщение, сгенерированное модулем распознавания потоков (203) и отправлять упомянутое сообщение в единое информационное пространство.

[0052] В предпочтительном варианте воплощения заявленного решения модуль обмена сообщениями (201) является XMPP-сервером.

[0053] Модуль пересылки потоков (202) устанавливает защищенное соединение с клиентскими устройствами. После установки защищенного соединения модуль пересылки потоков (202) принимает поток данных от по меньшей мере одного клиентского устройства и передаёт поток данных в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков (203). Далее модуль пересылки потоков (202) принимает от модуля распознавания потоков (203) по меньшей мере одно сгенерированное сообщение и отправляет упомянутое сообщение в единое информационное пространство.

[0054] В частном примере реализации в качестве защищенного соединения может выступать SRTP соединение.

[0055] Принятый от по меньшей мере одного клиентского устройства поток данных содержит по меньшей мере аудио и/или видео поток. Например, у клиентского устройства может отсутствовать камера, в связи с чем клиентское устройство будет передавать только аудио поток. В случае, если пользователь клиентского устройства общается только на языке жестов, то передача аудио потока не является обязательным условием для коммуникации

между участниками конференции и модуль пересылки потоков (202) может принять только видео поток.

[0056] В одном из частных примеров реализации модуль пересылки потоков (202) принимает от по меньшей мере одного клиентского устройства команду на распознавание аудио и/или видео потока и передаёт аудио и/или видео поток модулю распознавания потоков (203) после получения упомянутой команды на распознавание. Возможность модуля пересылки потоков (202) принимать команду на распознавание позволяет заявленному решению эффективно использовать вычислительные и временные ресурсы и либо не тратить их на распознавание аудио и/или видео потока, если такое распознавание не требуется, либо определять какой из потоков данных (аудио или видео поток) подлежит распознаванию.

[0057] Необходимо отметить, что в одном из альтернативных вариантов воплощения заявленного решения модуль пересылки потоков (202) отправляет сгенерированное сообщение, принятое от модуля распознавания потоков (203), в единое информационное пространство посредством отправки упомянутого сгенерированного сообщения модулю обмена сообщениями (201), который в свою очередь отправляет упомянутое сгенерированное сообщение в единое информационное пространство.

[0058] Модуль распознавания потоков (203) обрабатывает поток данных, полученный от модуля пересылки потоков (202), и генерирует по меньшей мере одно сообщение на основе обработанного потока данных и отправляет сгенерированное сообщение модулю пересылки потоков (202). Обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных модулем пересылки потоков (202) по меньшей мере одному другому клиентскому устройству. Возможность параллельного выполнения операций по распознаванию потока данных и передаче упомянутого потока данных клиентскому устройству для дальнейшего вывода позволяет заявленному решению сохранить скорость передачи потока данных клиентскому устройству с одновременным обеспечением функции распознавания потока данных.

[0059] В предпочтительном варианте воплощения заявленного решения в зависимости от типа потока данных модуль распознавания потоков (203) обрабатывает поток данных следующим образом:

- если поток данных включает аудио поток, то обработка аудио потока данных включает преобразование речи в текст (ASR),
- если поток данных включает видео поток, то обработка видео потока может включать как преобразование жестов в текстовую и/или графическую информацию (Gesture Recognition), так и распознавание языка жестов (SLR).

[0060] В частном примере реализации модуль распознавания потоков (203) дополнительно выполнен с возможностью определять тип потока данных и осуществлять соответствующую обработку потока данных в зависимости от типа потока данных. Процесс определения типа потока данных заключается в исследовании потока данных и определении его типа.

[0061] В общем случае, модуль распознавания потоков (203) дополнительно выполнен с возможностью осуществлять разбиение потока данных на фрагменты в процессе обработки потока данных, обрабатывать фрагменты, полученные в результате разбиения потока данных, и генерировать по меньшей мере одно сообщение на основе обработанных фрагментов. В качестве фрагментов могут выступать, но не ограничиваться указанными примерами: речевые фразы (для аудио потока) или кадры (для видео потока).

[0062] В предпочтительном варианте воплощения изобретения для случая преобразования речи в текст (ASR) модуль распознавания потоков (203) дополнительно выполнен с возможностью осуществлять разбиение аудио потока на фрагменты в процессе преобразования речи в текст, обрабатывать фрагменты, полученные в результате разбиения аудио потока, и генерировать по меньшей мере одно сообщение на основе обработанных фрагментов. При этом разбиение аудио потока на фрагменты включает детектирование окончания фрагмента путём детектирования тишины в аудио потоке, характерной, например, для окончания фразы или грамматической паузы.

[0063] Обработка фрагментов, полученных в результате разбиения аудио потока, посредством модуля распознавания потоков (203) дополнительно включает формирование набора гипотез по каждому слову во фрагменте исходя из языковой модели. При этом на каждое сказанное пользователем клиентского устройства слово может быть несколько вариантов распознанного текста. Гипотеза – это набор пар, состоящий из варианта распознанного слова и вероятности того, что это слово встречается в распознаваемом фрагменте. Например, для распознаваемого слова «Лук» в качестве гипотезы выступает следующий набор пар:

Лук - 0,5

Луг - 0,3

Люк - 0,2

[0064] Таким образом, модуль распознавания потоков (203) изначально не распознает каждое слово, а лишь формирует набор вариантов определения того или иного слова с их вероятностью (где сумма вероятностей равна 1 или 100%). При этом, в предпочтительном варианте реализации, из набора вариантов выбирается то слово, которое имеет наибольшую вероятность. В качестве примера, но не ограничиваясь им, оценка вероятности может

основываться на сходстве звука в слове с акустическими моделями ASR или местоположении слова в предложении (например, с учётом языковой или грамматической модели).

[0065] Упомянутая обработка фрагментов дополнительно включает выбор модулем распознавания потоков (203) для каждого слова из фрагмента наиболее вероятной гипотезы из набора гипотез. На основе выбранных слов модуль распознавания потоков (203) генерирует по меньшей мере одно сообщение, содержащее распознанный фрагмент аудио потока. Модуль распознавания потоков (203) может дополнительно осуществлять очищение сгенерированного сообщения от обсценной лексики. В частном примере реализации очищение сгенерированного сообщения от обсценной лексики осуществляют путём сравнения распознанных слов со словарем обсценной лексики, и замены обнаруженных слов, относящихся к обсценной лексике, на специальные символы, например, звёздочки, решётки или любые их комбинации. Кроме того, модуль распознавания потоков (203) может дополнительно восстановить пунктуацию (т.е. расставить знаки препинания) и капитализацию (т.е. расставить заглавные буквы там, где необходимо по правилам русского языка) сгенерированного сообщения. В результате модуль распознавания потоков (203) отправляет по меньшей мере одно сгенерированное сообщение, очищенное от обсценной лексики с восстановленной пунктуацией и капитализацией, модулю пересылки потоков (202).

[0066] В предпочтительном варианте воплощения изобретения преобразование речи в текст (ASR) реализовано на основе собственной технологии от SberDevices — SmartSpeech, которая обеспечивает высокую скорость и отличное качество распознавания, в том числе имён, сложных терминов и длинных слов. В SmartSpeech используются самые последние наработки в сфере Deep Learning (глубокое обучение). Нейросети обучаются на огромных объёмах данных с использованием мощностей суперкомпьютера «Кристофари» (до 1000+ GPU Tesla v100) от Сбера. Нейросети используют GPU для сверхбыстрой работы. В распознавании речи используются сверхточные архитектуры, такие как Jasper, QuartzNet и другие.

[0067] Предобученная модель для распознавания речи используемая в SmartSpeech обучалась на самом большом размеченном вручную датасете из 1240 часов аудиоданных на русском языке (речевой датасет Golos).

[0068] В одном из частных примеров реализации преобразование речи в текст может быть реализовано, но не ограничиваться, на основе следующих технологий: Microsoft Speech SDK, IBM Embedded Via Voice и т.п.

[0069] В предпочтительном варианте воплощения изобретения для случая преобразования жестов в текстовую и/или графическую информацию (Gesture Recognition) модуль распознавания потоков (203) выполняет разбивку видео потока на кадры, определяет область интереса (например, руку или ладонь) в по меньшей мере одном кадре видео потока, сегментирует определенную область интереса, выявляет ключевые точки на сегментированной области, и определяет жест путём сопоставления ключевых точек с одним из известных жестов из базы данных известных жестов.

[0070] В случае определения жеста модуль распознавания потоков (203) определяет тип жеста и в зависимости от типа жеста генерирует по меньшей мере одно соответствующее сообщение и отправляет модулю пересылки потоков (202).

[0071] В одном из частных примеров реализации, если тип жеста определен, как жест, характеризующий реакцию пользователя, то модуль распознавания потоков (203) генерирует по меньшей мере одно сообщение, содержащее текстовую и/или графическую информацию. Например, если модуль распознавания потоков (203) определил в видеопотоке пользователя, показывающего жест «большой палец вверх», то сгенерированное сообщение может содержать, как эмодзи, обозначающее большой палец вверх, так и текстовое сообщение «Превосходно!» или «Полностью поддерживаю!», либо комбинацию из эмодзи и текстового сообщения.

[0072] Если тип жеста определен, как управляющий жест, модуль распознавания потоков (203) дополнительно выполнен с возможностью отправки сообщения, содержащего распознанный управляющий ввод, модулю пересылки потоков (202). Распознанный управляющий ввод включает управляющую команду клиентскому устройству. К таким управляющим командам может относиться, но не ограничиваться: выполнение различных действий по управлению громкостью звука, включению/отключению звука, управлению камерой видеоконференции и/или доступу к настройкам для видеоконференции. Получив упомянутое сообщение, модуль пересылки потоков (202) отправляет его модулю обмена сообщениями (201), который извлекает из сообщения управляющую команду и отправляет её клиентскому устройству.

[0073] Например, если модуль распознавания потоков (203) определил в видеопотоке пользователя, показывающего жест «указательный палец вверх», то на клиентское устройство отправляется управляющая команда о повышении громкости звука.

[0074] В предпочтительном варианте воплощения изобретения для случая распознавания языка жестов (Sign Language Recognition) модуль распознавания потоков (203) выполняет разбивку видео потока на кадры, определяет по меньшей мере одну область интереса (например, руки) в по меньшей мере одном кадре видео потока,

сегментирует по меньшей мере одну определенную область интереса, выявляет ключевые точки на по меньшей мере одной сегментированной области интереса, нормализует выявленные ключевые точки, и определяет по меньшей мере один жест путём сопоставления нормализованных ключевых точек с одним из известных жестов из базы данных известных жестов. Далее модуль распознавания потоков (203) создаёт матрицу распознанных букв или слов из жестов, определенных на предыдущем этапе, и генерирует по меньшей мере одно сообщение на основе созданной матрицы распознанных букв или слов и отправляет упомянутое по меньшей мере одно сообщение модулю пересылки потоков (202).

[0075] Необходимо отметить, что в одном из альтернативных вариантов воплощения заявленного решения часть функций модуля распознавания потоков (203), связанных с предварительной обработкой видео потока, таких как разбивка видео потока на кадры, определение области интереса в кадре видео потока, сегментация области интереса, могут быть реализованы на стороне клиентского устройства. Таким образом, модуль распознавания потоков (203) будет принимать от модуля пересылки потоков (202) уже предварительно обработанный клиентским устройством видео поток.

[0076] Модуль фильтрации аудио потока (204) выполняет подавление шума. В частном варианте воплощения заявленного решения перед осуществлением обработки аудио потока модуль распознавания потоков (203) отправляет аудио поток модулю фильтрации аудио потока (204), который после выполнения операции по подавлению шума отправляет очищенный от шума аудио поток обратно модулю распознавания потоков (203).

[0077] В качестве примера, но не ограничиваясь им, возможность подавления шума в модуле фильтрации аудио потока (204) может быть реализована с помощью технологии умного шумоподавления на базе платформы Maxine от NVIDIA или на основе собственной технологии от SberDevices — SberNoise. Эти решения основаны на использовании искусственного интеллекта для автоматического удаления посторонних шумов.

[0078] Интеграция функции шумоподавления в заявленное решение, повышает комфортность общения, если человек находится в шумном месте во время звонка, удаляя из звучания посторонние шумы и резкие звуки. Например, если на фоне разговора плачет ребёнок, лает собака или из окна доносится строительный шум, благодаря использованию AI-модели заявленное решение заглушает эти звуки, оставляя только голос говорящего.

[0079] Модуль модификации видео потока (205) выполняет замену фона, применяет фильтры, улучшающие визуальные характеристики изображения пользователя, а также фильтры дополненной реальности (AR-фильтры).



[0080] Необходимость замены фона, отображаемого позади пользователя клиентского устройства во время осуществления видеоконференцсвязи, может возникнуть в связи с нахождением пользователя в нетрадиционном месте работы, например, в домашнем офисе, в отпуске и т.п., кроме того текущий фон видео потока пользователя может отвлекать других участников видеоконференции.

[0081] Процесс замены фона осуществляется модулем модификации видео потока (205) следующим образом:

- получают видео поток от модуля пересылки потоков (202);
- выполняют идентификацию фоновой части видео потока и изображения участника видеоконференцсвязи;
- выполняют сегментацию фона (background segmentation, BGS);
- заменяют сегментированный фон на по меньшей мере одно из: любое изображение, выбранное пользователем, фоновый шаблон или размытое изображение текущего фона;
- отправляют видео поток, содержащий замененный фон, в модуль пересылки потоков (202) для дальнейшей передачи клиентским устройствам.

[0082] В заявленном решении под улучшением визуальных характеристик изображения пользователя понимается сглаживание кожи, выравнивание цвета лица (например, удаление синяков под глазами) и коррекция погрешностей линзы камеры телефона. Этот процесс называется бьютификацией.

[0083] Процесс улучшения визуальных характеристик изображения пользователя (бьютификации) осуществляется модулем модификации видео потока (205) следующим образом:

- получают по меньшей мере один кадр видео потока от модуля пересылки потоков (202);
- идентифицируют лицо по меньшей мере в одном кадре видео потока;
- сегментируют части лица по оттенку кожи, например, глаза и/или рот;
- идентифицируют по меньшей мере одну дефектную область, которая отличаются по яркости на пороговую величину от участков оттенка кожи без дефектов;
- сглаживают яркость по меньшей мере одной дефектной области для создания сглаженной яркости;
- генерируют по меньшей мере один улучшенный кадр видео потока, включающий в себя улучшенную версию лица, в которой исходная яркость по меньшей мере одной дефектной области заменена сглаженной яркостью; и

- отправляют видео поток, содержащий улучшенные кадры, в модуль пересылки потоков (202) для дальнейшей передачи клиентским устройствам.

[0084] В заявленном решении под AR-фильтрами понимаются созданные компьютером визуальные эффекты, предназначенные для наложения на реальные изображения путем добавления дополнительного слоя к переднему или заднему плану реального изображения.

[0085] Процесс применения AR-фильтров осуществляется модулем модификации видео потока (205) следующим образом:

- получают по меньшей мере один кадр видео потока от модуля пересылки потоков (202);
- идентифицируют по меньшей мере одну область интереса (например, лицо) в по меньшей мере одном кадре видео потока;
- определяют координаты ориентиров по меньшей мере одной области интереса (например, лицевые ориентиры, причем к лицевым ориентирам относятся по меньшей мере нос, глаза, рот и очертания лица);
- осуществляют наложение по меньшей мере одного элемента AR на по меньшей мере одну соответствующую область интереса (например, наложение элемента AR «очки» на область лица, а конкретно на координаты, соответствующие глазам) с учетом определенных координат ориентиров и размера по меньшей мере одной области интереса;
- генерируют по меньшей мере один измененный кадр видео потока с примененным AR-фильтром, в котором элемент AR наложен на соответствующую область интереса; и
- отправляют видео поток, содержащий измененные кадры, в модуль пересылки потоков (202) для дальнейшей передачи клиентским устройствам.

[0086] В одном из частных примеров реализации процесс применения AR-фильтров дополнительно содержит отслеживание движения ориентиров по меньшей мере одной области интереса (например, движение головой в стороны, изменение мимики (моргание, улыбка), движение рук и т.п.) и добавление по меньшей мере одного элемента AR в по меньшей мере один кадр видео потока на основании изменения движения ориентиров по меньшей мере одной области интереса.

[0087] Необходимо отметить, что в одном из альтернативных вариантов воплощения заявленного решения модуль модификации видео потока (205) может быть реализован на стороне клиентского устройства. Таким образом, модуль пересылки потоков (202) будет принимать уже модифицированный видео поток от клиентских устройств.

[0088] Однако в предпочтительном варианте воплощения заявленного решения модуль модификации видео потока (205) реализуется на стороне сервера видеоконференцсвязи

(104). Такое воплощение позволяет модулю распознавания потоков (203) обрабатывать исходный, а не модифицированный видео поток, что повышает эффективность и точность обработки при преобразовании жестов в текстовую и/или графическую информацию (Gesture Recognition) и распознавании языка жестов (Sign Language Recognition), поскольку в модифицированном видео потоке могут содержаться визуальные эффекты, которые будут ухудшать качество обработки (распознавания).

[0089] В общем случае, модуль согласования (206) предназначен для согласования отправки сообщений в единое информационное пространство. Модуль согласования (206) определяет в режиме реального времени, осуществляет ли модуль распознавания потоков (203) обработку потока данных, в частности, аудио потока. Указанная возможность может быть реализована, например, с помощью отправки модулю распознавания потоков (203) запросов с определенным интервалом, где результатом отправки запроса является возврат модулю согласования (206) ответа о состоянии модуля распознавания потоков (203) в части осуществления обработки потока данных. Например, такое состояние может свидетельствовать о том, осуществляется ли в текущий момент голосовой ввод пользователя клиентского устройства или ввод с помощью языка жестов, обрабатываемый с помощью модуля распознавания потоков (203).

[0090] Затем модуль согласования (206) устанавливает значение флага, идентифицирующего состояние текущего голосового ввода пользователя клиентского устройства. В другом варианте воплощения заявленного решения, значение флага идентифицирует состояние текущего ввода с помощью языка жестов.

[0091] Далее модуль согласования (206) получает от модуля обмена сообщениями (201) запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства, отличного от клиентского устройства, осуществляющего текущий голосовой ввод или ввод с помощью языка жестов. Кроме того, модуль согласования (206) может отправлять модулю обмена сообщениями (201) разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства на основании значения упомянутого флага.

[0092] В частном примере реализации модуль согласования (206) выполнен с возможностью отправлять модулю пересылки потоков (202) запрос на отправку в единое информационное пространство сгенерированного сообщения, содержащего уже обработанные фрагменты, если значение флага идентифицирует наличие текущего голосового ввода пользователя клиентского устройства или ввода с помощью языка жестов. В качестве обработанных фрагментов могут выступать, например, текстовые сообщения, полученные в результате преобразования речи в текст или языка жестов в текст.

[0093] Модуль управления единым информационным пространством (207) определяет, что по меньшей мере одно сгенерированное сообщение получено от модуля пересылки потоков (202) и отображает при выводе по меньшей мере одного сгенерированного сообщения в единое информационное пространство метку, идентифицирующую формирование сгенерированного сообщения из аудио потока или видео потока. Такая метка позволяет пользователю однозначно определить к какому типу сообщения относится конкретное сообщение, отправленное в единое информационное пространство. В частном примере реализации упомянутая метка может быть различна для сгенерированного сообщения из аудио потока (т.е. для сообщения, являющегося результатом распознавания аудио потока путём преобразования речи в текст) и сгенерированного сообщения из видео потока (т.е. для сообщения, являющегося результатом распознавания видео потока путём преобразования языка жестов в текст).

[0094] Модуль управления единым информационным пространством (207) дополнительно позволяет сохранять сообщения, размещенные в едином информационном пространстве, и обеспечивать доступ подключенным клиентским устройствам к сохраненным сообщениям. Такая возможность позволяет пользователю клиентского устройства во время конференции больше концентрироваться на общении, а не на ведении заметок, поскольку ко всей совокупности сообщений в едином информационном пространстве обеспечен неограниченный доступ.

[0095] Кроме того, модуль управления единым информационным пространством (207) позволяет скачать полный текст разговора (т.е. всю совокупность сообщений, отправленных в единое информационное пространство) в любой момент конференции любому клиентскому устройству, которое прошло процесс регистрации подключения к конференции с помощью модуля обмена сообщениями (201). В качестве примера, но не ограничиваясь им, клиентское устройство отправляет запрос на скачивание полного текста разговора модулю управления единым информационным пространством (207). В ответ на упомянутый запрос, модуль управления единым информационным пространством (207) генерирует всю совокупность сообщений, отправленных в единое информационное пространство, например, в виде текстового файла, и отправляет клиентскому устройству.

[0096] В одном из частных примеров реализации модуль управления единым информационным пространством (207) обеспечивает управление отображением сообщений, размещенных в едином информационном пространстве. По запросу пользователя клиентского устройства в едином информационном пространстве могут быть отражены, но не ограничиваясь: только сообщения, отправленные непосредственно в единое информационное пространство; только сообщения, сгенерированные модуле

распознавания потоков (203), представляющие собой текстовую расшифровку аудио или видео потока; все сообщения, отправленные в единое информационное пространство.

[0097] Модуль управления единым информационным пространством (207) также осуществляет поиск по содержимому сообщений, размещенных в едином информационном пространстве, что позволяет ускорить процесс нахождения необходимой информации среди всей совокупности сообщений, сгенерированных участниками во время конференции.

[0098] Модуль перевода (208) выполняет перевод по меньшей мере одного сообщения, полученного от по меньшей мере одного клиентского устройства посредством модуля обмена сообщениями (201) и/или сообщения, сгенерированного на основе обработанного потока данных модулем распознавания потоков (203).

[0099] Процесс перевода осуществляется модулем перевода (208) следующим образом:

- получают запрос от по меньшей мере одного клиентского устройства на перевод по меньшей мере одного сообщения, при этом упомянутый запрос содержит указание на требуемый язык перевода и указание на по меньшей мере одно сообщение, которое необходимо перевести;
- осуществляют перевод указанного в запросе по меньшей мере одного сообщения на требуемый язык перевода;
- отправляют по меньшей мере одно переведенное сообщение клиентскому устройству посредством модуля обмена сообщениями (201).

[0100] На Фиг. 3 представлен способ (300) осуществления видеоконференцсвязи. Выполнение способа (300) осуществляется с помощью сервера видеоконференцсвязи (104) и по меньшей мере двух клиентских устройств (101, 102). На первом этапе (301) подключают по меньшей мере два клиентских устройства (101, 102) к серверу видеоконференцсвязи (401). В частном примере реализации по меньшей мере два клиентских устройства (101, 102) подключаются к модулю обмена сообщениями (201), входящему в состав сервера видеоконференцсвязи (104). Подключение к серверу видеоконференцсвязи (401) (в частном случае, к модулю обмена сообщениями (201)) может быть реализовано посредством получения клиентским устройством ссылки на присоединение к конференции и перехода по указанной ссылке, например, но не ограничиваясь, с помощью технологии WebRTC.

[0101] После подключения по меньшей мере двух клиентских устройств (101, 102), на этапе (302) регистрируют подключения упомянутых клиентских устройств (101, 102) к конференции посредством модуля обмена сообщениями (201). В частном примере реализации на этапе (302) модулем обмена сообщениями (201) дополнительно может

осуществляться отправка уведомлений клиентским устройствам о вновь подключенных клиентских устройствах.

[0102] На этапе (303) устанавливается защищенное соединение с клиентскими устройствами посредством модуля пересылки потоков (202). В частном примере реализации с клиентскими устройствами устанавливается SRTP соединение. В дальнейшем установленное защищенное соединение используется для двусторонней передачи потоков данных между по меньшей мере одним клиентским устройством и сервером видеоконференцсвязи (104) (в частном случае, модулем пересылки потоков (202), входящим в состав сервера видеоконференцсвязи (104)), а также между клиентскими устройствами через сервер видеоконференцсвязи (104) (в частном случае, через модуль пересылки потоков (202)).

[0103] После установки защищенного соединения, на этапе (304) посредством модуля пересылки потоков (202) принимаются поток данных от по меньшей мере одного клиентского устройства. Полученный поток данных посредством модуля пересылки потоков (202) на этапе (305) передается в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков (203).

[0104] В частном примере реализации на этапе (305) передача потока данных модулю распознавания потоков (203) включает отправку модулем пересылки потоков (202) gRPC-запроса модулю распознавания потоков (203). Упомянутый gRPC-запрос может содержать, но не ограничиваясь:

- информацию для аутентификации в модуле распознавания потоков (203) (например, размещенную в заголовке gRPC-запроса);
- параметры распознавания, такие как тип потока данных;
- поток данных для распознавания в модуле распознавания потоков (203).

[0105] После получения потока данных, на этапе (306) посредством модуля распознавания потоков (203) обрабатываются поток данных и генерируется по меньшей мере одно сообщение на основе обработанного потока данных, причем обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных модулем пересылки потоков (202) по меньшей мере одному другому клиентскому устройству.

[0106] В частном примере реализации на этапе (306) обработка потока данных модулем распознавания потоков (203) включает получение gRPC-запроса от модуля пересылки потоков (202), а также определение типа потока данных и осуществление соответствующей обработки потока данных в зависимости от типа потока данных. Процесс определения типа потока данных может быть осуществлен, например, путем считывания типа потока данных

из параметров распознавания, содержащихся в gRPC-запросе, полученном от модуля пересылки потоков (202).

[0107] После генерации по меньшей мере одного сообщения на основе обработанного потока данных посредством модуля распознавания потоков (203) передают упомянутое по меньшей мере одно сгенерированное сообщение модулю пересылки потоков (202). В частном примере реализации передача по меньшей мере одного сгенерированного сообщения модулю пересылки потоков (202) включает отправку модулем распознавания потоков (203) gRPC-ответа модулю пересылки потоков (202). Упомянутый gRPC-ответ содержит результаты распознавания, полученные в результате обработки потока данных посредством модуля распознавания потоков (203).

[0108] На этапе (307) посредством модуля пересылки потоков (202) принимают по меньшей мере одно сгенерированное сообщение от модуля распознавания потоков (203) и отправляют его в единое информационное пространство. В частном примере реализации по меньшей мере одно сгенерированное сообщение может содержать, но не ограничиваясь:

- идентификатор клиентского устройства;
- временная метка, идентифицирующая начало формирования сообщения;
- временная метка, идентифицирующая окончание формирования сообщения;
- сообщение, содержащее результаты распознавания;
- идентификатор сессии распознавания для конкретного клиентского устройства.

[0109] В одном из частных примеров реализации заявленного решения на этапе (307) посредством модуля пересылки потоков (202) отправляют сгенерированное сообщение, принятое от модуля распознавания потоков (203), в единое информационное пространство через отправку упомянутого сгенерированного сообщения модулю обмена сообщениями (201), который отправляет упомянутое сгенерированное сообщение в единое информационное пространство.

[0110] Один из альтернативных вариантов воплощения заявленного изобретения раскрывает решение проблемы некорректного вывода сообщений в единое информационное пространство. Рассмотрим случай, когда пользователь первого клиентского устройства передаёт поток данных (например, читает доклад), пользователи других клиентских устройств принимают этот поток данных (т.е. слышат и/или видят доклад первого пользователя) посредством модуля пересылки потоков (202). Поскольку модуль пересылки потоков (202) передаёт поток данных в параллельном режиме клиентским устройствам и модулю распознавания потоков (203), то пользователи других клиентских устройств могут получить поток данных раньше, чем модуль распознавания потоков (203) сгенерирует сообщение на основе обработанного потока данных, и оно будет

выведено в единое информационное пространство (далее - чат). В связи с этим может возникнуть ситуация, когда пользователь второго клиентского устройства, услышав часть доклада, задаёт вопрос по теме услышанного, например, генерирует сообщение и отправляет его в чат. Таким образом, сообщение от пользователя второго клиентского устройства может быть выведено в чат раньше, чем сообщение от пользователя первого клиентского устройства, преобразованное из потока данных (например, распознанный и преобразованный из аудио потока в текстовое сообщение доклад пользователя первого клиентского устройства), что приведёт к некорректному выводу сообщений в единое информационное пространство и как следствие, потере логической, структурной и причинно-следственной связи между совокупностью сообщений в едином информационном пространстве.

[0111] Для устранения указанной проблемы заявленное решение реализовано следующим образом.

[0112] На этапе (306) посредством модуля распознавания потоков (203) дополнительно осуществляют разбиение текста на фрагменты в процессе преобразования речи в текст, и генерируют по меньшей мере одно сообщение на основе фрагментов, полученных в результате разбиения текста. Затем в режиме реального времени посредством модуля согласования (206) определяют, осуществляет ли модуль распознавания потоков (203) обработку потока данных (например, аудио потока); и устанавливают значение флага, идентифицирующего состояние текущего ввода пользователя клиентского устройства (например, голосового ввода или ввода с помощью языка жестов). В качестве примера, но не ограничиваясь, устанавливают  $flag=1$ , если модуль распознавания потоков (203) осуществляет обработку потока данных в текущий момент времени, и устанавливают  $flag=0$ , если модуль распознавания потоков (203) не осуществляет обработку потока данных в текущий момент времени, т.е. ни один участник конференции не осуществляет голосовой ввод или ввод с помощью языка жестов, который необходимо распознать посредством модуля распознавания потоков (203).

[0113] Далее посредством модуля обмена сообщениями (201) принимают по меньшей мере одно сообщение от другого клиентского устройства (например, от пользователя второго клиентского устройства), и отправляют модулю согласования (206) запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства. Модуль согласования (206) получает от модуля обмена сообщениями (201) запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства и принимает решение о разрешении отправки на основании значения установленного флага.



[0114] Если значение флага идентифицирует отсутствие текущего ввода (голосового или с помощью языка жестов) пользователя клиентского устройства (например, flag=0), модуль согласования (206) отправляет модулю обмена сообщениями (201) разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства. Далее модуль обмена сообщениями (201) получает упомянутое разрешение от модуля согласования (206) и отправляет в единое информационное пространство сообщения от другого клиентского устройства.

[0115] Если значение флага идентифицирует наличие текущего ввода (голосового или с помощью языка жестов) пользователя клиентского устройства (например, flag=1), модуль согласования (206) отправляет модулю пересылки потоков (202) запрос на отправку в единое информационное пространство сгенерированного сообщения, содержащего уже обработанные фрагменты. Получив упомянутый запрос, модуль пересылки потоков (202) отправляет модулю распознавания потоков (203) запрос на генерацию сообщения, содержащего уже обработанные фрагменты. В ответ на запрос, модуль распознавания потоков (203) генерирует сообщение, содержащее уже обработанные фрагменты, и отправляет упомянутое сгенерированное сообщение модулю пересылки потоков (202), который отправляет сгенерированное сообщение, содержащее уже обработанные фрагменты, в единое информационное пространство. Затем модуль согласования (206) отправляет модулю обмена сообщениями (201) разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства. Далее модуль обмена сообщениями (201) получает упомянутое разрешение от модуля согласования (206) и отправляет в единое информационное пространство сообщение от другого клиентского устройства. Таким образом, будет обеспечена сохранность логической, структурной и причинно-следственной связи между совокупностью сообщений в едином информационном пространстве.

[0116] В одном из альтернативных вариантов воплощения заявленного решения система (100) для видеоконференцсвязи выполнена с возможностью бесшовного переключения между клиентскими устройствами во время осуществления видеоконференцсвязи. Упомянутое бесшовное переключение реализуется при помощи по меньшей мере одного клиентского устройства, которое прошло процесс регистрации подключения к конференции посредством модуля обмена сообщениями (201), следующим образом:

- активируют функцию по переносу конференции на другое клиентское устройство и осуществляют поиск клиентского устройства с поддержкой BLE путём сканирования ближайших устройств;

- получают результаты сканирования, содержащие по меньшей мере одно клиентское устройство с поддержкой BLE;
- проверяют расстояние до найденного клиентского устройства путём сравнения индикатора уровня принятого сигнала BLE (BLE RSSI) с пороговым значением;
- идентифицируют найденное клиентское устройство, как находящееся рядом с клиентским устройством, подключенным к конференции, если BLE RSSI меньше порогового значения;
- устанавливают BLE подключение между найденным клиентским устройством и клиентским устройством, подключенным к конференции;
- запускают конференцию на найденном клиентском устройстве путём передачи данных между упомянутыми клиентскими устройствами посредством BLE подключения.

[0117] На Фиг. 4 представлен общий вид вычислительного устройства (400), на базе которого могут быть реализованы элементы заявленной системы (100), такие как клиентские устройства (101, 102) и сервер видеоконференцсвязи (104). В частном примере реализации, каждый из модулей (201, 202, 203, 204, 205, 206, 207, 208), входящих в состав сервера видеоконференцсвязи (104) также может быть реализован на базе вычислительного устройства (400).

[0118] В общем случае вычислительное устройство (400) содержит объединенные общей шиной информационного обмена один или несколько процессоров (401), средства памяти, такие как ОЗУ (402) и ПЗУ (403), интерфейсы ввода/вывода (404), устройства ввода/вывода (405), и средство для сетевого взаимодействия (406).

[0119] Процессор (401) (или несколько процессоров, многоядерный процессор) могут выбираться из ассортимента устройств, широко применяемых в текущее время, например, компаний Intel™, AMD™, Apple™, Samsung Exynos™, MediaTek™, Qualcomm Snapdragon™ и т.п. В качестве процессора (401) может также применяться графический процессор, например, Nvidia, AMD, Graphcore и пр.

[0120] ОЗУ (402) представляет собой оперативную память и предназначено для хранения исполняемых процессором (401) машиночитаемых инструкций для выполнения необходимых операций по логической обработке данных. ОЗУ (402), как правило, содержит исполняемые инструкции операционной системы и соответствующих программных компонент (приложения, программные модули и т.п.).

[0121] ПЗУ (403) представляет собой одно или более устройств постоянного хранения данных, например, жесткий диск (HDD), твердотельный накопитель данных (SSD), флэш-память (EEPROM, NAND и т.п.), оптические носители информации (CD-R/RW, DVD-R/RW, BlueRay Disc, MD) и др.

[0122] Для организации работы компонентов устройства (400) и организации работы внешних подключаемых устройств применяются различные виды интерфейсов В/В (404). Выбор соответствующих интерфейсов зависит от конкретного исполнения вычислительного устройства, которые могут представлять собой, не ограничиваясь: PCI, AGP, PS/2, IrDa, FireWire, LPT, COM, SATA, IDE, Lightning, USB (2.0, 3.0, 3.1, micro, mini, type C), TRS/Audio jack (2.5, 3.5, 6.35), HDMI, DVI, VGA, Display Port, RJ45, RS232 и т.п.

[0123] Для обеспечения взаимодействия пользователя с вычислительным устройством (400) применяются различные устройства (405) В/В информации, например, клавиатура, дисплей (монитор), сенсорный дисплей, тач-пад, джойстик, манипулятор мышь, световое перо, стилус, сенсорная панель, трекбол, динамики, микрофон, средства дополненной реальности, оптические сенсоры, планшет, световые индикаторы, проектор, камера, средства биометрической идентификации (сканер сетчатки глаза, сканер отпечатков пальцев, модуль распознавания голоса) и т.п.

[0124] Средство сетевого взаимодействия (406) обеспечивает передачу данных устройством (400) посредством внутренней или внешней вычислительной сети, например, Интранет, Интернет, ЛВС и т.п. В качестве одного или более средств (406) может использоваться, но не ограничиваясь: Ethernet карта, GSM модем, GPRS модем, LTE модем, 5G модем, модуль спутниковой связи, NFC модуль, Bluetooth и/или BLE модуль, Wi-Fi модуль и др.

[0125] Дополнительно в составе устройства (400) могут также применяться средства спутниковой навигации, например, GPS, ГЛОНАСС, BeiDou, Galileo.

[0126] Представленные материалы заявки раскрывают предпочтительные примеры реализации технического решения и не должны трактоваться как ограничивающие иные, частные примеры его воплощения, не выходящие за пределы испрашиваемой правовой охраны, которые являются очевидными для специалистов соответствующей области техники.

## ФОРМУЛА

1. Система для видеоконференцсвязи, содержащая:

- по меньшей мере два клиентских устройства, выполненных с возможностью подключаться к модулю обмена сообщениями; и
- сервер видеоконференцсвязи, содержащий по меньшей мере модуль обмена сообщениями, модуль пересылки потоков и модуль распознавания потоков,

причем модуль обмена сообщениями выполнен с возможностью

- регистрировать подключения клиентских устройств к конференции,
- отправлять в единое информационное пространство по меньшей мере одно сообщение, полученное от по меньшей мере одного клиентского устройства,

модуль пересылки потоков выполнен с возможностью

- устанавливать защищенное соединение с клиентскими устройствами,
- принимать поток данных от по меньшей мере одного клиентского устройства,
- передавать поток данных в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков,
- принимать от модуля распознавания потоков по меньшей мере одно сгенерированное сообщение и отправлять его в единое информационное пространство,

модуль распознавания потоков выполнен с возможностью обрабатывать поток данных и генерировать по меньшей мере одно сообщение на основе обработанного потока данных, причем обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных модулем пересылки потоков по меньшей мере одному другому клиентскому устройству.

2. Система по п.1, характеризующаяся тем, что принятый от по меньшей мере одного клиентского устройства поток данных содержит по меньшей мере аудио и/или видео поток.

3. Система по п.2, характеризующаяся тем, что обработка аудио потока включает преобразование речи в текст (Automatic Speech Recognition).

4. Система по п.2, характеризующаяся тем, что обработка видео потока включает преобразование жестов в текстовую и/или графическую информацию (Gesture Recognition).

5. Система по п.2, характеризующаяся тем, что обработка видео потока включает распознавание языка жестов (Sign Language Recognition).

6. Система по любому из пп.1-5, характеризующаяся тем, что модуль распознавания потоков дополнительно выполнен с возможностью определять тип потока данных и

осуществлять соответствующую обработку потока данных в зависимости от типа потока данных.

7. Система по п.6, характеризующаяся тем, что сервер видеоконференцсвязи дополнительно содержит модуль фильтрации аудио потока, выполненный с возможностью подавления шума.

8. Система по п.1, характеризующаяся тем, что сообщения, сгенерированные на основе обработанного потока данных, содержат по меньшей мере текстовую и/или графическую информацию.

9. Система по любому из пп.1-5, характеризующаяся тем, что сервер видеоконференцсвязи дополнительно содержит модуль модификации видео потока, выполненный с возможностью

- замены фона,
- применения фильтров, улучшающих визуальные характеристики изображения пользователя,
- применения фильтров дополненной реальности (AR-фильтров).

10. Система по любому из пп.2-5, характеризующаяся тем, что модуль пересылки потоков дополнительно выполнен с возможностью принимать от по меньшей мере одного клиентского устройства команду на распознавание аудио и/или видео потока; и передавать аудио и/или видео поток модулю распознавания потоков после получения упомянутой команды на распознавание.

11. Система по п.2-3, характеризующаяся тем, что модуль распознавания потоков дополнительно выполнен с возможностью

- осуществлять разбиение аудио потока на фрагменты в процессе преобразования речи в текст,
- обрабатывать фрагменты, полученные в результате разбиения аудио потока,
- генерировать по меньшей мере одно сообщение на основе обработанных фрагментов;

сервер видеоконференцсвязи дополнительно содержит модуль согласования, выполненный с возможностью

- определять в режиме реального времени, осуществляет ли модуль распознавания потоков обработку аудио потока,
- устанавливать значение флага, идентифицирующего состояние текущего голосового ввода пользователя клиентского устройства,

- получать от модуля обмена сообщениями запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства,
  - отправлять модулю обмена сообщениями разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства на основании значения упомянутого флага, при этом
    - если значение флага идентифицирует отсутствие текущего голосового ввода пользователя клиентского устройства, модуль согласования отправляет модулю обмена сообщениями разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства;
    - если значение флага идентифицирует наличие текущего голосового ввода пользователя клиентского устройства, модуль согласования выполнен с возможностью
      - отправлять модулю пересылки потоков запрос на отправку в единое информационное пространство сгенерированного сообщения, содержащего уже обработанные фрагменты,
      - после отправки модулем пересылки потоков сгенерированного сообщения, содержащего уже обработанные фрагменты, в единое информационное пространство, отправлять модулю обмена сообщениями разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства;
- модуль обмена сообщениями дополнительно выполнен с возможностью
- отправлять модулю согласования запрос на разрешение отправки в единое информационное пространство сообщения от другого клиентского устройства,
  - отправлять в единое информационное пространство сообщения от другого клиентского устройства в ответ на прием разрешения отправки от модуля согласования.

12. Система по п.1, характеризующаяся тем, что сервер видеоконференцсвязи дополнительно содержит модуль управления единым информационным пространством, выполненный с возможностью определять, что по меньшей мере одно сгенерированное сообщение получено от модуля пересылки потоков; и отображать при выводе по меньшей мере одного сгенерированного сообщения метку, идентифицирующую формирование сгенерированного сообщения из аудио потока.

13. Система по п.1, характеризующаяся тем, что модуль управления единым информационным пространством дополнительно выполнен с возможностью

- сохранять сообщения, размещенные в упомянутом пространстве,
- обеспечивать доступ подключенным клиентским устройствам к сохраненным сообщениям.

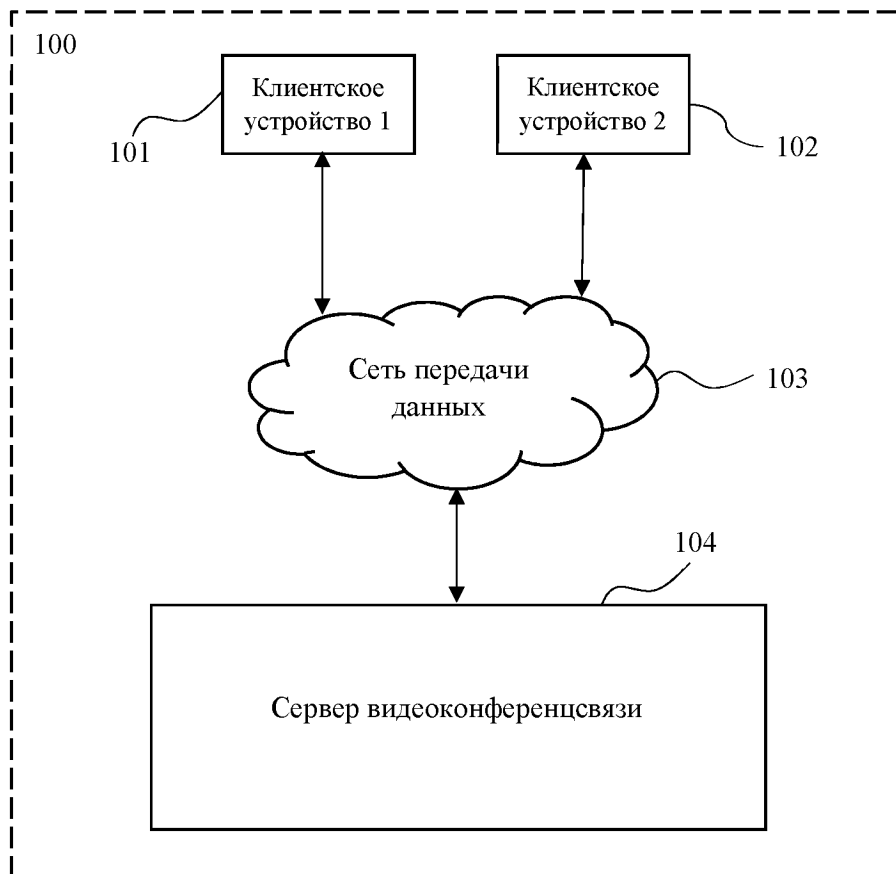
14. Система по п.1, характеризующаяся тем, что модуль управления единым информационным пространством дополнительно выполнен с возможностью осуществлять

- поиск по содержимому сообщений, размещенных в упомянутом пространстве,
- скачивание содержимого сообщений, размещенных в упомянутом пространстве,
- управление отображением сообщений, размещенных в упомянутом пространстве.

15. Система по п.1, характеризующаяся тем, что сервер видеоконференцсвязи дополнительно содержит модуль перевода, выполненный с возможностью перевода по меньшей мере одного сообщения, полученного от по меньшей мере одного клиентского устройства и/или сообщения, сгенерированного на основе обработанного потока данных.

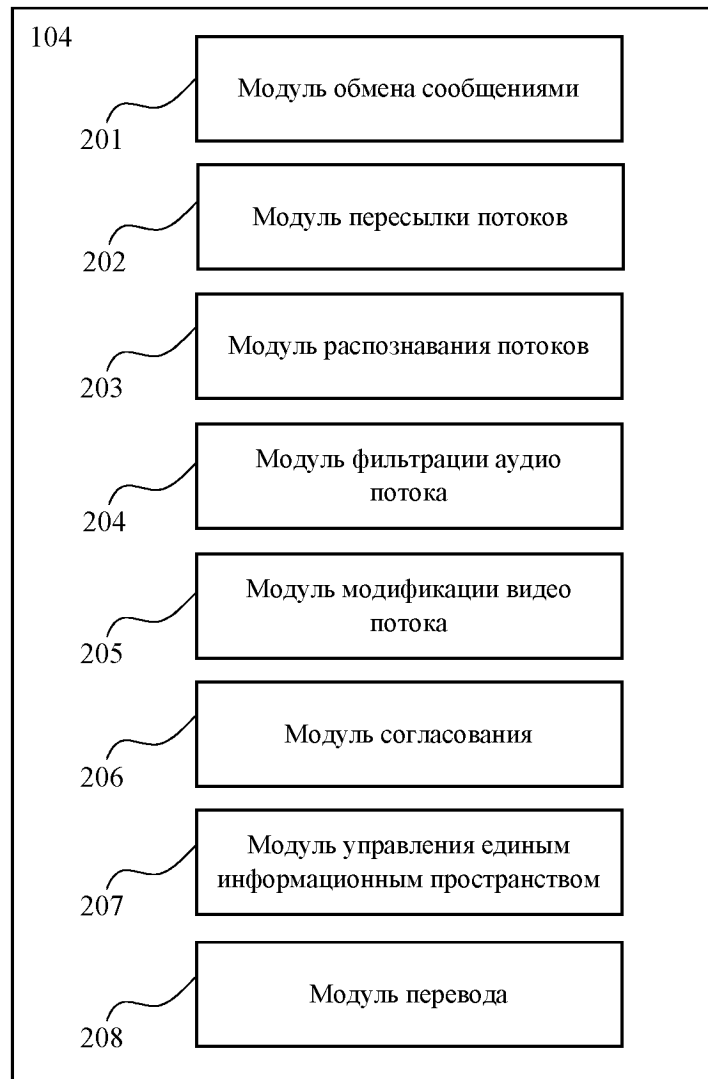
16. Способ осуществления видеоконференцсвязи, реализуемый с помощью сервера видеоконференцсвязи и содержащий этапы, на которых:

- подключают по меньшей мере два клиентских устройства к серверу видеоконференцсвязи,
- регистрируют подключения клиентских устройств к конференции,
- устанавливают защищенное соединение с клиентскими устройствами,
- принимают поток данных от по меньшей мере одного клиентского устройства,
- передают поток данных в параллельном режиме по меньшей мере одному другому клиентскому устройству и модулю распознавания потоков,
- в модуле распознавания потоков обрабатывают поток данных и генерируют по меньшей мере одно сообщение на основе обработанного потока данных, причем обработка потока данных и генерация по меньшей мере одного сообщения осуществляется параллельно с передачей потока данных по меньшей мере одному другому клиентскому устройству,
- принимают по меньшей мере одно сгенерированное сообщение и отправляют его в единое информационное пространство.



Фиг. 1

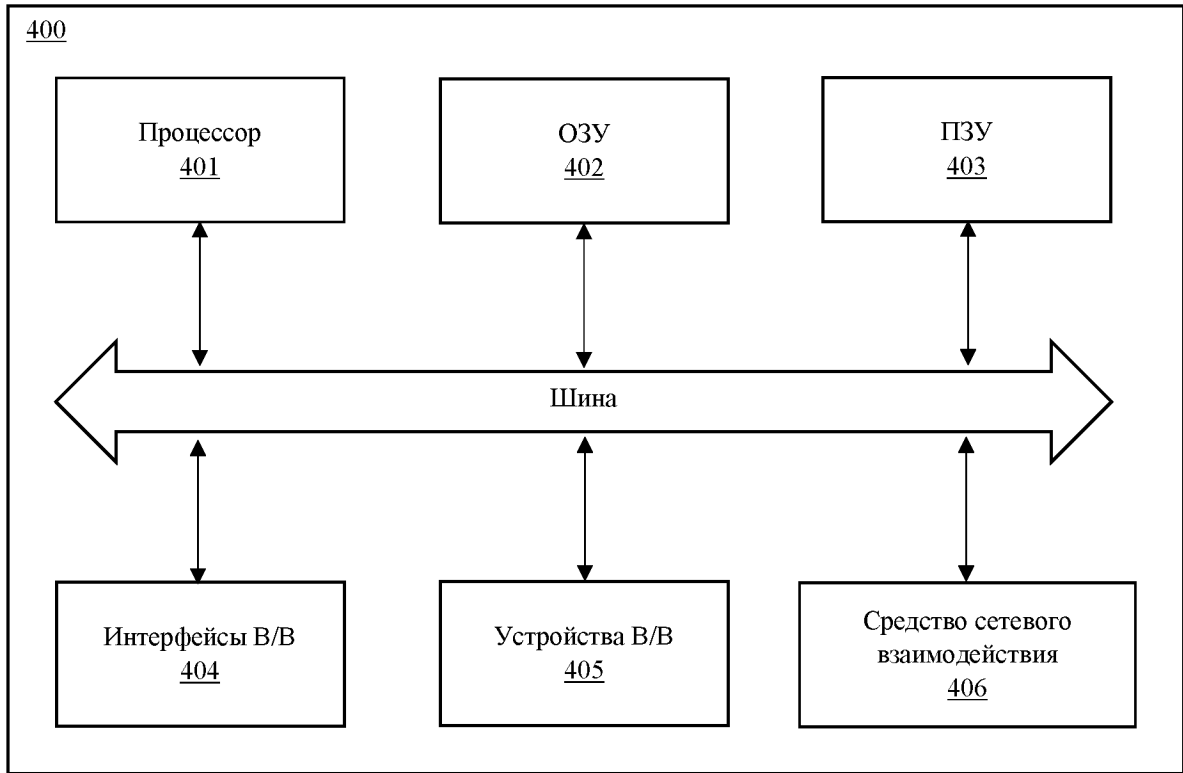




Фиг. 2



Фиг. 3



Фиг. 4

**ОТЧЕТ О ПАТЕНТНОМ ПОИСКЕ**  
(статья 15(3) ЕАПК и правило 42 Патентной инструкции к ЕАПК)

Номер евразийской заявки:  
**202291893**

**А. КЛАССИФИКАЦИЯ ПРЕДМЕТА ИЗОБРЕТЕНИЯ:**  
**H04N 21/4788 (2011.01)**  
**H04L 65/403 (2022.01)**

Согласно Международной патентной классификации (МПК)

**Б. ОБЛАСТЬ ПОИСКА:**  
Просмотренная документация (система классификации и индексы МПК)  
H04N 7/00, 7/14, 7/15, 21/00-21/60, H04L 12/00-12/18, 65/00-65/403

Электронная база данных, использовавшаяся при поиске (название базы и, если возможно, используемые поисковые термины)  
Espacenet, ЕАПАТИС, EPOQUE Net, Reaxys, Google

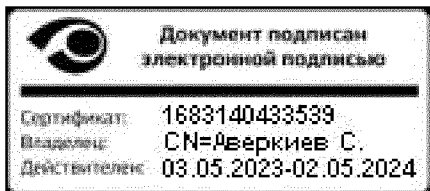
**В. ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ**

Категория*	Ссылки на документы с указанием, где это возможно, релевантных частей	Относится к пункту №
Y A	US 2020/0321007 A1 (SPEECH CLOUD INC) 08.10.2020, параграфы [0016], [0019]-[0021], [0038]-[0043], [0046]-[0048], [0058], [0064]-[0072], [0077]-[0080], фигуры 4, 6	1-10, 12-16 11
Y	US 2016/0352708 A1 (NAGRAVISION S.A.) 01.12.2016, параграфы [0017], [0023], [0027]-[0029], [0031], [0038], [0047], [0081]-[0082]	1-10, 12-16
Y	US 10139917 B1 (GOOGLE LLC) 27.11.2018, колонка 8, строка 49-колонка 9, строка 5, колонка 4, строка 60-колонка 5, строка 40	4, 6, 7, 9, 10
Y	SHURUTHI V. et al. Enabling Sign Language Recognition Feature in Video Conferencing. Springer, Lecture Notes in Networks and Systems, том 391, 2021, страницы 1227-1235 [онлайн] [найдено 2023-04-05]. Найдено в < <a href="https://doi.org/10.1007/978-3-030-94277-9_104">https://doi.org/10.1007/978-3-030-94277-9_104</a> >	5, 6, 7, 9, 10
Y	US 2021/0392231 A1 (TRUE MEETING INC.) 16.12.2021, параграфы [0213], [0214], [0361]-[0362], [0753]-[0756], [0797], [0847]-[0849]	7, 9, 15
Y	US 2011/0313754 A1 (INTERNATIONAL BUSINESS MACHINES CORPORATION) 22.12.2011, параграфы [0018]-[0019], [0022]-[0023]	10
A	ARKHANGORODSKY Arkady et al. MeetDot: Videoconferencing with Live Translation Captions. Conference on Empirical Methods in Natural Language Processing: System Demonstrations, November 2021, pages 195-201, [онлайн] [найдено 2023-04-05]. Найдено в < <a href="https://www.researchgate.net/publication/354779095_MeetDot_Videoconferencing_with_Live_Translation_Captions">https://www.researchgate.net/publication/354779095_MeetDot_Videoconferencing_with_Live_Translation_Captions</a> >	1-16

последующие документы указаны в продолжении

* Особые категории ссылочных документов: «А» - документ, определяющий общий уровень техники «D» - документ, приведенный в евразийской заявке «E» - более ранний документ, но опубликованный на дату подачи евразийской заявки или после нее «O» - документ, относящийся к устному раскрытию, экспонированию и т.д. "P" - документ, опубликованный до даты подачи евразийской заявки, но после даты испрашиваемого приоритета"	«Т» - более поздний документ, опубликованный после даты приоритета и приведенный для понимания изобретения «Х» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий новизну или изобретательский уровень, взятый в отдельности «У» - документ, имеющий наиболее близкое отношение к предмету поиска, порочащий изобретательский уровень в сочетании с другими документами той же категории «&» - документ, являющийся патентом-аналогом «L» - документ, приведенный в других целях
--	---

Дата проведения патентного поиска: 04 мая 2023 (04.05.2023)

Уполномоченное лицо: Начальник Управления экспертизы	 <p>Документ подписан электронной подписью</p> <p>Сертификат: 1683140433539 Владелец: СN=Аверкиев С. Действителен: 03.05.2023-02.05.2024</p>	С.Е. Аверкиев
---	--	---------------

**ОТЧЕТ О ПАТЕНТНОМ ПОИСКЕ**  
**(дополнительный лист)**

Номер евразийской заявки:

**202291893**

**ДОКУМЕНТЫ, СЧИТАЮЩИЕСЯ РЕЛЕВАНТНЫМИ (продолжение графы В)**

Категория*	Ссылки на документы с указанием, где это возможно, релевантных частей	Относится к пункту №
А	US 2021/0044635 A1 (MITEЛ NETWORKS (INTERNATIONAL) LIMITED) 11.02.2021	1-16