



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК

G06F 21/32 (2020.02); G06K 9/00335 (2020.02); G10L 17/06 (2020.02)

(21)(22) Заявка: 2018144787, 12.04.2017

(24) Дата начала отсчета срока действия патента:
12.04.2017

Дата регистрации:
11.12.2020

Приоритет(ы):

(30) Конвенционный приоритет:
19.05.2016 CN 201610340549.4

(43) Дата публикации заявки: 19.06.2020 Бюл. № 17

(45) Опубликовано: 11.12.2020 Бюл. № 35

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 19.12.2018

(86) Заявка РСТ:
CN 2017/080196 (12.04.2017)

(87) Публикация заявки РСТ:
WO 2017/198014 (23.11.2017)

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

(72) Автор(ы):

ЛИ, Пэн (CN),
СУНЬ, Ипэн (CN),
СЕ, Юнсян (CN),
ЛИ, Лян (CN)

(73) Патентообладатель(и):

АЛИБАБА ГРУП ХОЛДИНГ ЛИМИТЕД
(КУ)

(56) Список документов, цитированных в отчете
о поиске: US 2013/0267204 A1, 10.10.2013. CN
104598796 A, 06.05.2015. CN 105426723 A,
23.03.2016. US 2015/0199502 A1, 16.07.2015. RU
2543958 C2, 10.03.2015.

(54) СПОСОБ И УСТРОЙСТВО АУТЕНТИФИКАЦИИ ЛИЧНОСТИ

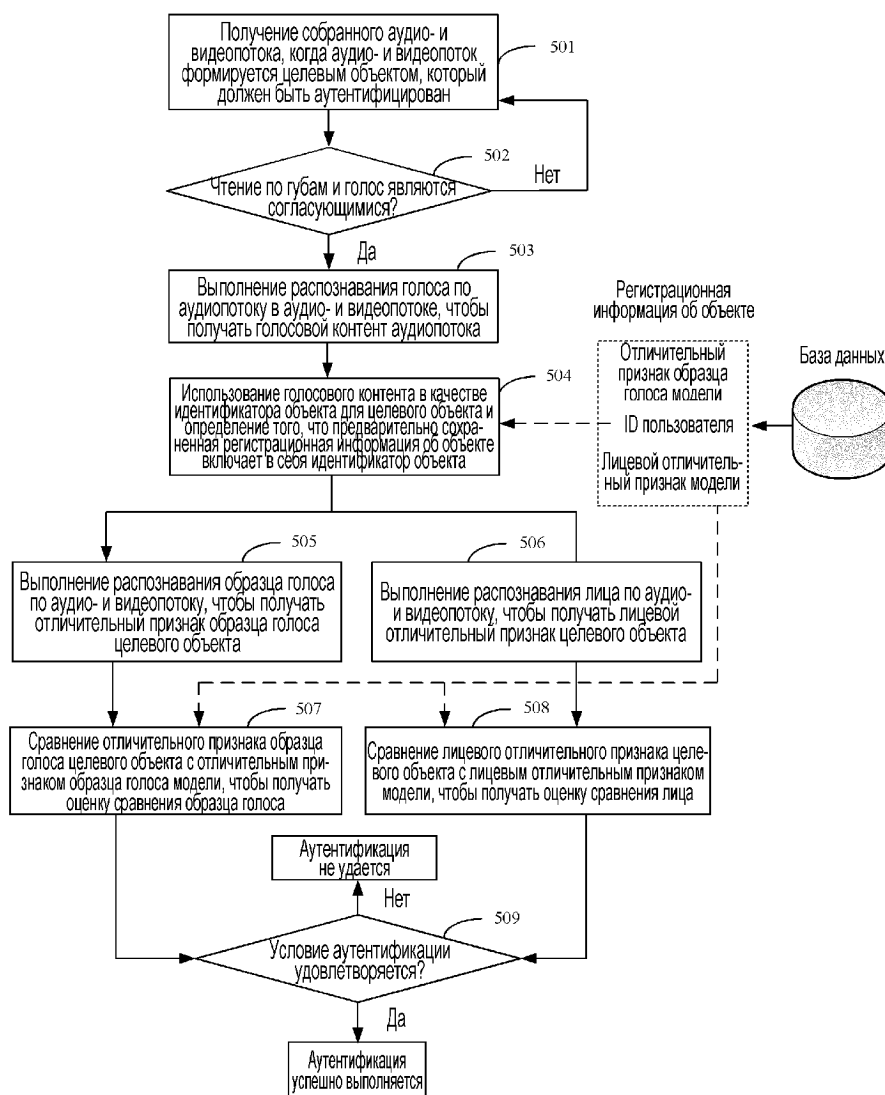
(57) Реферат:

Изобретение относится к вычислительной технике. Технический результат заключается в повышении эффективности и надежности аутентификации личности. Способ аутентификации личности, в котором получают собранный аудио- и видеопоток, когда аудио- и видеопоток формируется целевым объектом, который должен быть аутентифицирован; определяют, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, и если чтение по губам и голос являются согласующимися, получают идентификатор

объекта для целевого объекта путем использования голосового контента; определяют, содержит ли предварительно сохраненная регистрационная информация идентификатор объекта; получают физиологический признак модели, соответствующий идентификатору объекта, из регистрационной информации об объекте, если предварительно сохраненная регистрационная информация об объекте содержит идентификатор объекта; выполняют физиологическое распознавание по аудио- и видеопотоку, чтобы получать физиологический

признак целевого объекта; и сравнивают физиологический признак целевого объекта с физиологическим признаком модели, чтобы получить результат сравнения, и если результат

сравнения удовлетворяет условию аутентификации, определяют, что целевой объект был аутентифицирован. 2 н. и 8 з.п. ф-лы, 7 ил.



ФИГ. 5



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC

G06F 21/32 (2020.02); *G06K 9/00335* (2020.02); *G10L 17/06* (2020.02)(21)(22) Application: **2018144787, 12.04.2017**(24) Effective date for property rights:
12.04.2017Registration date:
11.12.2020

Priority:

(30) Convention priority:
19.05.2016 CN 201610340549.4(43) Application published: **19.06.2020 Bull. № 17**(45) Date of publication: **11.12.2020 Bull. № 35**(85) Commencement of national phase: **19.12.2018**(86) PCT application:
CN 2017/080196 (12.04.2017)(87) PCT publication:
WO 2017/198014 (23.11.2017)Mail address:
**129090, Moskva, ul. B. Spasskaya, 25, str. 3, OOO
"Yuridicheskaya firma Gorodisskij i Partnery"**

(72) Inventor(s):

**LI, Peng (CN),
SUN, Yipeng (CN),
XIE, Yongxiang (CN),
LI, Liang (CN)**

(73) Proprietor(s):

Alibaba Group Holding Limited (KY)(54) **METHOD AND DEVICE FOR AUTHENTICATING AN INDIVIDUAL**

(57) Abstract:

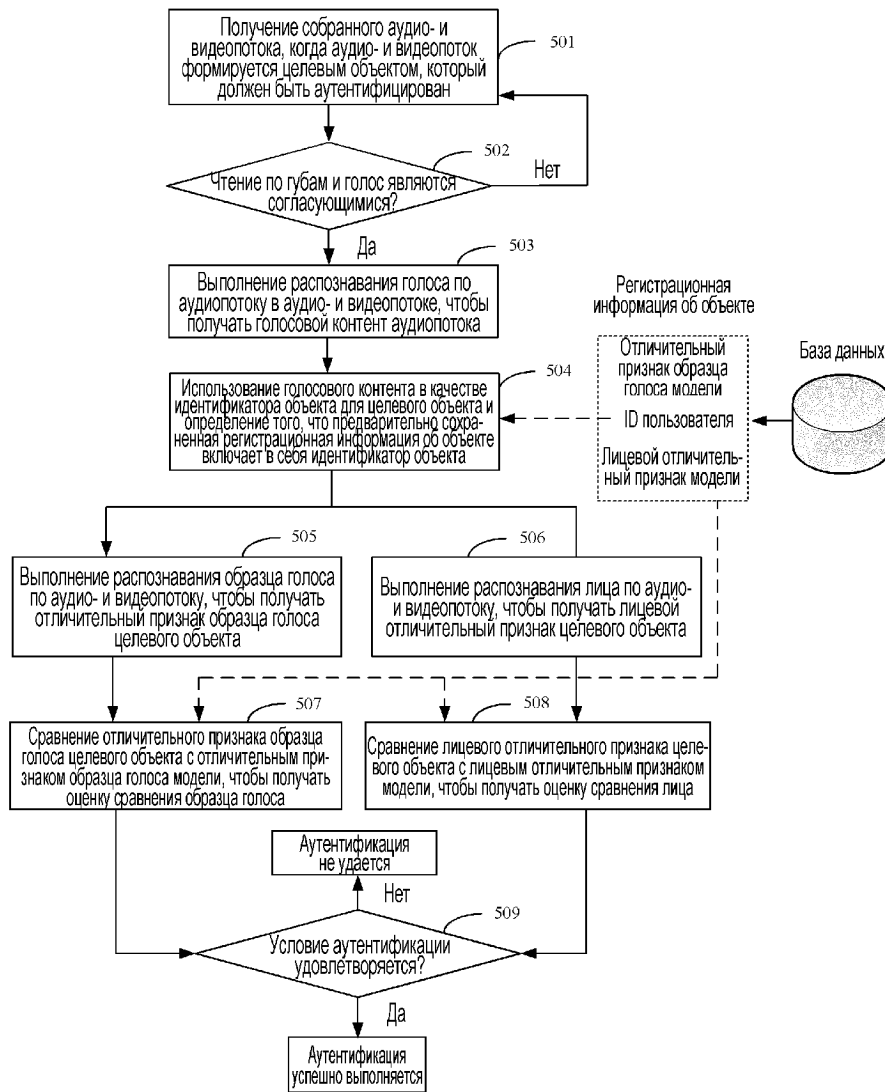
FIELD: computer equipment.

SUBSTANCE: method of authenticating a person, in which a collected audio and video stream is received, when an audio and video stream is formed by a target object to be authenticated; determining whether lip-matching and voice are consistent in the audio and video stream, and if lip-reading and voice are consistent, obtaining object identifier for target object by using voice content; determining whether pre-stored registration information comprises an object identifier; obtaining a physiological feature of the model corresponding to the object identifier from the registration information about the object, if previously

stored object registration information contains object identifier; performing physiological recognition by audio and video stream, to obtain a target object physiological feature; and comparing the physiological attribute of the target object with the physiological feature of the model to obtain a comparison result, and if the comparison result satisfies the authentication condition, it is determined that the target object has been authenticated.

EFFECT: technical result consists in improvement of efficiency and reliability of identity authentication.

10 cl, 7 dwg



ФИГ. 5

Область техники, к которой относится изобретение

[0001] Настоящая заявка относится к сетевым технологиям и, в частности, к способу и устройству аутентификации личности.

Уровень техники

5 [0002] С развитием Интернет-технологий сетевые службы все более широко используются людьми. Например, люди могут отправлять и принимать электронные письма с помощью почтового ящика, осуществлять онлайн-покупки или даже работать в режиме онлайн. Некоторые приложения имеют высокие требования безопасности, и личность пользователя должна быть идентифицирована. Например, личность
10 пользователя должна быть аутентифицирована, прежде чем оплата авторизуется для онлайн-покупки; или пользователь может регистрироваться в приложении с относительно высоким требованием безопасности, после того как личность пользователя была аутентифицирована. В технологиях предшествующего уровня существуют несколько способов аутентификации личности, используемых посредством Интернета, таких как распознавание лица и распознавание образца голоса. Однако, эти повсеместно
15 используемые способы аутентификации являются относительно сложными. Например, пользователю необходимо вводить ID пользователя и затем проверять образец голоса для распознавания образца голоса. Кроме того, существующие способы аутентификации имеют относительно низкую надежность. Например, злоумышленник может
20 осуществлять распознавание лица с помощью аналогового видео или записи. Даже если проверка выполняется на основе двух объединенных способов аутентификации, например, со ссылкой на лицо и образец голоса, поскольку эти способы аутентификации являются относительно независимыми друг от друга, злоумышленник может прорываться сквозь процессы аутентификации. Ранее описанные недостатки способов
25 аутентификации личности могут создавать некоторые риски для безопасности приложения.

Сущность изобретения

[0003] Принимая во внимание вышеописанное, настоящая заявка предоставляет способ и устройство аутентификации личности, с тем, чтобы улучшить эффективность
30 и надежность аутентификации личности.

[0004] В частности, настоящая заявка реализуется с помощью следующих технических решений.

[0005] Согласно первому аспекту предоставляется способ аутентификации личности, и способ включает в себя получение собранного аудио- и видеопотока, когда аудио- и
35 видеопоток формируется целевым объектом, который должен быть аутентифицирован; определение того, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, и если чтение по губам и голос являются согласующимися, использование голосового контента, полученного посредством выполнения распознавания голоса по аудиопотоку для аудио- и видеопотока, в качестве идентификатора объекта для целевого
40 объекта; получение физиологического признака модели, соответствующего идентификатору объекта, из регистрационной информации об объекте, если предварительно сохраненная регистрационная информация об объекте включает в себя идентификатор объекта; выполнение физиологического распознавания по аудио- и видеопотоку, чтобы получать физиологический признак целевого объекта; и сравнение
45 физиологического признака целевого объекта с физиологическим признаком модели, чтобы получать результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определение того, что целевой объект был аутентифицирован.

[0006] Согласно второму аспекту предоставляется устройство аутентификации

личности, и устройство включает в себя модуль получения информации, сконфигурированный, чтобы получать собранный аудио- и видеопоток, когда аудио- и видеопоток формируется целевым объектом, который должен быть аутентифицирован; модуль определения идентификатора, сконфигурированный, чтобы определять, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, и если чтение по губам и голос являются согласующимися, использовать голосовой контент, полученный посредством выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке, в качестве идентификатора объекта для целевого объекта; модуль управления информацией, сконфигурированный, чтобы получать физиологический признак модели, соответствующий идентификатору объекта, из регистрационной информации об объекте, если предварительно сохраненная регистрационная информация об объекте включает в себя идентификатор объекта; модуль распознавания отличительного признака, сконфигурированный, чтобы выполнять физиологическое распознавание по аудио- и видеопотоку, чтобы получать физиологический признак целевого объекта; и модуль обработки аутентификации, сконфигурированный, чтобы сравнивать физиологический признак целевого объекта с физиологическим признаком модели, чтобы получать результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определять, что целевой объект был аутентифицирован.

[0007] Согласно способу и устройству аутентификации личности, предоставленным в настоящей заявке, идентификатор пользователя получается посредством распознавания аудио- и видеопотока во время аутентификации пользователя, и лицевой отличительный признак и отличительный признак образца голоса могут также быть проверены с помощью того же аудио- и видеопотока. Это упрощает пользовательскую операцию, улучшает эффективность аутентификации, поддерживает модель аутентификации один к одному и гарантирует точность распознавания. Кроме того, в способе, согласованность между чтением по губам и голосом определяется, чтобы гарантировать, что целевой объект является живым объектом вместо поддельной видеозаписи злоумышленника, тем самым, улучшая безопасность и надежность аутентификации.

Краткое описание чертежей

[0008] Фиг. 1 иллюстрирует процедуру регистрации личности согласно примерной реализации настоящей заявки;

[0009] Фиг. 2 иллюстрирует процедуру определения согласованности между чтением по губам и голосом, согласно примерной реализации настоящей заявки;

[0010] Фиг. 3 иллюстрирует процедуру распознавания лицевого отличительного признака, согласно примерной реализации настоящей заявки;

[0011] Фиг. 4 иллюстрирует процедуру распознавания отличительного признака образца голоса, согласно примерной реализации настоящей заявки;

[0012] Фиг. 5 иллюстрирует процедуру аутентификации личности согласно примерной реализации настоящей заявки;

[0013] Фиг. 6 является структурной схемой устройства аутентификации личности согласно примерной реализации настоящей заявки; и

[0014] Фиг. 7 является структурной схемой устройства аутентификации личности согласно примерной реализации настоящей заявки.

Описание вариантов осуществления

[0015] Примерные реализации описываются подробно в настоящем документе, и примеры примерных реализаций представляются на сопровождающих чертежах. Когда последующее описание соотносится с сопровождающими чертежами, пока не указано

иное, одинаковые номера на различных сопровождающих чертежах представляют одинаковые или аналогичные элементы. Реализации, описанные в последующих примерных реализациях, не представляют все реализации, согласующиеся с настоящей заявкой. Вместо этого, они являются лишь примерами устройств и способов, согласующихся с некоторыми аспектами настоящей заявки, которые описываются подробно в прилагаемой формуле изобретения.

[0016] Способ аутентификации личности, предоставленный в реализации настоящей заявки, может быть применен к Интернет-аутентификации личности. Например, пользователь может входить в сетевое приложение после аутентификации с помощью способа, с тем, чтобы гарантировать безопасность приложения.

[0017] Последующее использует приложение с относительно высоким требованием безопасности в качестве примера. Предположим, что приложение может работать на пользовательском интеллектуальном устройстве, например, смартфоне или интеллектуальном планшете. Когда пользователь должен входить в приложение на интеллектуальном устройстве, аудио- и видеопоток может быть собран с помощью камеры и микрофона на интеллектуальном устройстве. Например, пользователь может читать ID приложения для пользователя в камеру и микрофон мобильного телефона. ID приложения может быть учетным номером "123456", который регистрируется пользователем в приложении. После того как пользователь читает ID приложения, мобильный телефон может собирать аудио- и видеопоток пользователя, в том числе видеоизображение и считывание голоса.

[0018] Согласно способу аутентификации личности в этой реализации настоящей заявки собранный аудио- и видеопоток может быть обработан. Перед аутентификацией пользователю необходимо выполнять процедуру регистрации личности, чтобы выполнять аутентификацию личности позже. Процедура регистрации также выполняется на основе собранного аудио- и видеопотока. Последующее отдельно описывает процедуру регистрации личности и процедуру аутентификации личности. Кроме того, эта реализация не накладывает ограничение на устройство, которое выполняет регистрацию личности или обработку аутентификации в фактических приложениях. Например, после того как аудио- и видеопоток пользователя собирается, смартфон может передавать аудио- и видеопоток внутреннему серверу приложения, или часть обработки может быть выполнена на стороне клиентского программного обеспечения смартфона, а другие части обработки могут быть выполнены на стороне сервера; или могут быть использованы другие способы.

Регистрация личности

[0019] В способе в этой реализации, когда пользователь выполняет регистрацию личности, могут содержаться два типа информации. Один тип информации является идентификатором объекта. Например, когда пользователь входит в приложение, пользователь может упоминаться как целевой объект. Когда пользователь регистрируется в приложении, информация, используемая для различения пользователя от другого пользователя, является идентификатором объекта в приложении, например, может быть учетным номером 123456 пользователя в приложении, и учетный номер 123456 является идентификатором объекта для целевого объекта. Другой тип информации является физиологической информацией, которая может уникально идентифицировать пользователя, например, отличительным признаком образца голоса пользователя или лицевым отличительным признаком пользователя. Обычно, образцы голоса и лица различных людей являются различными, и физиологическая информация, которая идентифицирует каждого пользователя, может называться физиологическим

признаком модели.

[0020] Сопоставляющее соотношение устанавливается между двумя типами информации: идентификатором объекта и физиологическим признаком модели, и сопоставляющее соотношение сохраняется. Соответственно сохраненный идентификатор объекта и физиологический признак модели целевого объекта могут называться "регистрационной информацией об объекте". Например, пользователь Сяо Жэнг может сохранять регистрационную информацию об объекте для пользователя как "123456 - физиологический признак А модели". Для более точной идентификации пользователя физиологическая информация, включенная в физиологический признак модели, используемый в этом примере, может быть, по меньшей мере, двумя типами физиологической информации, например, лицом и образцом голоса.

[0021] Фиг. 1 иллюстрирует примерную процедуру регистрации личности. Процедура включает в себя следующую обработку:

[0022] Этап 101: Получение аудио- и видеопотока, который должен быть зарегистрирован для целевого объекта.

[0023] Например, пользователь регистрируется в приложении, и пользователь может читать учетный номер "123456" приложения в интеллектуальное устройство пользователя, такое как мобильный телефон. В этом примере пользователь, который выполняет регистрацию, может называться целевым объектом; и камера и микрофон интеллектуального устройства могут собирать аудио- и видеопоток, сформированный, когда пользователь читает учетный номер. Аудио- и видеопоток, собранный во время регистрации, может называться аудио- и видеопотоком, который должен быть зарегистрирован, включающим в себя аудиопоток и видеопоток. Аудиопоток является считыванием голоса пользователя, а видеопоток является считыванием видеоизображения пользователя.

[0024] После того как аудио- и видеопоток на настоящем этапе получен, три последующих аспекта обработки могут быть выполнены, чтобы завершать регистрацию пользователя. Более подробно, ссылки могут все еще выполняться на фиг. 1.

[0025] В одном аспекте обработка является следующей: Этап 102: Определение того, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, который должен быть зарегистрирован.

[0026] Согласованность здесь указывает сопоставляющее соотношение между движением губ и движением, указываемым голосом. Например, предположим, что голос является фразой "погода сегодня солнечная", когда голос формируется с низкой скоростью, но движение губ соответствует фразе "погода сегодня солнечная", формируемой с высокой скоростью. Очевидно, что голос и движение губ не соответствуют друг другу: когда движение губ остановилось (контент полностью считан), голос продолжается (... сегодня солнечная). Это может происходить во множестве возможных ситуаций, таких как, когда злоумышленник пытается пройти определение ID пользователя и лица, злоумышленник может атаковать определение лица с помощью предыдущей видеозаписи пользователя (атакованного пользователя), и злоумышленник считывает ID пользователя, чтобы атаковать распознавание ID голосового контента. По существу, злоумышленник может аутентифицироваться. Однако, при такой атаке, чтение по губам и голос обычно являются несогласующимися, и может быть определено, что читающий не является пользователем.

[0027] Как показано на этапе 102, если определяется, что чтение по губам и голос в аудио- и видеопотоке, который должен быть сформирован, являются несогласующимися, неудача регистрации может быть сообщена пользователю. Альтернативно, как показано

на фиг. 1, выполняется переход к этапу 101, чтобы снова собирать аудио- и видеопоток, чтобы предотвращать ошибку обработки.

[0028] Иначе, если определяется, что чтение по губам и голос в аудио- и видеопотоке, который должен быть зарегистрирован, являются согласующимися, выполняется этап 103: Использование голосового контента, полученного посредством выполнения распознавания голоса по аудиопотоку в собранном аудио- и видеопотоке, в качестве идентификатора объекта для целевого объекта. Распознавание голоса использует компьютерные технологии, чтобы автоматически распознавать голосовой контент человека, т.е., процесс распознавания для преобразования голоса в контент. Например, после того как распознавание голоса выполняется по аудиопотоку чтения "123456" пользователем, который должен выполнять регистрацию, полученным голосовым контентом аудиопотока является "123456", и контент, полученный посредством распознавания, может быть использован в качестве идентификатора пользователя, а именно, ID пользователя.

[0029] Предыдущее распознавание голоса аудиопотока может быть выполнено, после того как определяется, что чтение по губам и голос являются согласующимися, чтобы получать идентификатор объекта, или может быть выполнено в процессе определения того, являются ли согласующимися чтение по губам и голос, чтобы получать идентификатор объекта.

[0030] В другом аспекте физиологическое распознавание выполняется по аудио- и видеопотоку, который должен быть зарегистрирован, чтобы получать физиологический признак модели аудио- и видеопотока, который должен быть зарегистрирован. В этом примере физиологический признак включает в себя лицевой отличительный признак и отличительный признак образца голоса, но не ограничивается двумя отличительными признаками, при условии, что примерный отличительный признак может уникально идентифицировать пользователя, и может быть использован для различения между физиологическими признаками различных пользователей. В этом аспекте, как показано на этапе 104, распознавание образца голоса может быть выполнено по аудиопотоку в аудио- и видеопотоке, который должен быть зарегистрирован, чтобы получать отличительный признак образца голоса целевого объекта.

[0031] В еще одном аспекте обнаружение лица выполняется по видеопотоку в аудио- и видеопотоке, который должен быть зарегистрирован, чтобы получать лицевой отличительный признак целевого объекта.

[0032] В процедуре регистрации обнаруженный лицевой отличительный признак может называться лицевым отличительным признаком модели и использоваться в качестве критерия в последующем процессе аутентификации. Аналогично, обнаруженный отличительный признак образца голоса может также называться отличительным признаком образца голоса модели, и отличительный признак образца голоса модели и лицевой отличительный признак модели могут совместно называться физиологическим признаком модели.

[0033] В этой реализации физиологический признак модели и идентификатор объекта для целевого объекта также называются регистрационной информацией об объекте. После того как определяется, что данные в регистрационной информации об объекте являются полными, на этапе 106, идентификатор объекта для целевого объекта и соответствующий физиологический признак модели сохраняются в базе данных в качестве регистрационной информации об объекте.

[0034] Кроме того, порядок выполнения трех аспектов, показанный на фиг. 1, не ограничивается. Например, после того как аудио- и видеопоток, который должен быть

зарегистрирован, получается на этапе 101, три аспекта могут выполняться параллельно. Если чтение по губам и голос являются несогласующимися, распознанный отличительный признак образца голоса и распознанный лицевой отличительный признак могут не сохраняться. Альтернативно, после того как определяется, что чтение по губам и голос являются согласующимися, обнаружение и распознавание затем выполняются по отличительному признаку образца голоса и лицевому отличительному признаку.

[0035] Фиг. 2 иллюстрирует процедуру определения согласованности между чтением по губам и голосом на фиг. 1. Процедура может включать в себя следующие этапы:

[0036] Этап 201: Выполнение обнаружения конечной точки на основе аудиопотока в аудио- и видеопотоке, который должен быть зарегистрирован. На текущем этапе время начала и время окончания аудиопотока могут быть обнаружены в непрерывном аудиопотоке.

[0037] Этап 202: Извлечение постоянных голосовых отличительных признаков на основе аудиопотока, когда отличительные признаки включают в себя, но не только, MFCC-признак и LPCC-признак. Отличительные признаки, извлеченные на текущем этапе, могут быть использованы для распознавания голоса.

[0038] Этап 203: Идентификация слога голоса и соответствующего момента времени в аудиопотоке. На текущем этапе каждый слог в аудиопотоке может быть идентифицирован на основе голосовых отличительных признаков, извлеченных на этапе 202, и соответствующий момент времени появления и соответствующий момент времени исчезновения слога могут быть определены. Способ распознавания голоса включает в себя, но не только, способы, такие как скрытая Марковская модель (HMM), глубокая нейронная сеть (DNN) и долгая краткосрочная модель (LSTM).

[0039] Этап 204: Обнаружение местоположений губ на основе видеопотока в аудио- и видеопотоке, который должен быть зарегистрирован. На текущем этапе местоположения губ могут быть обнаружены из видеоизображения.

[0040] Этап 205: Определение качества обнаруженного изображения губ. Например, параметры, такие как четкость и экспозиция местоположений губ, могут быть определены. Если четкость является низкой, или степень экспозиции является слишком высокой, определяется, что изображение является неответчающим требованиям; и в этом случае сбор аудио- и видеопотока, который должен быть зарегистрирован, выполняется снова. Если изображение отвечает требованиям, выполняется переход к этапу 206, чтобы выполнять распознавание чтения по губам.

[0041] Этап 206: Извлечение постоянных отличительных признаков губ. На текущем этапе отличительный признак может быть извлечен из непрерывных изображений губ, и отличительный признак включает в себя, но не только, локальный описатель изображения, такой как пустой пиксел, LBP, Gabor, SIFT или Surf.

[0042] Этап 207: Идентификация слога чтения по губам и соответствующего момента времени в видеопотоке. На текущем этапе слог чтения по губам может быть идентифицирован с помощью способов, таких как скрытая Марковская модель (HMM) и долгой краткосрочной модели. Момент времени, соответствующий слогу чтения по губам во временной последовательности видео, также определяется с помощью модели во время распознавания чтения по губам.

[0043] Этап 208: Определение того, являются ли слог чтения по губам и слог голоса согласующимися с соответствующим моментом времени. Например, на текущем этапе, информация о моменте времени для слога голоса может быть сравнена с информацией о моменте времени для слога чтения по губам. Если результат сравнения указывает

согласованность, считается, что аудиопоток формируется реальным человеком; в этом случае выполняется переход к этапу 209. Если результат сравнения указывает на несогласованность, предполагается, что аудиопоток является атакующим поведением; и в этом случае выполняется возврат к процедуре регистрации. В этой реализации способ обнаружения согласованности между слогом чтения по губам и слогом голоса и соответствующим моментом времени является более детализированным, и, следовательно, существует более высокая точность в определении реального голоса человека.

[0044] Этап 209: Выполнение распознавания голоса по отличительным признакам голоса, извлеченным на этапе 202, чтобы получать ID пользователя, а именно, идентификатор объекта. Способ распознавания голоса включает в себя, но не только, способы, такие как скрытая Марковская модель (HMM), глубокая нейронная сеть (DNN) и долгая краткосрочная модель (LSTM).

[0045] Кроме того, в примере, показанном на фиг. 2, распознавание голоса для аудиопотока может быть выполнено на этапе 209, после того как определяется, что чтение по губам и голос являются согласующимися. Альтернативно, когда момент времени слога в аудиопотоке идентифицируется на этапе 203, ID пользователя получается посредством выполнения распознавания голоса на основе отличительного признака голоса. В этом случае, после того как определяется, что чтение по губам и голос являются согласующимися, на этапе 208, ID пользователя, полученный посредством распознавания, может быть непосредственно использован в качестве идентификатора объекта.

[0046] Фиг. 3 иллюстрирует процедуру распознавания лицевого отличительного признака на фиг. 1. Процедура может включать в себя следующие этапы:

[0047] Этап 301: Обнаружение изображений лица на основе видеопотока в аудио- и видеопотоке, который должен быть зарегистрирован. На текущем этапе изображение видеокadra может быть извлечено из видеопотока в аудио- и видеопотоке, и может быть обнаружено, появляется ли лицо в изображении видеокadra. Если да, выполняется переход к этапу 302. Иначе, выполняется возврат к процедуре определения.

[0048] Этап 302: Обнаружение качества изображений лица. На текущем этапе обнаружение точки лицевого отличительного признака может быть выполнено на лице, обнаруженном на этапе 301, и углы лица в горизонтальном направлении и в вертикальном направлении, могут быть определены на основе результата обнаружения точки отличительного признака. Если оба угла не превышают некоторые углы наклона, требование к качеству удовлетворяется. Иначе, требование к качеству не удовлетворяется. Кроме того, четкость, экспозиция и т.д. области лица определяются, которые также должны быть в пределах некоторых пороговых значений. Если изображения лица имеют хорошее качество, лицевой отличительный признак может быть лучше распознан.

[0049] Этап 303: Для изображений лица, которые удовлетворяют требованию по качеству, выполняется извлечение собственных векторов из изображений лица, где собственные векторы включают в себя, но не только, локальный бинарный шаблон (LBP), Gabor, сверточную нейронную сеть (CNN), и т.д.

[0050] Этап 304: Объединение множества лицевых собственных факторов, извлеченных на этапе 303, чтобы формировать уникальный лицевой отличительный признак пользователя, а именно, лицевой отличительный признак модели.

[0051] Фиг. 4 иллюстрирует процедуру распознавания отличительного признака образца голоса на фиг. 1. Процедура может включать в себя следующие этапы.

[0052] Этап 401: Получение аудиопотока в аудио- и видеопотоке, который должен быть зарегистрирован.

[0053] В этом примере распознавание отличительного признака образца голоса может быть выполнено на основе аудиопотока в аудио- и видеопотоке, который должен
5 быть зарегистрирован.

[0054] Этап 402: Определение того, что качество звука аудиопотока удовлетворяет критерию качества.

[0055] На текущем этапе может быть определено качество звука. Лучшее качество собранного аудиопотока ведет к лучшему результату выполнения распознавания
10 образца голоса по звуку. Следовательно, прежде чем выполняется последующее распознавание образца голоса, сначала может быть определено качество аудиопотока. Например, информация, такая как сила сигнала и соотношение сигнал-шум для голоса в аудиопотоке, может быть вычислена, чтобы определять, удовлетворяет ли голос критерию качества. Например, критерием качества может быть то, что соотношение
15 сигнал-шум попадает в некоторый диапазон, или может быть то, что сила сигнала голоса больше порогового значения силы. Если аудиопоток соответствует требованиям, выполняется переход к этапу 403. Иначе, сбор аудио- и видеопотока, который должен быть зарегистрирован, выполняется снова.

[0056] Этап 403: Извлечение собственных векторов образца голоса из аудиопотока.

[0057] В этом примере может быть множество аудио- и видеопотоков, которые
20 должны быть зарегистрированы. Например, пользователь может читать ID пользователя дважды, и, следовательно, два аудио- и видеопотока собираются. На текущем этапе собственный вектор образца голоса аудиопотока в каждом аудио- и видеопотоке может быть извлечен. Собственный вектор может быть извлечен множеством способов, и
25 подробности пропускаются здесь ради простоты. Например, коэффициент косинусного преобразования Фурье для частот чистых тонов (MFCC) параметра отличительного признака голоса может быть извлечен из голосового сигнала аудиопотока, и затем собственный вектор вычисляется с помощью способа, такого как i-вектор (алгоритм распознавания говорящего) или вероятностный линейный дискриминантный анализ
30 (PLDA, т.е., алгоритм корректировки канала для распознавания образца голоса).

[0058] Этап 404: Определение того, являются ли согласующимися собственные векторы образца голоса для множества аудиопотоков.

[0059] Например, когда пользователь читает ID пользователя, по меньшей мере, дважды во время регистрации, существуют, соответственно, по меньшей мере, два
35 собранных аудиопотока. Чтобы гарантировать, что различие между отличительными признаками образца голоса для множества аудиопотоков является не слишком большим, определение согласованности образца голоса может быть выполнено между множеством аудиопотоков. Например, оценка сходства между множеством аудиопотоков может быть вычислена на основе собственного вектора образца голоса, извлеченного из
40 каждого аудиопотока на этапе 403.

[0060] Если оценка сходства попадает в пределы некоторого порогового значения оценки, это указывает, что аудиопотоки удовлетворяют требованию по сходству; и в этом случае выполняется переход к этапу 405. Иначе, это указывает, что существует большое различие между множеством звуков, введенных пользователем, и пользователь,
45 который выполняет регистрацию, может быть проинструктирован, чтобы читать ID пользователя снова, т.е., собрать аудиопоток снова.

[0061] Этап 405: Формирование отличительного признака образца голоса модели на основе собственных векторов образца голоса для множества аудиопотоков.

[0062] На текущем этапе взвешенное суммирование может быть выполнено по собственным векторам образца голоса, извлеченным из аудиопотоков на предыдущем этапе, чтобы получать отличительный признак образца голоса модели.

[0063] После того как предыдущая процедура регистрации завершается, регистрационная информация об объекте для целевого объекта сохраняется в базе данных. Регистрационная информация об объекте может включать в себя идентификатор объекта и соответствующий физиологический признак модели. Физиологический признак модели может включать в себя отличительный признак образца голоса модели и лицевой отличительный признак модели, и обработка аутентификации личности объекта может быть выполнена ниже на основе регистрационной информации об объекте.

Аутентификация личности

[0064] Фиг. 5 иллюстрирует примерную процедуру аутентификации личности. В этой процедуре физиологический признак, используемый для аутентификации, описывается с помощью сочетания лицевого отличительного признака и отличительного признака образца голоса в качестве примера. Кроме того, физиологические признаки могут быть сравнены, после того как определяется, что целевой объект, который аутентифицируется, является живым объектом вместо видеозаписи. Как показано на фиг. 5, процедура аутентификации включает в себя следующую обработку:

[0065] Этап 501: Получение собранного аудио- и видеопотока, когда аудио- и видеопоток формируется целевым объектом, который должен быть аутентифицирован.

[0066] Например, пользователь может входить в приложение с относительно высоким требованием по безопасности, только после того как личность пользователя была аутентифицирована приложением. На текущем этапе пользователь может запускать приложение на интеллектуальном устройстве, например, смартфоне пользователя, и пользователь может собирать аудио- и видеопоток, который должен быть аутентифицирован, с помощью камеры и микрофона смартфона. Аудио- и видеопоток может быть сформирован, когда пользователь читает ID приложения пользователя.

[0067] Этап 502: Определение того, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке.

[0068] В этом примере, сначала может быть определено, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке. Что касается конкретной процедуры определения согласованности, ссылки могут быть выполнены на фиг. 2, и подробности пропускаются здесь ради простоты.

[0069] Если чтение по губам и голос являются согласующимися, это указывает, что целевой объект, который аутентифицируется, является живым объектом вместо видеозаписи, и т.д. В этом случае выполняется переход к этапу 503. Иначе, выполняется возврат к этапу 501, чтобы выполнять сбор снова.

[0070] Этап 503: Выполнение распознавания голоса по аудиопотоку в аудио- и видеопотоке, чтобы получать голосовой контент аудиопотока. Например, голосовой контент, полученный посредством распознавания, может быть ID пользователя "123456".

[0071] Этап 504: Использование голосового контента в качестве идентификатора объекта для целевого объекта и определение того, включает ли в себя предварительно сохраненная регистрационная информация об объекте идентификатор объекта.

[0072] Например, если предварительно определенная регистрационная информация об объекте включает в себя идентификатор объекта, физиологический признак модели, соответствующий идентификатору объекта, может быть получен в регистрационной информации об объекте, например, лицевой отличительный признак модели и отличительный признак образца голоса модели. Физиологическое распознавание

дополнительно выполняется по аудио- и видеопотоку, который должен быть аутентифицирован, чтобы получать физиологический признак целевого объекта, с тем, чтобы сравнивать физиологический признак с физиологическим признаком модели. Если предварительно сохраненная регистрационная информация об объекте не включает в себя идентификатор объекта, пользователь может быть уведомлен о том, что пользователь не выполнил регистрацию.

[0073] Этап 505: Выполнение распознавания образца голоса по аудио- и видеопотоку, чтобы получать отличительный признак образца голоса целевого объекта. Что касается извлечения отличительного признака образца голоса на текущем этапе, ссылки могут быть выполнены на фиг. 4.

[0074] Этап 506: Выполнение распознавания лица по аудио- и видеопотоку, чтобы получать лицевой отличительный признак целевого объекта.

[0075] Затем физиологический признак целевого объекта может быть сравнен с физиологическим признаком модели, чтобы получать результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определяется, что целевой объект был аутентифицирован. Например, содержатся этапы 507-509.

[0076] Этап 507: Сравнение отличительного признака образца голоса целевого объекта с отличительным признаком образца голоса модели, чтобы получать оценку сравнения образца голоса.

[0077] Этап 508: Сравнение лицевого отличительного признака целевого объекта с лицевым отличительным признаком модели, чтобы получать оценку сравнения лица.

[0078] Этап 509: Определение того, удовлетворяют ли оценка сравнения образца голоса и оценка совпадения лица условию аутентификации.

[0079] Например, определяется, что целевой объект был аутентифицирован, если оценка сравнения образца голоса и оценка сравнения лица удовлетворяют, по меньшей мере, одному из следующего: оценка сравнения образца голоса больше порогового значения оценки образца голоса, оценка сравнения лица больше порогового значения оценки лица; или произведение оценки сравнения образца голоса и оценки сравнения лица больше соответствующего порогового значения произведения; или взвешенная сумма оценки сравнения образца голоса и оценки сравнения лица больше соответствующего взвешенного порогового значения.

[0080] Если определяется, что оценка сравнения образца голоса и оценка сравнения лица удовлетворяют условию аутентификации на текущем этапе, определяется, что целевой объект был аутентифицирован. Иначе, определяется, что целевой объект не удастся аутентифицировать.

[0081] Кроме того, в этом примере аутентификации личности, аналогично предыдущей процедуре регистрации личности, распознавание голоса может быть выполнено по аудиопотоку, чтобы получать ID пользователя, после того как определяется, что чтение по губам и голос являются согласующимися, или ID пользователя может быть получен, когда момент времени слога в аудиопотоке идентифицируется. В предыдущем примере ID пользователя идентифицируется, после того как определяется, что чтение по губам и голос являются согласующимися.

[0082] Согласно способу аутентификации личности в этой реализации настоящей заявки, аудио- и видеопоток должен быть сформирован только однократно во время аутентификации пользователя. Например, пользователь должен читать ID пользователя только один раз. В способе ID пользователя может быть получен посредством выполнения распознавания голоса по аудио- и видеопотоку, и лицевой отличительный признак и отличительный признак образца голоса могут также быть проверены с

помощью того же аудио- и видеопотока. Это упрощает пользовательскую операцию, улучшает эффективность аутентификации, поддерживает модель аутентификации один к одному и гарантирует точность распознавания. Другими словами, распознанный физиологический признак сравнивается только с отличительным признаком, соответствующим идентификатору объекта в базе данных, тем самым, обеспечивая точность распознавания. Кроме того, в способе, согласованность между чтением по губам и голосом определяется, чтобы гарантировать, что целевой объект является живым объектом вместо поддельной видеозаписи злоумышленника, тем самым, улучшая безопасность и надежность аутентификации. В способе ID пользователя и физиологический признак, полученный посредством распознавания, получают на основе того же аудио- и видеопотока. До некоторой степени, поддельный аудио- и видеопоток злоумышленника может быть распознан.

[0083] Чтобы реализовать предыдущий способ аутентификации личности, реализация настоящей заявки дополнительно предоставляет устройство аутентификации личности. Как показано на фиг. 6, устройство может включать в себя модуль 61 получения информации, модуль 62 определения идентификатора, модуль 63 управления информацией, модуль 64 распознавания отличительного признака и модуль 65 обработки аутентификации.

[0084] Модуль 61 получения информации конфигурируется, чтобы получать собранный аудио- и видеопоток, когда аудио- и видеопоток формируется целевым объектом, который должен быть аутентифицирован.

[0085] Модуль 62 определения идентификатора конфигурируется, чтобы определять, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, и если чтение по губам и голос являются согласующимися, использовать голосовой контент, полученный посредством выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке, в качестве идентификатора объекта для целевого объекта.

[0086] Модуль 63 управления информацией конфигурируется, чтобы получать физиологический признак модели, соответствующий идентификатору объекта, из регистрационной информации об объекте, если предварительно сохраненная регистрационная информация об объекте включает в себя идентификатор объекта.

[0087] Модуль 64 распознавания отличительного признака конфигурируется, чтобы выполнять физиологическое распознавание по аудио- и видеопотоку, чтобы получать физиологический признак целевого объекта.

[0088] Модуль 65 обработки аутентификации конфигурируется, чтобы сравнивать физиологический признак целевого объекта с физиологическим признаком модели, чтобы получать результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определять, что целевой объект был аутентифицирован.

[0089] Обращаясь к фиг. 7, в примере модуль 64 распознавания отличительного признака может включать в себя подмодуль 641 распознавания образца голоса и подмодуль 642 распознавания лица.

[0090] Подмодуль 641 распознавания образца голоса конфигурируется, чтобы выполнять распознавание образца голоса по аудио- и видеопотоку, чтобы получать отличительный признак образца голоса целевого объекта.

[0091] Подмодуль 642 распознавания лица конфигурируется, чтобы выполнять распознавание лица по аудио- и видеопотоку, чтобы получать лицевой отличительный признак целевого объекта.

[0092] Модуль 65 обработки аутентификации конфигурируется, чтобы сравнивать отличительный признак образца голоса целевого объекта с отличительным признаком

образца голоса модели, чтобы получать оценку сравнения образца голоса, и сравнивать
лицевой отличительный признак целевого объекта с лицевым отличительным признаком
модели, чтобы получать оценку сравнения лица; и если оценка сравнения образца голоса
и оценка сравнения лица удовлетворяют условию аутентификации, определять, что
5 целевой объект был аутентифицирован.

[0093] В примере определяется, что целевой объект был аутентифицирован, если
оценка сравнения образца голоса и оценка сравнения лица удовлетворяют, по меньшей
мере, одному из следующего: оценка сравнения образца голоса больше порогового
значения оценки образца голоса, оценка сравнения лица больше порогового значения
10 оценки лица; или произведение оценки сравнения образца голоса и оценки сравнения
лица больше соответствующего порогового значения произведения; или взвешенная
сумма оценки сравнения образца голоса и оценки сравнения лица больше
соответствующего взвешенного порогового значения.

[0094] В примере, как показано на фиг. 7, модуль 62 определения идентификатора
15 может включать в себя следующее: подмодуль 621 распознавания слога,
сконфигурированный, чтобы идентифицировать слог голоса и соответствующий момент
времени в аудиопотоке в аудио- и видеопотоке и идентифицировать слог чтения по
губам и соответствующий момент времени в видеопотоке в аудио- и видеопотоке; и
подмодуль 622 определения согласованности, сконфигурированный, чтобы определять,
20 что чтение по губам и голос являются согласующимися, если и слог голоса, и слог
чтения по губам согласуются с соответствующим моментом времени.

[0095] В примере модуль 61 получения информации дополнительно конфигурируется,
чтобы получать аудио- и видеопоток, который должен быть зарегистрирован для
целевого объекта.

[0096] Модуль 62 определения идентификатора дополнительно конфигурируется,
чтобы использовать голосовой контент, полученный посредством выполнения
распознавания голоса по аудиопотоку в аудио- и видеопотоке, в качестве
идентификатора объекта для целевого объекта, когда чтение по губам и голос в аудио-
и видеопотоке, который должен быть зарегистрирован, являются согласующимися.

[0097] Модуль 64 распознавания отличительного признака дополнительно
30 конфигурируется, чтобы выполнять физиологическое распознавание по аудио- и
видеопотоку, который должен быть зарегистрирован, чтобы получать физиологический
признак модели аудио- и видеопотока, который должен быть зарегистрирован.

[0098] Модуль 63 управления информацией дополнительно конфигурируется, чтобы
35 соответствующим образом хранить идентификатор объекта для целевого объекта и
соответствующий физиологический признак модели в регистрационной информации
об объекте.

[0099] Предыдущие описания являются просто примерными реализациями настоящей
заявки и не предназначены, чтобы ограничивать настоящую заявку. Любые
40 модификации, эквивалентные замены, улучшения и т.д., выполненные в духе и принципе
настоящей заявки, должны попадать в рамки защиты настоящей заявки.

(57) Формула изобретения

1. Способ аутентификации личности, при этом способ содержит этапы, на которых:
45 получают собранный аудио- и видеопоток, когда аудио- и видеопоток формируется
целевым объектом, который должен быть аутентифицирован;

определяют, являются ли согласующимися чтение по губам и голос в аудио- и
видеопотоке, и если чтение по губам и голос являются согласующимися, получают

идентификатор объекта для целевого объекта путем использования голосового контента, полученного путем выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке;

определяют, содержит ли предварительно сохраненная регистрационная информация идентификатор объекта;

получают физиологический признак модели, соответствующий идентификатору объекта, из регистрационной информации об объекте, если предварительно сохраненная регистрационная информация об объекте содержит идентификатор объекта;

выполняют физиологическое распознавание по аудио- и видеопотоку, чтобы получить физиологический признак целевого объекта; и

сравнивают физиологический признак целевого объекта с физиологическим признаком модели, чтобы получить результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определяют, что целевой объект был аутентифицирован.

2. Способ по п. 1, в котором физиологический признак целевого объекта содержит отличительный признак образца голоса целевого объекта и лицевой отличительный признак целевого объекта, и физиологический признак модели содержит лицевой отличительный признак модели и отличительный признак образца голоса модели;

выполнение физиологического распознавания по аудио- и видеопотоку, чтобы получать физиологический признак целевого объекта, содержит этапы, на которых:

выполняют распознавание образца голоса по аудио- и видеопотоку, чтобы получить отличительный признак образца голоса целевого объекта; и

выполняют распознавание лица по аудио- и видеопотоку, чтобы получить лицевой отличительный признак целевого объекта; и

сравнение физиологического признака целевого объекта с физиологическим признаком модели, чтобы получить результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определение того, что целевой объект был аутентифицирован, содержат этапы, на которых:

сравнивают отличительный признак образца голоса целевого объекта с отличительным признаком образца голоса модели, чтобы получить оценку сравнения образца голоса, и сравнивают лицевой отличительный признак целевого объекта с лицевым отличительным признаком модели, чтобы получить оценку сравнения лица; и

если оценка сравнения образца голоса и оценка сравнения лица удовлетворяют условию аутентификации, определяют, что целевой объект был аутентифицирован.

3. Способ по п. 2, в котором определяется, что целевой объект был аутентифицирован, если оценка сравнения образца голоса и оценка сравнения лица удовлетворяют, по меньшей мере, одному из следующего:

оценка сравнения образца голоса больше порогового значения оценки образца голоса, оценка сравнения лица больше порогового значения оценки лица; или произведение оценки сравнения образца голоса и оценки сравнения лица больше соответствующего порогового значения произведения; или взвешенная сумма оценки сравнения образца голоса и оценки сравнения лица больше соответствующего взвешенного порогового значения.

4. Способ по п. 1, в котором определение того, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, и если чтение по губам и голос являются согласующимися, использование голосового контента, полученного посредством выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке, в качестве

идентификатора объекта для целевого объекта содержат этапы, на которых:

идентифицируют слог голоса и соответствующий момент времени в аудиопотоке в аудио- и видеопотоке;

идентифицируют слог чтения по губам и соответствующий момент времени в видеопотоке в аудио- и видеопотоке; и

определяют, что чтение по губам и голос являются согласующимися, если и слог голоса, и слог чтения по губам согласуются с соответствующим моментом времени.

5. Способ по п. 1, в котором перед получением собранного аудио- и видеопотока способ дополнительно содержит этапы, на которых:

получают аудио- и видеопоток, который должен быть зарегистрирован для целевого объекта;

используют голосовой контент, полученный посредством выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке, в качестве идентификатора объекта для целевого объекта, когда чтение по губам и голос в аудио- и видеопотоке, который должен быть зарегистрирован, являются согласующимися;

выполняют физиологическое распознавание в аудио- и видеопотоке, который должен быть зарегистрирован, чтобы получать физиологический признак модели для аудио- и видеопотока, который должен быть зарегистрирован; и

соответственно сохраняют идентификатор объекта для целевого объекта и соответствующий физиологический признак модели в регистрационной информации об объекте.

6. Устройство аутентификации личности, при этом устройство содержит:

модуль получения информации, сконфигурированный, чтобы получать собранный аудио- и видеопоток, когда аудио- и видеопоток формируется целевым объектом,

который должен быть аутентифицирован;

модуль определения идентификатора, сконфигурированный, чтобы определять, являются ли согласующимися чтение по губам и голос в аудио- и видеопотоке, и если чтение по губам и голос являются согласующимися, получать идентификатор объекта для целевого объекта путем использования голосового контента, полученного путем выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке;

модуль управления информацией, сконфигурированный, чтобы определять содержит ли предварительно сохраненная регистрационная информация идентификатор объекта и получать физиологический признак модели, соответствующий идентификатору объекта, из регистрационной информации об объекте, если предварительно сохраненная регистрационная информация об объекте содержит идентификатор объекта;

модуль распознавания отличительного признака, сконфигурированный, чтобы выполнять физиологическое распознавание по аудио- и видеопотоку, чтобы получать физиологический признак целевого объекта; и

модуль обработки аутентификации, сконфигурированный, чтобы сравнивать физиологический признак целевого объекта с физиологическим признаком модели, чтобы получать результат сравнения, и если результат сравнения удовлетворяет условию аутентификации, определять, что целевой объект был аутентифицирован.

7. Устройство по п. 6, в котором модуль распознавания отличительного признака содержит подмодуль распознавания образца голоса и подмодуль распознавания лица, при этом

подмодуль распознавания образца голоса конфигурируется, чтобы выполнять распознавание образца голоса по аудио- и видеопотоку, чтобы получать отличительный признак образца голоса целевого объекта;

подмодуль распознавания лица конфигурируется, чтобы выполнять распознавание лица по аудио- и видеопотоку, чтобы получать лицевой отличительный признак целевого объекта; и

модуль обработки аутентификации конфигурируется, чтобы сравнивать отличительный признак образца голоса целевого объекта с отличительным признаком образца голоса модели, чтобы получать оценку сравнения образца голоса, и сравнивать лицевой отличительный признак целевого объекта с лицевым отличительным признаком модели, чтобы получать оценку сравнения лица; и если оценка сравнения образца голоса и оценка сравнения лица удовлетворяют условию аутентификации, определять, что целевой объект был аутентифицирован.

8. Устройство по п. 7, в котором определяется, что целевой объект был аутентифицирован, если оценка сравнения образца голоса и оценка сравнения лица удовлетворяют, по меньшей мере, одному из следующего:

оценка сравнения образца голоса больше порогового значения оценки образца голоса, оценка сравнения лица больше порогового значения оценки лица; или произведение оценки сравнения образца голоса и оценки сравнения лица больше соответствующего порогового значения произведения; или взвешенная сумма оценки сравнения образца голоса и оценки сравнения лица больше соответствующего взвешенного порогового значения.

9. Устройство по п. 6, в котором модуль определения идентификатора содержит: подмодуль распознавания слога, сконфигурированный, чтобы идентифицировать слог голоса и соответствующий момент времени в аудиопотоке в аудио- и видеопотоке и идентифицировать слог чтения по губам и соответствующий момент времени в видеопотоке в аудио- и видеопотоке; и

подмодуль определения согласованности, сконфигурированный, чтобы определять, что чтение по губам и голос являются согласующимися, если и слог голоса, и слог чтения по губам согласуются с соответствующим моментом времени.

10. Устройство по п. 6, в котором

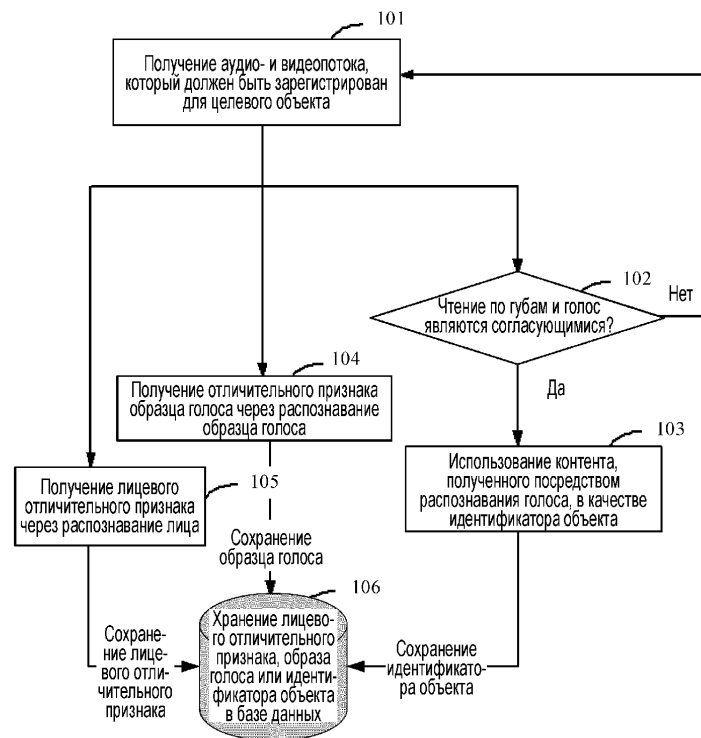
модуль получения информации дополнительно конфигурируется, чтобы получать аудио- и видеопоток, который должен быть зарегистрирован для целевого объекта;

модуль определения идентификатора дополнительно сконфигурирован, чтобы использовать голосовой контент, полученный посредством выполнения распознавания голоса по аудиопотоку в аудио- и видеопотоке, в качестве идентификатора объекта для целевого объекта, когда чтение по губам и голос в аудио- и видеопотоке, который должен быть зарегистрирован, являются согласующимися;

модуль распознавания отличительного признака дополнительно сконфигурирован, чтобы выполнять физиологическое распознавание по аудио- и видеопотоку, который должен быть зарегистрирован, чтобы получать физиологический признак модели аудио- и видеопотока, который должен быть зарегистрирован; и

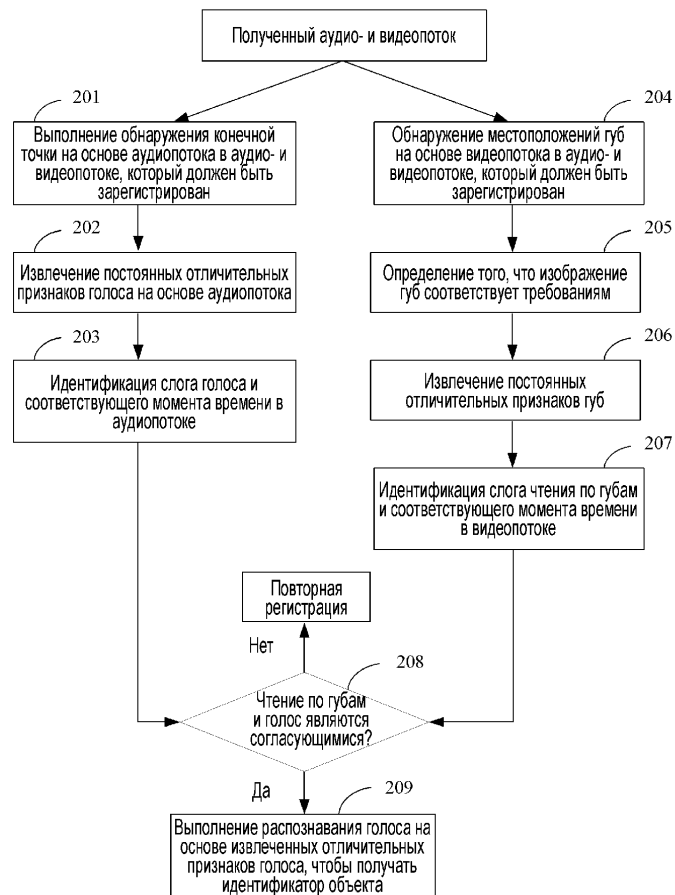
модуль управления информацией дополнительно сконфигурирован, чтобы соответствующим образом хранить идентификатор объекта для целевого объекта и соответствующий физиологический признак модели в регистрационной информации об объекте.

1/7



ФИГ. 1

2/7



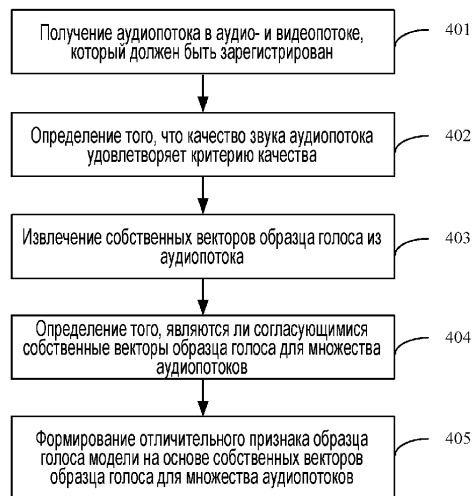
ФИГ. 2

3/7



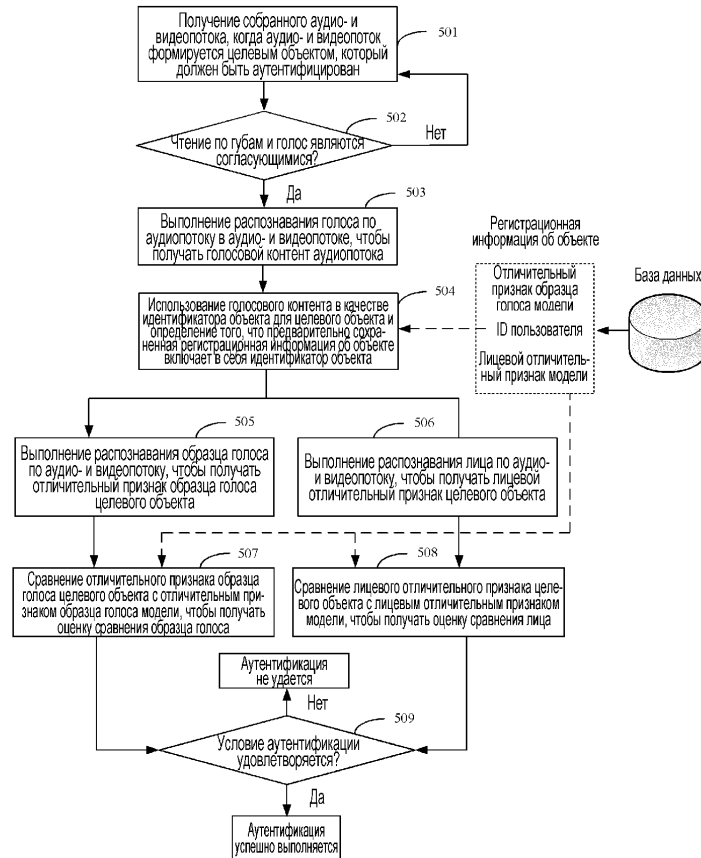
ФИГ. 3

4/7



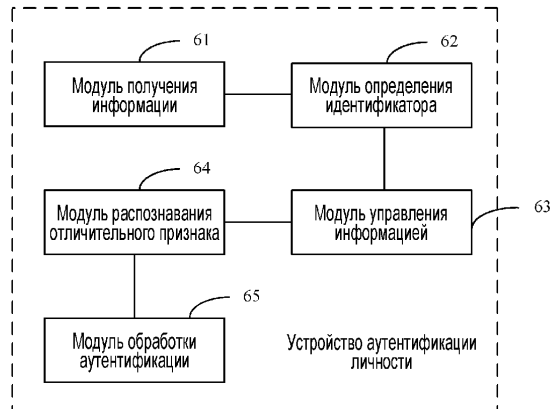
ФИГ. 4

5/7



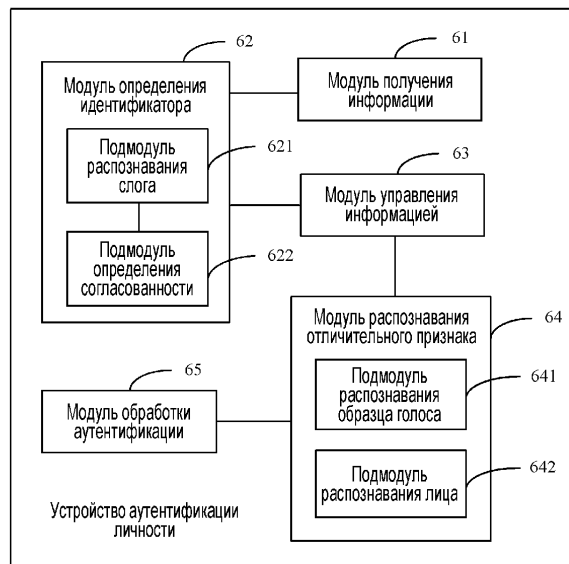
ФИГ. 5

6/7



ФИГ. 6

7/7



ФИГ. 7