



ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ

(52) СПК
G06F 17/20 (2019.02)

(21)(22) Заявка: 2016144699, 14.05.2015

(24) Дата начала отсчета срока действия патента:
14.05.2015

Дата регистрации:
21.05.2019

Приоритет(ы):

(30) Конвенционный приоритет:
16.05.2014 US 14/279,346

(43) Дата публикации заявки: 15.05.2018 Бюл. № 14

(45) Опубликовано: 21.05.2019 Бюл. № 15

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 15.11.2016

(86) Заявка РСТ:
US 2015/030687 (14.05.2015)

(87) Публикация заявки РСТ:
WO 2015/175736 (19.11.2015)

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

(72) Автор(ы):

ЭЛЬ-САБАН Мотаз Ахмад (US),
ТАВФИК Ахмед Иассин (US),
ЧАЛАБИ Ачраф Абдел Монеим Тавфик
(US),
САЙЕД Сайед Хассан (US)

(73) Патентообладатель(и):

МАЙКРОСОФТ ТЕКНОЛОДЖИ
ЛАЙСЕНСИНГ, ЭлЭлСи (US)

(56) Список документов, цитированных в отчете
о поиске: EP 2402867 A1, 04.01.2012. RU
2345414 C1, 27.01.2009. RU 2421775 C2,
20.06.2011. EP 1033662 A2, 06.09.2000. US
20060224569 A1, 05.10.2006. RU 2510935 C2,
10.04.2014.

(54) ПОИСК ИЗОБРАЖЕНИЙ НА ЕСТЕСТВЕННОМ ЯЗЫКЕ

(57) Реферат:

Изобретение относится к средствам поиска изображений на естественном языке. Достижимый технический результат заключается в расширении арсенала средств для действенного и эффективного поиска на естественном языке в коллекциях изображений. Технический результат достигается за счет того, что в способе

предусмотрена возможность принимать запрос на естественном языке, вычислять расстояние между запросами и выбирать тег изображения на основе вычислений, при этом онтология представляет собой лексическую базу данных слов, сгруппированных в когнитивные синонимы. 3 н. и 16 з.п. ф-лы, 9 ил.



FEDERAL SERVICE
FOR INTELLECTUAL PROPERTY

(12) **ABSTRACT OF INVENTION**

(52) CPC
G06F 17/20 (2019.02)

(21)(22) Application: **2016144699, 14.05.2015**

(24) Effective date for property rights:
14.05.2015

Registration date:
21.05.2019

Priority:

(30) Convention priority:
16.05.2014 US 14/279,346

(43) Application published: **15.05.2018 Bull. № 14**

(45) Date of publication: **21.05.2019 Bull. № 15**

(85) Commencement of national phase: **15.11.2016**

(86) PCT application:
US 2015/030687 (14.05.2015)

(87) PCT publication:
WO 2015/175736 (19.11.2015)

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

(72) Inventor(s):

**EL-SABAN Motaz Akhmad (US),
TAVFIK Akhmed Iassin (US),
CHALABI Achraf Abdel Monejm Tavfik (US),
SAJED Sajed Khassan (US)**

(73) Proprietor(s):

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

(54) **IMAGE SEARCH IN NATURAL LANGUAGE**

(57) Abstract:

FIELD: data processing.

SUBSTANCE: invention relates to means of searching for images in natural language. Method provides for the possibility to receive a request in a natural language, calculate the distance between requests and select an image tag based on calculations,

wherein the ontology is a lexical database of words grouped into cognitive synonyms.

EFFECT: wider range of means for effective and efficient natural language search in image collections.

19 cl, 9 dwg

R U 2 6 8 8 2 7 1 C 2

R U 2 6 8 8 2 7 1 C 2

УРОВЕНЬ ТЕХНИКИ

[1] Пользователи собирают большое количество изображений на своих различных устройствах, например камерах, цифровых камерах, видеокамерах и пр. Изображения обычно сохраняются или резервируются на персональном компьютере, в облаке или в других местах.

[2] Действительный и эффективный поиск в коллекциях изображений требует много времени и сложен для пользователей. Обычно пользователи способны только осуществлять прокрутку миниатюр изображений. Это затрудняет пользователям просмотр или поиск изображений, необходимых для конкретной задачи.

[3] Предыдущие подходы предусматривали тегирование изображений метаданными, например, метками даты и времени или ключевыми словами. Тегирование осуществляется вручную или автоматически. После тегирования, пользователи способны использовать теги в качестве запросов для определения местоположения изображений. Ограничение такого подхода состоит в том, что пользователи часто не могут вспомнить или не знают или не понимают, какие теги использовать для извлечения.

[4] Описанные ниже варианты осуществления не ограничиваются реализациями, которые исправляют некоторые или все недостатки известных систем поиска изображений.

СУЩНОСТЬ ИЗОБРЕТЕНИЯ

[5] Ниже в упрощенной форме представлена сущность изобретения для обеспечения понимания его основных принципов. Эта сущность изобретения не является обширным обзором изобретения и не идентифицирует ключевые/критические элементы и не ограничивает объем описания изобретения. Его единственной целью является представление в упрощенной форме ряда раскрытых здесь принципов, предваряющее более подробное описание, представленное ниже.

[6] В порядке примера, описан поиск изображений на естественном языке, что позволяет использовать запросы на естественном языке для извлечения изображений из хранилища изображений, автоматически тегированных тегами изображений, которые являются концептами онтологии (которая может содержать иерархию концептов). В различных примерах, запрос на естественном языке соотносится с одним или более из множества тегов изображений, и соотнесенный запрос используется для извлечения. В различных примерах, запрос соотносится путем вычисления одной или более мер расстояния между запросом и тегами изображений, причем меры расстояния вычисляются относительно онтологии и/или относительно семантического пространства слов, вычисленного из корпуса естественного языка. Семантическое пространство слов можно вычислять с использованием нейронной сети. В примерах, теги изображений могут быть связаны с ограничивающими окнами объектов, присутствующих в изображениях, и пользователь может осуществлять навигацию по хранилищу изображений путем выбора ограничивающего окна и/или изображения.

[7] Многие сопутствующие признаки будут более понятны, если обратиться к нижеследующему подробному описанию и рассматривать его совместно с прилагаемыми чертежами.

ОПИСАНИЕ ЧЕРТЕЖЕЙ

[8] Настоящее описание будет более понятно, если обратиться к нижеследующему подробному описанию, приведенному со ссылками на прилагаемые чертежи, в которых:

фиг. 1 - схема системы для поиска множества изображений с использованием естественного языка;

фиг. 2 - схема иллюстративного пользовательского интерфейса для поиска множества

изображений с использованием естественного языка;

фиг. 3 - схема другого иллюстративного пользовательского интерфейса для поиска множества изображений с использованием естественного языка;

фиг. 4 - блок-схема сервера тегирования изображений, показанного на фиг. 1;

5 фиг. 5 - блок-схема блока соотнесения запросов на естественном языке, показанного на фиг. 1;

фиг. 6 - блок-схема операций способа соотнесения термина запроса на естественном языке с одним или более тегами;

10 фиг. 7 - блок-схема операций способа поиска множества изображений с использованием естественного языка;

фиг. 8 - блок-схема операций способа навигации по множеству изображений; и

фиг. 9 демонстрирует иллюстративное вычислительное устройство, в котором можно реализовать варианты осуществления описанных здесь систем и способов.

15 Для указания аналогичных частей в прилагаемых чертежах используются одинаковые ссылочные позиции.

ПОДРОБНОЕ ОПИСАНИЕ

[9] Подробное описание, приведенное ниже совместно с прилагаемыми чертежами, призвано описывать настоящие примеры и не призвано представлять только формы, в которых можно построить или использовать настоящий пример. В описании изложены 20 функции примера и последовательность этапов для построения и действия примера. Однако такие же или эквивалентные функции и последовательности могут осуществляться в других примерах.

[10] В описанных здесь примерах используются изображения, например, цифровые фотографии. Изображения также могут быть видеозаписями.

25 [11] Здесь описаны системы и способы для поиска множества изображений с использованием запросов на естественном языке. Изображения автоматически тегятся одним или более тегами изображений, которые описывают содержание изображения. Поиск может выполняться путем соотнесения запроса на естественном языке с одним или более тегами изображений с использованием комбинации внедрения 30 онтологии и семантики. Например, в ряде случаев запрос на естественном языке соотносится путем вычисления одной или более мер расстояния между запросом и тегами изображений, причем меры расстояния вычисляются относительно онтологии и/или относительно семантического пространства слов, вычисленного из корпуса естественного языка. Затем вычисленные меры расстояния объединяются для 35 идентификации одного или более тегов, которые представляют запрос на естественном языке. Затем идентифицированные теги изображений используются для идентификации изображений, отвечающих критериям поиска (например, изображений, тегированных идентифицированными тегами изображений).

[12] Сохранение множества изображений в связи с одним или более тегами 40 изображений, описывающих содержание и/или признаки изображений, позволяет легко и эффективно извлекать изображения без необходимости анализировать каждое изображение во время извлечения или вручную редактировать или обеспечивать метаданные для каждого изображения. Извлечение изображений из множества изображений с использованием описанных способов и систем позволяет пользователям 45 быстро и легко извлекать подходящие изображения с использованием естественного языка. Это избавляет пользователей от необходимости вручную прокручивать список изображений для определения местоположения изображений с конкретным содержанием, что не только требует много времени, но и допускает возможность ошибки.

[13] Кроме того, автоматическое соотнесение терминов запроса на естественном языке с одним или более тегами изображений позволяет легко и интуитивно осуществлять поиск, поскольку пользователю не нужно знать конкретные теги изображений, и он может просто использовать язык, привычный и интуитивно понятный ему. Установлено, что использование внедрения онтологии и семантики для соотнесения терминов или фраз запроса на естественном языке с одним или более тегами создает более точное соотнесение, чем использование внедрения онтологии или семантики на нем самом.

[14] Различные описанные здесь примеры допускают поиск изображений на естественном языке (т.е. без ограничений обученными концептами/тегами) и навигацию между изображениями либо согласно сходству полных изображений, либо согласно сходству отдельных их областей.

[15] Хотя настоящие примеры описаны и проиллюстрированы здесь как реализуемые в распределенной системе извлечения изображений, описанная система предусмотрена в порядке примера, но не ограничения. Как очевидно специалистам в данной области техники, настоящие примеры пригодны для применения в системах извлечения изображений разных типов.

[16] Обратимся сначала к фиг. 1, где показана иллюстративная система для поиска множества изображений с использованием запросов на естественном языке.

[17] Система содержит сервер 102 тегирования изображений, выполненный с возможностью автоматически анализировать множество нетегированных изображений 114 и генерировать тегированное изображение 112 для каждого из нетегированных изображений 114. Нетегированные изображения 114 могут представлять собой любую коллекцию или множество изображений. Например, нетегированные изображения могут представлять собой: все изображения на конкретном устройстве (например, смартфоне), все изображения, связанные с конкретным пользователем на конкретном устройстве, или все изображения, связанные с конкретным пользователем на множестве устройств (например, смартфоне и портативном компьютере). Все изображения могут располагаться в одном месте или распределяться, например, по сети 100 связи.

[18] Каждому нетегированному изображению назначается один или более тегов для описания признаков и/или содержания изображения. Признаком может быть, например, объект, сцена, и/или ориентир в изображении. Каждый тег является концептом онтологии 108. Онтология 108 является графом из узлов, представляющих концепты, причем узлы связаны ребрами согласно отношениям между концептами. В некоторых примерах онтология может иметь иерархическую структуру с множеством подкатегорий.

[19] В частности, сервер 102 тегирования изображений выполнен с возможностью анализировать каждое нетегированное изображение 114 для идентификации признаков в изображении и назначать один или более тегов изображений каждому идентифицированному признаку для создания тегированного изображения.

Иллюстративный сервер 102 тегирования изображений будет описан ниже со ссылкой на фиг. 4.

[20] Система также содержит модуль 104 поиска и навигации изображений, что позволяет пользователю осуществлять поиск тегированных изображений 112 на естественном языке. В частности, модуль 104 поиска и навигации изображений выполнен с возможностью принимать термины и/или фразы запроса на естественном языке от пользователя через устройство 116 конечного пользователя, и выдавать термины запроса на естественном языке на блок 106 соотнесения запросов на естественном языке. Устройством 116 конечного пользователя может быть, например, смартфон,

персональный компьютер, планшетный компьютер или портативный компьютер.

[21] Блок 106 соотнесения запросов на естественном языке соотносит каждый термин или фразу запроса на естественном языке с одним или более тегами. В различных примерах блок 106 соотнесения запросов на естественном языке может быть выполнен с возможностью определять, совпадает ли термин или фраза запроса на естественном языке с одним из тегов в списке тегов. Если термин или фраза совпадает с одним из тегов в списке, блок 106 соотнесения запросов на естественном языке может выдавать термин или фразу запроса обратно на модуль 104 поиска и навигации изображений в качестве выходного сигнала процесса отображения.

[22] Если же термин или фраза запроса на естественном языке не совпадает ни с одним из тегов в списке, блок 106 соотнесения запросов на естественном языке может быть выполнен с возможностью выбирать тег или теги, наиболее сходный/е с термином или фразой запроса на естественном языке. В ряде случаев, блок 106 соотнесения запросов на естественном языке выполнен с возможностью выбирать тег или теги, наиболее сходный с термином или фразой запроса с использованием комбинации онтологического и семантического анализа. Например, блок 106 соотнесения запросов на естественном языке может вычислять одно или более расстояний между термином или фразой запроса и тегами, причем каждое расстояние представляет сходство между термином запроса и тегом. В некоторых примерах, блок 106 соотнесения запросов на естественном языке выполнен с возможностью вычислять расстояние в онтологии между термином или фразой запроса и тегами; и одно или более расстояний в семантическом пространстве между термином или фразой запроса и тегами. Затем вычисленные расстояния используются для выбора тега(ов), ближайшего или наиболее сходного с термином или фразой запроса.

[23] Например, модуль 104 поиска и навигации изображений может быть выполнен с возможностью взаимодействия с графическим пользовательским интерфейсом 118 на модуле отображения устройства 116 конечного пользователя. Графический пользовательский интерфейс 118 позволяет пользователю вводить один или более терминов и/или фраз запроса (например, в окне 120 ввода термина запроса) и инициировать поиск тегированных изображений 114 с использованием введенных терминов и/или фраз запроса (например, путем кликанья или иного выбора кнопки 122 поиска). После инициирования поиска (например, путем кликанья или иного выбора кнопки 122 поиска) термины и/или фразы запроса на естественном языке (например, введенные в окне 120 ввода термина запроса) поступают на модуль 104 поиска и навигации изображений. Затем модуль 104 поиска и навигации изображений выдает термины и/или фразы запроса на естественном языке на блок 106 соотнесения запросов на естественном языке.

[24] Если пользователь выдает термин или фраза запроса на естественном языке (например, "транспортное средство"), который не совпадает ни с одним из тегов, то блок 106 соотнесения запросов на естественном языке может соотносить термин запроса на естественном языке (например, "транспортное средство") с одним или более тегами (например "автомобиль") и выдавать соотнесенные теги (например, "автомобиль") на модуль 104 поиска и навигации изображений.

[25] Иллюстративный блок 106 соотнесения запросов на естественном языке описан со ссылкой на фиг. 5, и иллюстративный способ соотнесения термина или фразы запроса на естественном языке с одним или более тегами, который может выполняться блоком 106 соотнесения запросов на естественном языке описан со ссылкой на фиг. 6.

[26] После приема одного или более тегов от блока 106 соотнесения запросов на

естественном языке модуль 104 поиска и навигации изображений использует принятые теги для выбора изображений из тегированных изображений 122, которые отвечают терминам и/или фразам поискового запроса. В ряде случаев модуль 104 поиска и навигации изображений выполнен с возможностью выбирать изображения, которые
 5 были тегированы или связаны с принятым тегом(ами). Затем модуль 104 поиска и навигации изображений выдает выбранные изображения (например, изображения, совпадающие с терминами и/или фразами поискового запроса).

[27] Например, когда пользователь предоставляет термин запроса "транспортное средство" и блок 106 соотнесения запросов на естественном языке соотносит этот
 10 термин запроса с тегом "автомобиль", модуль 104 поиска и навигации изображений может искать тегированные изображения 112 для изображений, которым назначен тег "автомобиль". Затем модуль 104 поиска и навигации изображений может отображать результаты поиска 130 (например, изображения, совпадающие с терминами и/или фразами запроса) пользователю, например, через графический пользовательский
 15 интерфейс 124, отображаемый устройством 116 конечного пользователя. В ряде случаев, модуль 104 поиска и навигации изображений может быть выполнен с возможностью ранжировать результаты поиска до отображения их пользователю.

[28] Теперь обратимся к фиг. 2, где представлен иллюстративный графический пользовательский интерфейс 124 позволяющий пользователю искать множество
 20 изображений с использованием запросов на естественном языке.

[29] Как описано со ссылкой на фиг. 1, графический пользовательский интерфейс 124 может содержать окно 126 ввода термина запроса, которое выполнено с возможностью принимать термины и/или фразы запроса на естественном языке от
 25 пользователя. Термины и/или фразы запроса могут содержать одно или более ключевых слов или ключевых фраз (например, "автомобиль" и "человек") и один, более или ни одного термина отношений. Термином отношения является, например, термин "и", "не", "или", который указывает отношение между ключевыми словами. Также можно использовать термины пространственных отношений, например, "рядом", "справа", "слева", "вблизи". В ряде случаев графический пользовательский интерфейс может
 30 предусматривать термин отношения по умолчанию, например, если термины отношений не указаны.

[30] Графический пользовательский интерфейс 124 также содержит кнопку 128 поиска, которая, при активации (например, путем кликанья по кнопке 128 или иного выбора кнопки 128) инициирует поиск тегированных изображений 114 с использованием
 35 терминов и/или фраз запроса на естественном языке в окне 126 ввода термина запроса.

[31] Как описано выше, при инициировании поиска, термины и/или фразы запроса на естественном языке в окне 126 ввода термина запроса отправляются на модуль 104 поиска и навигации изображений, затем они преобразуются или соотносятся с одним или более тегами блоком 106 соотнесения запросов на естественном языке, затем
 40 соотнесенные теги используются для идентификации и извлечения изображений, которые отвечают терминам и/или фразам запроса на естественном языке. Затем идентифицированные изображения (или их часть или их версия) предоставляются пользователю (например, через устройство 116 конечного пользователя).

[32] В примере, показанном на фиг. 2, пользователь ищет множество изображений
 45 с использованием термина запроса на естественном языке "автомобиль". Изображения 130 (или их миниатюры или версии), отвечающие запросу (например, изображения, связанные с тегом "автомобиль"), отображаются пользователю через графический пользовательский интерфейс 124.

[33] В ряде случаев пользователь способен находить дополнительную информацию о конкретном отображаемом изображении путем кликанья по изображению или иного его выбора. Например, как показано на фиг. 2, если пользователь кликает по первому отображаемому изображению 130 или иначе выбирает его, изображение может

5 отображаться в окне 200 совместно с информацией об изображении 130 или связанной с ним (например, теги, связанные изображения и т.д.). Окно 200 может входить в состав главного графического пользовательского интерфейса 124 или может быть отделено от главного графического пользовательского интерфейса.

[34] В некоторых примерах, окно 200 может отображать список тегов 202, связанных с изображением 130. Например, на фиг. 2, окно 200 показывает, что выбранное изображение 130 связано (или тегировано) с тегами 202 "человек", "автомобиль" и "улица". В ряде случаев теги можно категоризовать, и при их отображении пользователю (например, в окне 200) они отображаются в соответствии с их категорией. Например, теги, связанные с объектами, идентифицированными в изображении, можно

15 идентифицировать как теги "объект"; теги, связанные с конкретной сценой, идентифицированной в изображении, можно идентифицировать как теги "сцена"; и теги, связанные с конкретным ориентиром, идентифицированным в изображении, можно идентифицировать как теги "ориентир". Теги, связанные с конкретной областью (или ограничивающим окном) в изображении можно идентифицировать как теги "область".

20 В ряде случаев пользователь может автоматически обновлять термины запроса путем кликанья по одному из тегов или иного его выбора. Например, если пользователь кликает по тегу "человек" или иначе выбирает его, термин "человек" может добавляться в окно ввода термина запроса.

[35] В некоторых примерах, окно 200 может также, или альтернативно, отображать

25 одно или более изображений 204 и 206, которые аналогичны выбранному изображению 130. Сходство двух изображений можно определять, например, на основе количества тегов изображений, которые они совместно используют (т.е. имеют сообща). Например, в ряде случаев чем больше тегов изображений имеют сообща два изображения, тем более они аналогичны. Сходство двух изображений также может быть основано на

30 значении достоверности, назначенном тегам изображений. Например, помимо тегирования нетегированных изображений 114 одним или более тегами изображений, сервер 102 тегирования изображений может быть выполнен с возможностью назначать значение достоверности каждому тегу, который назначается изображению. Значение достоверности указывает точность тега изображения относительно изображения

35 (например, оно указывает, что вероятность того, что изображение содержит признак (например, объект, сцену и т.д.), указанный тегом изображения). Пользователь может получать дополнительные сведения о конкретном аналогичном изображении путем кликанья по аналогичному изображению или иного его выбора. В другом примере сходство двух изображений можно определять на основе визуальных признаков,

40 извлеченных из изображений. Признаки можно извлекать с использованием глубокой нейронной сети или другими способами.

[36] В некоторых примерах, пользователь имеет возможность видеть, какие объекты были идентифицированы в выбранном изображении 130, путем перемещения курсора, например, по выбранному изображению 130, отображаемому в окне 200. При наведении

45 курсора на идентифицированный объект, идентифицированный объект может указываться или выделяться как таковой. Например, как показано на фиг. 2, вокруг идентифицированного объекта может изображаться прямоугольная рамка 208 (также именуемая ограничивающей рамкой). Ограничивающая рамка вокруг объекта может

просто всплывать поверх изображения, без фактической отрисовки. Рамка 208, при кликании, может использоваться для навигации между изображениями путем поиска изображений с соответствующими тегами области. Например, если ограничивающая рамка 208 содержит человека, то тегом области может быть "человек". В случае приема пользовательского ввода для выбора ограничивающей рамки 208 тег области можно использовать как запрос для извлечения изображений.

[37] Пользователь может автоматически добавлять термины в запрос путем кликания по объекту или иного его выбора в выбранном изображении 130. Например, если пользователь наводит курсор на одного из людей, показанных в выбранном изображении 130, прямоугольная рамка будет отображаться поверх человека. Если затем пользователь кликает в любом месте прямоугольной рамки, термин "человек" может добавляться в окно ввода термина запроса, в результате чего, оно будет содержать два термина запроса - "автомобиль" и "человек". Затем, при инициировании запроса, запрос может осуществляться для определения местоположения изображений, которые отвечают одному или обоим терминам запроса, в зависимости от конфигурации системы. Например, при автоматическом обновлении или модификации запрос для включения терминов "автомобиль" и "человек", графический пользовательский интерфейс 210 может обновляться для отображения изображений 212, которые отвечают обоим терминалам запроса (например "автомобиль" и "человек").

[38] Возможность, таким образом, автоматически обновлять термины запроса позволяет пользователю быстро и эффективно редактировать запрос и осуществлять навигацию по множеству изображений.

[39] Результаты поиска изображений можно представлять в виде множества миниатюрных изображений, размещенных в сетке или другом шаблоне. В некоторых примерах изображение высшего ранга (полученное в результате поиска) представляется в центре области результатов графического пользовательского интерфейса, и изображения более низкого ранга представляются вокруг центрального изображения, причем центральное изображение соединено с каждым из изображений более низкого ранга дугами. Дуги могут иметь ширину, цвет или другой признак, который представляет степень сходства между центральным изображением и изображениями более низкого ранга.

[40] Теперь обратимся к фиг. 3, где показан другой иллюстративный графический пользовательский интерфейс 300, позволяющий пользователю искать множество изображений с использованием естественного языка. В этом примере, графический пользовательский интерфейс 300, как и графический пользовательский интерфейс 124, показанный на фиг. 1 и 2, содержит окно 302 ввода термина запроса и кнопку 304 поиска, которыми можно пользоваться таким же образом, как окном 126 ввода термина запроса и кнопкой 128 поиска, показанными на фиг. 1 и 2.

[41] Графический пользовательский интерфейс 300, показанный на фиг. 3, также содержит инструмент 306 выбора по принципу близости. Инструмент 306 выбора по принципу близости позволяет пользователю искать изображения, которые имеют указанные термины запроса, ближайшие друг к другу в изображении. Такой поиск именуется здесь поиском или запросом по принципу близости. Например, как показано на фиг. 3, если термины запроса включают в себя "человек" и "велосипед", поиск или запрос, инициированный (например, путем кликания по кнопке 304 поиска или иного ее выбора) с использованием этих терминов, будет идентифицировать изображения, которые содержат человека, ближайшего (или находящегося в непосредственной близости) к велосипеду.

[42] В ряде случаев сервер 102 тегирования изображений может быть выполнен с возможностью записи местоположения любых объектов, идентифицированных в изображении в связи с тегированным изображением. Затем эту информацию можно использовать для динамического определения расстояния между объектами в изображениях, когда инициируется поиск по принципу близости. Например, когда модуль 104 поиска и навигации изображений принимает поиск по принципу близости от пользователя (например, через устройство 116 конечного пользователя), модуль 104 поиска и навигации изображений может быть выполнен с возможностью определения местоположения или идентификации изображений во множестве тегированных изображений, которые отвечают терминам запроса; определения расстояния между указанными объектами в идентифицированных изображениях с использованием информации местоположения; и исключения любых идентифицированных изображений, где вычисленное расстояние превышает заранее определенный порог.

[43] Альтернативно, сервер 102 тегирования изображений может быть выполнен с возможностью автоматически определять расстояние между любыми объектами в изображении и сохранять эту информацию расстояния в связи с тегированным изображением. Это позволяет быстрее извлекать изображения, отвечающие запросу по принципу близости, поскольку модулю 104 поиска и навигации изображений не требуется вычислять расстояния прежде, чем он сможет вернуть список изображений, отвечающих запросу, однако для этого требуется больше пространства для хранения дополнительной информации расстояния.

[44] Теперь обратимся к фиг. 4, где показан иллюстративный сервер 102 тегирования изображений. Как описано выше, сервер 102 тегирования изображений принимает нетегированное изображение 402 и генерирует тегированное изображение 404.

Тегированное изображение 404 представляет собой изображение, с которым связаны один или более тегов, где тег описывает признак изображения. В ряде случаев сервер 102 тегирования изображений может принимать только нетегированное изображение. В других случаях сервер 102 тегирования изображений также может принимать метаданные, связанные с изображением. Когда сервер тегирования изображений принимает метаданные помимо нетегированного изображения 102, сервер 102 тегирования изображений может использовать метаданные для помощи в тегировании изображения. Например, глобальная система позиционирования (GPS) может использоваться для извлечения близлежащих ориентиров из базы данных ориентиров. Имена близлежащих ориентиров можно использовать в качестве тегов. В другом примере, использование вспышки при фотографировании может повышать вероятность тега "ночь" или может использоваться для выбора между моделями вне помещения в дневное время и вне помещения в ночное время.

[45] Сервер 102 тегирования изображений содержит один или более модулей распознавания. Например, модуль распознавания ориентира использует данные GPS и базу данных ориентиров. Некоторые из модулей распознавания заранее обучаются для идентификации определенных признаков в изображении и связывания одного или более тегов с каждым идентифицированным признаком. Например, сервер 102 тегирования изображений, показанный на фиг. 4, содержит модуль 406 распознавания объектов, модуль 408 распознавания сцены, модуль 410 распознавания ориентира, модуль 412 распознавания активности, модуль 414 распознавания текста в изображении, модуль 416 распознавания лиц, модуль 418 распознавания пола, модуль 420 распознавания возраста, модуль 422 распознавания выражения лица. Модуль 412 распознавания активности может использовать правила или обученную систему

машинное обучение для обнаружения действий, изображенных в изображениях. Модуль распознавания текста в изображении может содержать компонент OCR. Модули распознавания возраста и пола действуют, когда получено согласие от любых людей, изображенных в изображениях. Они используют машинное обучение и/или правила для классификации людей, изображенных в изображениях, по половому и возрастному признаку. модуль распознавания выражения лица может содержать компоненты распознавания жестов и распознавания выражения лица, которые могут быть компонентами машинного обучения.

[46] В других примерах, сервер 402 тегирования изображений может содержать только один из этих модулей распознавания, другую комбинацию этих модулей распознавания или другие подходящие модули распознавания.

[47] Модуль 406 распознавания объектов выполнен с возможностью идентификации объектов в изображениях, классификации идентифицированных объектов и назначения объектов одного или более тегов на основе классификации. Модуль 404 распознавания объектов может быть выполнен с возможностью отнесения элементов изображения к одному из фиксированного количества классов объектов с использованием метода различения. Например, обученный случайный лес решений можно использовать для классификации пикселей изображения с использованием признаков пиксельной разности. В ряде случаев, каждый узел деревьев случайного леса решений связан с внешним видом или формой. Затем изображению или элементу изображения, например, ограничивающей рамке, пикселю или группе пикселей, на основе классификации, назначаются один или более тегов.

[48] Модуль 408 распознавания сцены выполнен с возможностью классификации сцены изображения и назначения одного или более тегов на основе классификации.

Классификатор сцены можно обучать на основе помеченных данных (изображений с известными сценами) для построения модели машинного обучения для данной сцены, содержащего извлечение визуальных признаков из изображений с последующим обучением классификатора (например, случайного леса или нейронной сети). Извлечение признаков может осуществляться с использованием глубокой нейронной сети, которая выполнена с возможностью осуществлять извлечение признаков и классификацию на первичных пиксельных значениях.

[49] Модуль 410 распознавания ориентира выполнен с возможностью идентификации известных ориентиров (например, пизанской башни) в изображении и назначения одного или более тегов на основе идентификации. В ряде случаев модуль 410

распознавания ориентира может работать совместно с модулем 406 распознавания объектов. Например, модуль 410 распознавания ориентира может принимать информацию от модуля 406 распознавания объектов на объектах, идентифицированных в изображении. Затем модуль 410 распознавания ориентира может использовать форму объекта и информацию местоположения в метаданных для идентификации объекта как ориентира. Информация местоположения может автоматически генерироваться устройством (например, камерой), которая сгенерировала изображение, или может вручную вводиться в метаданные. После того, как модуль 410 распознавания ориентира идентифицирует объект как ориентир, один или более тегов назначаются изображению или связываются с ним. В другом примере, метаданные GPS, связанные с изображениями, используется для поиска потенциальных ориентиров в базе данных ориентиров. При наличии более одного близкого ориентира, визуальное содержание изображения можно использовать для выбора одного из потенциальных ориентиров с использованием канонических изображений ориентиров, хранящихся в базе данных.

[50] Теперь обратимся к фиг. 5, где показан иллюстративный блок 106 соотнесения запросов на естественном языке. Как описано выше, блок 106 соотнесения запросов на естественном языке принимает термины и/или фразы 500 запроса на естественном языке от модуля 104 поиска и навигации изображений и соотносит каждый термин или фразу запроса на естественном языке с одним или более тегами 502 изображений из множества тегов 503 изображений (именуемых здесь соотнесенными тегами). В частности, блок 106 соотнесения запросов на естественном языке использует комбинацию семантического анализа и онтологии (где каждый тег является концептом в онтологии) для соотнесения каждого термина и/или фразы 500 запроса на естественном языке с одним или более тегами 502 изображений. Затем соотнесенные теги 502 изображений поступают на модуль 104 поиска и навигации изображений для идентификации изображений, которые были тегированы соотнесенными тегами 502 изображений.

[51] Блок 106 соотнесения запросов на естественном языке, показанный на фиг. 5, содержит модуль 504 семантического расстояния, выполненный с возможностью вычислять, по меньшей мере, одно расстояние в семантическом пространстве между термином или фразой запроса на естественном языке и каждым из возможных тегов изображений. Каждое расстояние указывает семантическое сходство между термином или фразой запроса на естественном языке и соответствующим тегом изображения. Семантическое сходство двух слов или фраз базируется на том, имеют ли они аналогичное смысловое значение (например, используются ли они для обозначения аналогичных вещей в одном и том же контексте).

[52] В ряде случаев семантические расстояния вычисляются модулем 504 семантического расстояния из внедрения 506 семантики слов и/или фраз, которое является семантическим пространством слов, где каждое слово или фраза соотносится с вектором внедрения низкой или высокой размерности, который представляет семантическое сходство между словами и/или фразами.

[53] В ряде случаев внедрение 506 семантики генерируется путем применения семантического кодирования 508 к корпусу 510 естественного языка. Корпус 510 естественного языка представляет собой большой объем текстов. Семантическое кодирование 508 является компонентом машинного обучения, который обучается захватывать семантическую информацию между словами.

[54] В ряде случаев семантическое кодирование является нейронной сетью, например, рекурсивной нейронной сетью (RNN), которая обучается предсказать слово на основе окружающих слов (или контекста). Следовательно, слова, которые появляются в аналогичном контексте, заканчиваются аналогичными векторами внедрения. Применение, например, нейронной сети к корпусу 510 естественного языка приводит к внедрению высокой размерности каждого слова на основе сходства использования слов в предложениях, встречающихся в корпусе естественного языка. Например, слова "warm" и "hot" могут встречаться в предложениях наподобие следующих:

The soup was still hot...
The soup was still warm...
The hot weather...
The warm weather...

[55] В результате, слова "hot" и "warm" имеют аналогичные или идентичные векторы внедрения.

[56] Модуль 504 семантического расстояния может быть выполнен с возможностью вычисления одного или более расстояний в семантическом пространстве слов (т.е.

внедрении 506 семантики). В частности, модуль 504 семантического расстояния может содержать один или более модулей расстояния, причем каждый модуль расстояния вычисляет расстояние по-разному. Например, модуль 504 семантического расстояния, показанный на фиг. 5, содержит модуль 512 косинусного сходства, который вычисляет косинусное сходство между термином и/или фразой запроса на естественном языке и отдельными тегами; модуль 514 скалярного произведения, который вычисляет скалярное произведение термина и/или фразы запроса на естественном языке и отдельными тегами; модуль 516 дайсова сходства, который вычисляет дайсово сходство термина и/или фразы запроса на естественном языке и отдельными тегами; модуль 518 расстояния Хемминга, который вычисляет расстояние Хемминга между термином и/или фразой запроса на естественном языке и отдельными тегами; и модуль 520 расстояния городских кварталов, который вычисляет расстояние городских кварталов между термином и/или фразой запроса на естественном языке и отдельными тегами. Однако в других примерах модуль 504 семантического расстояния может содержать только один из этих модулей расстояния, другую комбинацию этих модулей расстояния или различные типы модулей расстояния.

[57] Каждый модуль 512-520 расстояния вычисляет расстояние по разному, таким образом, каждый из них по разному определяет сходство между словами и/или фразами. Для получения наилучшего результата расстояния, вычисленные различными модулями 512-520 расстояния, объединяются для поиска согласия в результатах. В частности, расстояния могут поступать на модуль 522 порога, который может отвергать любое расстояние, которое превышает заранее определенный порог (указывая, что термин запроса на естественном языке и тег не вполне аналогичны). Порог может различаться для различных типов расстояний. Любое расстояние, которое не превышает соответствующий заранее определенный порог, поступает на модуль 524 выбора, где расстояния, которые превышают порог, обеспечивают голос в пользу соответствующего тега. Затем голоса объединяются для выбора тегов или тегов с наибольшим количеством голосов. В ряде случаев голоса взвешиваются на основе степени сходства (например, значения расстояния). Такое объединение расстояний повышает точность соотнесения, поскольку разные расстояния используют разные критерии. В общем случае, чем больше отличаются используемые методы вычисления расстояния, тем точнее соотнесения. Однако компромиссом является увеличенное время и ресурсы обработки.

[58] Хотя вычисление и объединение разных значений семантического расстояния приводят к весьма точным результатам соотнесения, время от времени, слово и его антоним (или не связанное слово) обычно используются в идентичном контексте. Например, антонимы и не связанные слова "быстро", "медленно" и "едва" можно использовать в аналогичном контексте, например, "медленно движущийся поезд" и "быстро движущийся поезд"; и "едва движущийся поезд". Соответственно, дополнительная информация (т.е. информация, отличная от информации семантического анализа) может быть полезна для различия этих ситуаций и, таким образом, повышения точности соотнесения.

[59] В некоторых примерах, дополнительная информация получается из онтологии. В частности, иллюстративный блок 106 соотнесения запросов на естественном языке, показанный на фиг. 5, содержит модуль 526 онтологического расстояния, который выполнен с возможностью вычислять расстояние в онтологии между термином или фразой 500 запроса на естественном языке и каждым из тегов изображений. Как описано выше, онтология 108 является графом из узлов, представляющих концепты (каждый тег является концептом в онтологии), где узлы связаны ребрами согласно отношениям

между концептами. Каждое онтологическое расстояние вычисляется путем обхода узлов в онтологии.

[60] Онтология может быть коммерчески доступной онтологией, например WordNet®, или специально разработанной онтологией. WordNet® представляет собой большую лексическую базу данных английских слов, которые группируются во множества когнитивных синонимов (синсеты), каждое из которых выражает отдельный концепт. Синсеты взаимосвязаны посредством концептуально-семантических и лексических отношений.

[61] Онтологические расстояния, сгенерированные модулем 526 онтологического расстояния, также поступают на модуль 522 порога, где любые расстояния, превышающие определенный порог, отвергаются или игнорируются, и любые расстояния, которые ниже заранее определенного порога, поступают на модуль 524 выбора, где они обеспечивают голос в пользу соответствующего тега.

[62] Теперь обратимся к фиг. 6, где показан способ соотнесения термина или фразы запроса на естественном языке с одним или более тегами изображений с использованием комбинации онтологического и семантического анализа, который может выполняться блоком 106 соотнесения запросов на естественном языке. на этапе 600, блок соотнесения запросов на естественном языке принимает термин или фразу запроса на естественном языке. Как описано выше, термин или фраза запроса на естественном языке может приниматься от модуля 104 поиска и навигации изображений после того, как модуль 104 поиска и навигации изображений принимает поисковый запрос (указывающий один или более терминов и/или фраз запроса) от пользователя через устройство 116 конечного пользователя.

[63] После приема термина и/или фразы запроса на естественном языке блок 106 соотнесения запросов на естественном языке определяет 602, присутствует ли термин и/или фраза в контрольном множестве (например, является одним из тегов изображений). Если термин и/или фраза запроса на естественном языке присутствует в контрольном множестве, то блок 106 соотнесения запросов на естественном языке выдает тег изображения на модуль 104 поиска и навигации изображений, который затем использует его для идентификации и извлечения изображений, отвечающих запросу на естественном языке 604.

[64] Если термин и/или фраза запроса на естественном языке не присутствует в контрольном множестве (например, не совпадает с тегом изображения), то способ переходит к этапам 606 и 608, на которых вычисляются онтологическое расстояние и одно или более расстояний в семантическом пространстве между термином или фразой запроса на естественном языке и отдельными тегами изображений. Как описано выше со ссылкой на фиг. 5, вычисление онтологического расстояния может содержать вычисление расстояния в онтологии (например WordNet®) между термином или фразой запроса на естественном языке и отдельными тегами изображений, причем каждый тег изображения является концептом в онтологии.

[65] Как описано выше со ссылкой на фиг. 5, вычисление одного или более семантических расстояний может содержать вычисление расстояния в семантическом пространстве слов между термином или фразой запроса на естественном языке и отдельными тегами изображений. Семантическое пространство слов может формироваться путем применения обученного компонента машинного обучения, например, нейронной сети, к корпусу текста на естественном языке. Семантические расстояния могут включать в себя одно или более из косинусного сходства, скалярного произведения, дайсова сходства, расстояния Хемминга и расстояния городских

кварталов.

[66] После генерации или вычисления онтологических и семантических расстояний способ переходит к этапу 610, на котором один или более порогов применяются к онтологическим и семантическим расстояниям для исключения или отвержения расстояний, превышающих заранее определенный порог. Для каждого типа расстояния может быть предусмотрен конкретный заранее определенный порог (например, один для онтологических расстояний и по одному для каждого типа семантического расстояния) или может быть предусмотрен один заранее определенный порог, который применяется ко всем расстояниям. Целью применения порога(ов) является исключение расстояний, которые указывают такое отдаленное сходство между термином или фразой запроса, которое не следует рассматривать при выборе надлежащего тега изображения. Благодаря исключению этих расстояний на этой стадии, можно снизить вычислительную мощность, необходимую для выбора наилучших кандидатов в теги изображения.

[67] После применения порога(ов) к вычисленным расстояниям, способ переходит к этапу 612, на котором оставшиеся расстояния используются для выбора одного или более тегов изображений, ближайших к термину или фразе запроса на естественном языке. В ряде случаев каждое оставшееся расстояние считается голосом в пользу соответствующего тега изображения. Затем голоса для каждого тега изображения накапливаются для получения счета голосов или значения для каждого тега изображения. Теги изображений с большинством голосов можно выбирать на наилучших кандидатах в теги изображения 614 и пересылать на модуль 104 поиска и навигации изображений.

[68] В ряде случаев, до накопления голосов взвешивается каждый голос. Весовые коэффициенты могут базироваться на величине связанного значения расстояния. Например, в ряде случаев меньшее значение расстояния будет давать более высокий весовой коэффициент.

[69] Теперь обратимся к фиг. 7, где показан способ поиска множества изображений с использованием терминов и/или фраз запроса на естественном языке который может выполняться модулем 104 поиска и навигации. На этапе 700 модуль 104 поиска и навигации принимает поисковый запрос (включающий в себя термины и/или фразы запроса на естественном языке и, в необязательном порядке, указатель близости) от пользователя через устройство 116 конечного пользователя. После приема терминов и/или фраз запроса на естественном языке модуль 104 поиска и навигации выдает термины и/или фразы запроса на естественном языке на блок 106 соотнесения запросов на естественном языке для соотнесения терминов и/или фраз запроса на естественном языке с одним или более тегами 702 изображений. Блок 106 соотнесения запросов на естественном языке может соотносить термины и/или фразы запроса на естественном языке с одним или более тегами изображений с использованием, например, способа, показанного на фиг. 6. Затем блок 106 соотнесения запросов на естественном языке выдает соотнесенные теги изображений на модуль 104 поиска и навигации изображений.

[70] В некоторых примерах, после приема соотнесенных тегов изображений, способ переходит к этапу 204, на котором модуль 104 поиска и навигации изображений выводит теги изображений на графический пользовательский интерфейс, отображаемый на устройстве 116 конечного пользователя. Однако выводить теги изображений на GUI не обязательно. Затем способ переходит к этапу 206.

[71] На этапе 206 модуль 104 поиска и навигации изображений использует соотнесенные теги изображений для идентификации и извлечения одного или более изображений из базы данных тегированных изображений, которые отвечают терминам и/или фразам запроса на естественном языке. Например, модуль 104 поиска и навигации

изображений может извлекать изображения, которые были тегированы соотнесенными тегами изображений. Когда поисковый запрос, содержащий указатель близости, может извлекать только изображения, которые были тегированы соотнесенными тегами изображений и имеют объекты, идентифицированные соотнесенными тегами изображений в непосредственной близости. После извлечения изображений, отвечающих запросу из базы данных тегированных изображений, способ может переходить к этапу 208 или способ может переходить непосредственно к этапу 210.

[72] На этапе 208, модуль 104 поиска и навигации изображений ранжирует извлеченные изображения на основе того, насколько они отвечают критериям поиска. Например, как описано выше, в ряде случаев сервер 102 тегирования изображений может быть выполнен с возможностью назначать значение достоверности каждому тегу изображения, назначенному изображению. Значение достоверности указывает точность тега (например, вероятность того, что изображение содержит элемент, идентифицированный тегом). В этих случаях модуль 104 поиска и навигации изображений может быть выполнен с возможностью ранжировать извлеченные изображения. Например, чем выше значение достоверности для соотнесенных тегов изображений (которые пересекаются с соотнесенными терминами запроса) тем выше ранг изображения. В других случаях, для ранжирования извлеченных изображений можно использовать другие критерии. Например, блок ранжирования машинного обучения можно обучать ранжировать порядок результатов поиска на основе пар запрос-изображение, вручную определенных человеком-комментатором.

[73] На этапе 210 модуль 104 поиска и навигации изображений может выводить ранжированные или не ранжированные извлеченные изображения на графический пользовательский интерфейс устройства 116 конечного пользователя.

[74] Теперь обратимся к фиг. 8, где показан способ навигации по множеству изображений, который может выполняться модулем 104 поиска и навигации изображений. На этапе 800 модуль 104 поиска и навигации изображений принимает от устройства 116 конечного пользователя указание, что пользователь выбрал одно из отображаемого изображения или объекта в отображаемом изображении (указанного, например, ограничивающей рамкой).

[75] Модуль 104 поиска и навигации изображений извлекает теги, связанные с выбранным изображением или выбранным объектом 802, и отображает теги изображений для выбранного изображения или объекта в графическом пользовательском интерфейсе 804. Когда пользователь выбирает изображение, теги изображений для изображения может отображаться в виде списка в графическом пользовательском интерфейсе, как показано на фиг. 2. Однако, когда пользователь выбирает объект в изображении, тег изображения, связанный с объектом, может отображаться, например, поверх ограничивающей рамки или в окне ввода термина запроса, как показано на фиг. 2.

[76] Модуль 104 поиска и навигации изображений также извлекает изображения с использованием тегов изображений для выбранного изображения или выбранного объекта. Когда пользователь выбирает изображение, извлеченные изображения могут быть изображениями, которые аналогичны выбранному изображению. Сходство может базироваться на тегах изображений, которые совместно используются. Чем больше тегов изображений совместно используется, тем более аналогичны два изображения. Соответственно, когда пользователь выбирает изображение, модуль 104 поиска и навигации изображений может быть выполнен с возможностью извлечения изображений из базы данных тегированных изображений, которые были тегированы теми же тегами

изображений, что и выбранное изображение. Однако, когда пользователь выбирает изображение, извлеченные изображения могут быть изображениями, которые содержат термины запроса в окне ввода термина запроса (которое теперь включает в себя тег изображения, связанный с выбранным объектом). После извлечения изображений из
 5 базы данных тегированных изображений, способ может переходить к этапу 808 или может переходить непосредственно к этапу 810.

[77] На этапе 808 извлеченные изображения ранжируются на основе того, насколько точные теги изображений используются для извлечения. Например, как описано выше, в ряде случаев сервер 102 тегирования изображений может быть выполнен с
 10 возможностью назначать значение достоверности каждому тегу изображения, назначенному изображению. Значение достоверности указывает точность тега (например, вероятность того, что изображение содержит элемент, идентифицированный тегом). В этих случаях модуль 104 поиска и навигации изображений может быть выполнен с возможностью ранжировать извлеченные изображения с использованием
 15 значений достоверности. Например, чем выше значение достоверности для соотнесенных тегов изображений, тем выше ранг изображения. В других случаях, для ранжирования извлеченных изображений можно использовать другие критерии. После ранжирования извлеченных изображений способ переходит к этапу 810.

[78] На этапе 810 модуль 104 поиска и навигации изображений выводит
 20 ранжированный или не ранжированный список извлеченных изображений на графический пользовательский интерфейс, отображаемый на устройстве 116 конечного пользователя. Когда пользователь выбирает изображение, извлеченные изображения (изображения, аналогичные выбранным изображениям) могут отображаться во вторичном окне GUI, как показано на фиг. 2. Однако, когда пользователь выбирает
 25 объект, извлеченные изображения (изображения, отвечающие терминам запроса) могут отображаться в главном окне результатов GUI, как показано на фиг. 2.

[79] На этапе 812 модуль поиска и навигации изображений может принимать от устройства 116 конечного пользователя указание, что пользователь указал, что желает совместно использовать отображаемые изображения с другой стороной. Когда модуль
 30 104 поиска и навигации изображений принимает такое указание, модуль 104 поиска и навигации изображений может переходить к этапу 814, на котором извлеченные изображения предоставляются указанным сторонам, например, инструментом социальной сети, доступном пользователю и/или устройству 116 конечного пользователя.

[80] Фиг. 9 демонстрирует различные компоненты иллюстративного вычислительного
 35 устройства 900, которое можно реализовать как любую форму вычислительного и/или электронного устройства, и в котором можно реализовать варианты осуществления описанных здесь систем и способов.

[81] Вычислительное устройство 900 содержит один или более процессоров 902, которые могут быть микропроцессорами, контроллерами или любым другим
 40 подходящим типом процессоров для обработки компьютерно-исполнимых инструкций для управления работой устройства для поиска множества изображений с использованием естественного языка. В некоторых примерах, например, где используется архитектура системы на кристалле, процессоры 902 может включать в себя один или более фиксированных функциональных блоков (также именуемых
 45 ускорителями), которые реализуют часть способа поиска множества изображений с использованием естественного языка в оборудовании (а не в программном обеспечении или программно-аппаратном обеспечении). Программное обеспечение платформы, содержащее операционную систему 904 или любое другое подходящее программное

обеспечение платформы, может обеспечиваться на вычислительном устройстве 900, чтобы на устройстве могло выполняться прикладное программное обеспечение, например, блок 906 соотнесения запросов и модуль 912 поиска и навигации изображений.

[82] Компьютерно-исполнимые инструкции могут обеспечиваться с использованием любых компьютерно-считываемых носителях, к которым может осуществлять доступ вычислительное устройство 900. Компьютерно-считываемые носители может включать в себя, например, компьютерные носители данных, например, память 910 и среды связи. Компьютерные носители данных, например память 910, включает в себя энергозависимые и энергонезависимые, сменные и стационарные носители, реализованные согласно любому способу или по любой технологии для хранения информации, например, компьютерно-считываемых инструкций, структур данных, программных модулей или других данных. Компьютерные носители данных включают в себя, но без ограничения, RAM, ROM, EPROM, EEPROM, флэш-память или другую технологию памяти, CD-ROM, цифровые универсальные диски (DVD) или другое оптическое запоминающее устройство, магнитные кассеты, магнитную ленту, запоминающее устройство на основе магнитного диска или другие магнитные запоминающие устройства, или любой другой носитель, не связанный с передачей, который можно использовать для хранения информации для доступа вычислительным устройством. Напротив, среды передачи данных могут воплощать компьютерно-считываемые инструкции, структуры данных, программные модули, или другие данные в модулированном сигнале данных, например, несущей волне, или другом механизме переноса. По определению, компьютерные носители данных не включают в себя среды передачи данных. Таким образом, компьютерный носитель данных не следует интерпретировать как распространяющийся сигнал сам по себе. Распространяющиеся сигналы могут присутствовать в компьютерных носителях данных, но распространяющиеся сигналы сами по себе не являются примерами компьютерных носителей данных. Хотя компьютерные носители данных (память 910) показаны в вычислительном устройстве 900 очевидно, что хранилище может быть распределенным или располагаться удаленно, и доступ к нему может осуществляться через сеть или другую линию связи (например, с использованием интерфейса 916 связи).

[83] Вычислительное устройство 900 также содержит контроллер 914 ввода/вывода, выполненный с возможностью выводить отображаемую информацию на устройство 920 отображения, которое может быть отдельным от вычислительного устройства 800 или объединенным с ним. Отображаемая информация может обеспечивать графический пользовательский интерфейс. Контроллер 914 ввода/вывода также выполнен с возможностью принимать и обрабатывать ввод от одного или более устройств, например, устройства 922 пользовательского ввода (например, мыши, клавиатуры, камеры, микрофона или другого датчика). В некоторых примерах устройство 922 пользовательского ввода может обнаруживать голосовой ввод, жесты пользователя или другие действия пользователя и может обеспечивать естественный пользовательский интерфейс (NUI). Этот пользовательский ввод можно использовать для управления работой вычислительного устройства 900. Согласно варианту осуществления, устройство 920 отображения также может действовать как устройство 922 пользовательского ввода, если оно является устройством отображения, чувствительным к прикосновению. Контроллер 914 ввода/вывода также может выводить данные на устройства, отличные от устройства отображения, например, локально подключенное печатающее устройство (не показанный на фиг. 9).

[84] Любой из контроллер 914 ввода/вывода, устройства 920 отображения и

устройство 922 пользовательского ввода может содержать технологию NUI, которая позволяет пользователю взаимодействовать с вычислительным устройством естественным образом, без каких-либо искусственных ограничений, налагаемых устройствами ввода, например, мышами, клавиатурами, органами удаленного управления и пр. Примеры технологии NUI, которые можно обеспечить, включают в себя, но без ограничения, распознавание голоса и/или речи, распознавание прикосновения и/или стилуса (дисплеи, чувствительные к прикосновению), распознавание жестов, как на экране, так и рядом с экраном, трехмерных жестов, отслеживание движения головы и глаз, голоса и речи, зрения, прикосновения, жестов и машинный интеллект. Другие примеры технологии NUI, которые можно использовать, включают в себя системы понимания намерения и цели, системы обнаружения с использованием камер с восприятием глубины сцены (например, системы стереоскопической камеры, системы инфракрасной камеры, системы rgb-камеры и их комбинации), обнаружение жестовых движений с использованием акселерометров/гироскопов, распознавание лиц, 3D дисплеи, отслеживание головы, глаз и взгляда, системы погружения в дополненную реальность и виртуальную реальность и технологии для регистрации мозговой активности с использованием зондов электрического поля (ЭЭГ и родственные способы).

[85] Альтернативно или дополнительно, описанные здесь функциональные возможности могут осуществляться, по меньшей мере, частично, одним или более аппаратными логическими компонентами. Например, и без ограничения, иллюстративные типы аппаратных логических компонентов, которые можно использовать, включают в себя вентильные матрицы, программируемые пользователем (FPGA), программно-зависимые интегральные схемы (ASIC), программно-зависимые стандартные продукты (ASSP), однокристальные системы (SOC), сложные программируемые логические устройства (CPLD).

[86] Термин 'компьютер' или 'вычислительное устройство' используется здесь в отношении любого устройства с возможностями обработки, таким образом, что оно может выполнять инструкции. Специалистам в данной области техники очевидно, что такие возможности обработки включены во многие разные устройства, и, таким образом, термины 'компьютер' или 'вычислительное устройство' включают в себя ПК, серверы, мобильные телефоны (включая смартфоны), планшетные компьютеры, телевизионные приставки, медиапроигрыватели, игровые консоли, карманные персональные компьютеры и многие другие устройства.

[87] Описанные здесь способы могут осуществляться программным обеспечением в машиночитаемой форме на вещественном носителе данных, например, в форме компьютерной программы, содержащей средство компьютерного программного кода, адаптированного для осуществления всех этапов любого из описанных здесь способов, когда программа выполняется на компьютере и где компьютерная программа может быть воплощена на компьютерно-считываемом носителе. Примеры вещественных носителей данных включают в себя компьютерные запоминающие устройства, содержащие компьютерно-считываемые носители, например, диски, флэш-накопители, память и т.д. и не включают в себя распространяющиеся сигналы. Распространяющиеся сигналы могут присутствовать в вещественных носителях данных, но распространяющиеся сигналы сами по себе не являются примерами вещественных носителей данных. Программное обеспечение может быть пригодно для выполнения на параллельном процессоре или последовательном процессоре таким образом, что этапы способа могут осуществляться в любом пригодном порядке или одновременно.

[88] Это подтверждает, что программное обеспечение может быть ценным, отдельно

продаваемым товаром. Оно призвано охватывать программное обеспечение, которое выполняется на "немом" или стандартном оборудовании или управляет им, для осуществления желаемых функций. Оно также призвано охватывать программное обеспечение, которое "описывает" или задает конфигурацию оборудования, например, программное обеспечение HDL (язык описания аппаратного обеспечения), используемое для проектирования кремниевых кристаллов, или для конфигурирования универсальных программируемых микросхем, для осуществления желаемых функций.

[89] Специалистам в данной области техники очевидно, что запоминающие устройства, используемые для хранения программных инструкций, могут быть распределены по сети. Например, на удаленном компьютере может храниться пример процесса, описанного как программное обеспечение. Локальный или оконечный компьютер может осуществлять доступ к удаленному компьютеру и загружать часть или все программное обеспечение для выполнения программы. Альтернативно, локальный компьютер при необходимости может загружать фрагменты программного обеспечения или выполнять некоторые команды на локальном терминале и некоторые - на удаленном компьютере (или в компьютерной сети). Специалисты в данной области техники также очевидно, что, согласно традиционным подходам, известным специалистам в данной области техники, все или часть команд может осуществляться специализированной схемой, например, DSP, программируемой логической матрицей и т.п.

[90] Любой приведенный здесь диапазон или значение устройства можно расширить или изменить без потери искомого результата, что очевидно специалисту в данной области техники.

[91] Хотя предмет изобретения описан в отношении структурных признаков и/или методологических действий, следует понимать, что предмет изобретения, заданный в нижеследующей формуле изобретения, не обязан ограничиваться конкретными признаками или действиями, описанными выше. Напротив, вышеописанные конкретные признаки и действия раскрыты как иллюстративные формы реализации формулы изобретения.

[92] Следует понимать, что вышеописанные достоинства и преимущества могут относиться к одному варианту осуществления или могут относиться к нескольким вариантам осуществления. Варианты осуществления не ограничиваются теми, которые решают некоторые или все из указанных проблем, или теми, которые обладают некоторыми или всеми из указанных достоинств и преимуществ. Следует также понимать, что употребление наименования элемента в единственном числе подразумевает один или более из этих элементов.

[93] Этапы описанных здесь способов могут осуществляться в любом пригодном порядке или одновременно, когда это целесообразно. Дополнительно, отдельные блоки можно удалять из любого способа не выходя за рамки сущности и объема описанного здесь предмета изобретения. Аспекты любого из вышеописанных примеров можно комбинировать с аспектами любого из других описанных примеров, для формирования дополнительных примеров без потери искомого результата.

[94] Термин 'содержащий' используется здесь в смысле включения идентифицированных блоков способа или элементов, но при этом такие блоки или элементы не содержат исключительный перечень, и способ или устройство может содержать дополнительные блоки или элементы.

[95] Следует понимать, что вышеприведенное описание приведено исключительно в порядке примера, и что специалисты в данной области техники могут вносить

различные изменения. Вышеприведенное описание изобретения, примеры и данные обеспечивают полное описание структуры и использования иллюстративных вариантов осуществления. Хотя различные варианты осуществления были описаны выше с некоторой степенью детализации или со ссылкой на один или более отдельных вариантов осуществления, специалисты в данной области техники могут вносить многочисленные изменения в раскрытые варианты осуществления, не выходя за рамки сущности или объема этого описания изобретения.

(57) Формула изобретения

1. Компьютерно-реализуемый способ поиска изображений на естественном языке, содержащий этапы, на которых:

принимают запрос на естественном языке на поиск изображений в базе данных изображений, которые тегированы одним или более тегами изображений из множества тегов изображений;

вычисляют первое расстояние в онтологии между запросом на естественном языке и множеством тегов изображений, при этом онтология представляет собой лексическую базу данных слов, сгруппированных в наборы когнитивных синонимов, причем каждый набор связан с концептом, соответствующим когнитивным синонимам данного набора, при этом каждый концепт представлен как узел в графе узлов, соединенных ребрами согласно концептуально-семантическим отношениям между концептами, причем каждый тег изображения из множества тегов изображений является концептом онтологии;

вычисляют второе расстояние в семантическом пространстве слов между запросом на естественном языке и множеством тегов изображений, при этом второе расстояние показывает семантическое сходство между запросом на естественном языке и множеством тегов изображений, причем семантическое пространство слов формируется программой с машинным обучением на основе смыслового значения слов в корпусе слов;

выбирают по меньшей мере один тег изображения из множества тегов изображений на основе обоих из первого расстояния и второго расстояния, причем выбранного по меньшей мере одного тега изображения нет в запросе на естественном языке;

осуществляют поиск по базе данных изображений с использованием выбранного по меньшей мере одного тега изображения;

извлекают, в ответ на упомянутый поиск, из базы данных изображений одно или более изображений, тегированных выбранным по меньшей мере одним тегом изображения; и

обеспечивают представление извлеченных одного или более изображений в пользовательском интерфейсе.

2. Способ по п. 1, в котором первое расстояние вычисляется путем обхода узлов в онтологии.

3. Способ по п. 1, в котором программой с машинным обучением является нейронная сеть.

4. Способ по п. 1, в котором второе расстояние вычисляется с использованием любого из косинусного сходства, скалярного произведения, дайсова сходства, расстояния Хемминга и расстояния городских кварталов.

5. Способ по п. 1, дополнительно содержащий этап, на котором вычисляют третье расстояние в семантическом пространстве слов между запросом на естественном языке и множеством тегов изображений, причем третье расстояние вычисляется с использованием операции, отличающейся от таковой для второго расстояния, при этом

упомянутый выбор по меньшей мере одного тега изображения дополнительно основывается на третьем расстоянии.

6. Способ по п. 1, в котором упомянутый выбор по меньшей мере одного тега изображения из множества тегов изображений дополнительно содержит этап, на котором выбирают теги изображений, которые находятся в пределах заранее определенного первого порога расстояния и в пределах заранее определенного второго порога расстояния.

7. Способ по п. 1, в котором упомянутый выбор по меньшей мере одного тега изображения из множества тегов изображений дополнительно содержит этапы, на которых:

представляют каждое вычисленное первое и второе расстояние как голос для конкретного тега изображения из множества тегов изображений;

объединяют голоса за каждый тег изображения и

выбирают по меньшей мере один тег изображения на основе голосов.

8. Способ по п. 7, в котором каждому голосу назначают вес на основе величины соответствующего расстояния до объединения голосов.

9. Способ по п. 1, в котором упомянутое обеспечение представления в пользовательском интерфейсе дополнительно содержит этапы, на которых:

отображают по меньшей мере часть одного или более извлеченных изображений;

принимают информацию, указывающую выбор одного из извлеченных изображений; и

отображают выбранное изображение и информацию, относящуюся к выбранному изображению.

10. Способ по п. 9, в котором информация, относящаяся к выбранному изображению, содержит одно или более изображений, которые сходны с выбранным изображением.

11. Способ по п. 10, в котором сходство двух изображений основывается на тегах изображений, которые совместно используются этими двумя изображениями, и значениях достоверности, связанных с каждым совместно используемым тегом.

12. Способ по п. 9, дополнительно содержащий этапы, на которых:

принимают информацию, указывающую позицию курсора относительно выбранного изображения, причем курсором управляет пользователь;

определяют, располагается ли курсор поверх объекта, идентифицированного в выбранном изображении; и

в качестве реакции на определение того, что курсор располагается поверх объекта, идентифицированного в выбранном изображении, отображают ограничивающую рамку вокруг идентифицированного объекта.

13. Способ по п. 12, дополнительно содержащий этапы, на которых:

принимают указание выбора ограничивающего окна и

обновляют запрос на естественном языке включением в него тега изображения, связанного с идентифицированным объектом, соответствующим ограничивающему окну.

14. Способ по п. 1, в котором запрос на естественном языке содержит множество терминов запроса и указание того, должны ли термины из множества терминов запроса располагаться в пределах третьего расстояния друг от друга, и в качестве реакции на определение того, что термины должны располагаться в пределах третьего расстояния друг от друга, извлекают из базы данных изображений одно или более изображений, тегированных каждым из выбранных тегов изображений, причем объекты, связанные с выбранными тегами изображений, располагаются в пределах третьего расстояния

друг от друга.

15. Способ по п. 1, дополнительно содержащий этапы, на которых:
автоматически идентифицируют по меньшей мере один признак в каждом изображении из множества нетегированных сохраненных изображений и

связывают соответственный тег изображения из множества тегов изображений с каждым изображением из множества нетегированных сохраненных изображений, причем каждый соответственный тег изображения соответствует по меньшей мере одному признаку, идентифицированному в соответствующем изображении из множества нетегированных сохраненных изображений.

16. Способ по п. 1, дополнительно содержащий этапы, на которых:

принимают данные, указывающие, что одно или более извлеченных изображений должны совместно использоваться; и

делают одно или более извлеченных изображений доступными для одной или более других сторон.

17. Вычислительное устройство, выполненное с возможностью поиска изображений на естественном языке, при этом вычислительное устройство включает в себя процессор, дисплей, подключенный к процессору, и память, хранящую инструкции, которые при их исполнении процессором предписывают процессору:

принимать запрос на естественном языке на поиск изображений в базе данных

изображений, которые тегированы одним или более тегами изображений из множества тегов изображений;

вычислять первое расстояние в онтологии между запросом на естественном языке и множеством тегов изображений, при этом онтология представляет собой лексическую базу данных слов, сгруппированных в наборы когнитивных синонимов, причем каждый набор связан с концептом, соответствующим когнитивным синонимам данного набора, при этом каждый концепт представлен как узел в графе узлов, соединенных ребрами согласно концептуально-семантическим отношениям между концептами, причем каждый тег изображения из множества тегов изображений является концептом онтологии;

вычислять второе расстояние в семантическом пространстве слов между запросом на естественном языке и множеством тегов изображений, при этом второе расстояние показывает семантическое сходство между запросом на естественном языке и множеством тегов изображений, причем семантическое пространство слов формируется программой с машинным обучением на основе смыслового значения слов в корпусе слов;

выбирать по меньшей мере один тег изображения из множества тегов изображений на основе обоих из первого расстояния и второго расстояния, причем выбранного по меньшей мере одного тега изображения нет в запросе на естественном языке;

осуществлять поиск по базе данных изображений с использованием выбранного по меньшей мере одного тега изображения;

извлекать, в ответ на упомянутый поиск, из базы данных изображений одно или более изображений, тегированных выбранным по меньшей мере одним тегом изображения; и

обеспечивать представление извлеченных одного или более изображений в пользовательском интерфейсе

18. Вычислительное устройство по п. 17, при этом вычислительное устройство по меньшей мере отчасти реализовано с использованием аппаратных логических средств, выбранных из одного или более из программируемой пользователем вентильной матрицы, специализированной интегральной схемы, специализированного стандартного

продукта, однокристалльной системы, сложного программируемого логического устройства.

19. Долговременный машиночитаемый носитель, включающий в себя инструкции, которые при их исполнении вычислительным устройством предписывают
5 вычислительному устройству выполнять операции, содержащие:

прием запроса на естественном языке на поиск изображений в базе данных изображений, которые тегированы одним или более тегами изображений из множества тегов изображений;

вычисление первого расстояния в онтологии между запросом на естественном языке
10 и множеством тегов изображений, при этом онтология представляет собой лексическую базу данных слов, сгруппированных в наборы когнитивных синонимов, причем каждый набор связан с концептом, соответствующим когнитивным синонимам данного набора, при этом каждый концепт представлен как узел в графе узлов, соединенных ребрами согласно концептуально-семантическим отношениям между концептами, причем каждый
15 тег изображения из множества тегов изображений является концептом онтологии;

вычисление второго расстояния в семантическом пространстве слов между запросом на естественном языке и множеством тегов изображений, при этом второе расстояние показывает семантическое сходство между запросом на естественном языке и
множеством тегов изображений, причем семантическое пространство слов формируется
20 посредством применения обученной нейронной сети к корпусу документов на естественном языке;

выбор по меньшей мере одного тега изображения из множества тегов изображений на основе обоих из первого расстояния и второго расстояния, причем выбранного по
по меньшей мере одного тега изображения нет в запросе на естественном языке;

25 осуществление поиска по базе данных изображений с использованием выбранного по меньшей мере одного тега изображения;

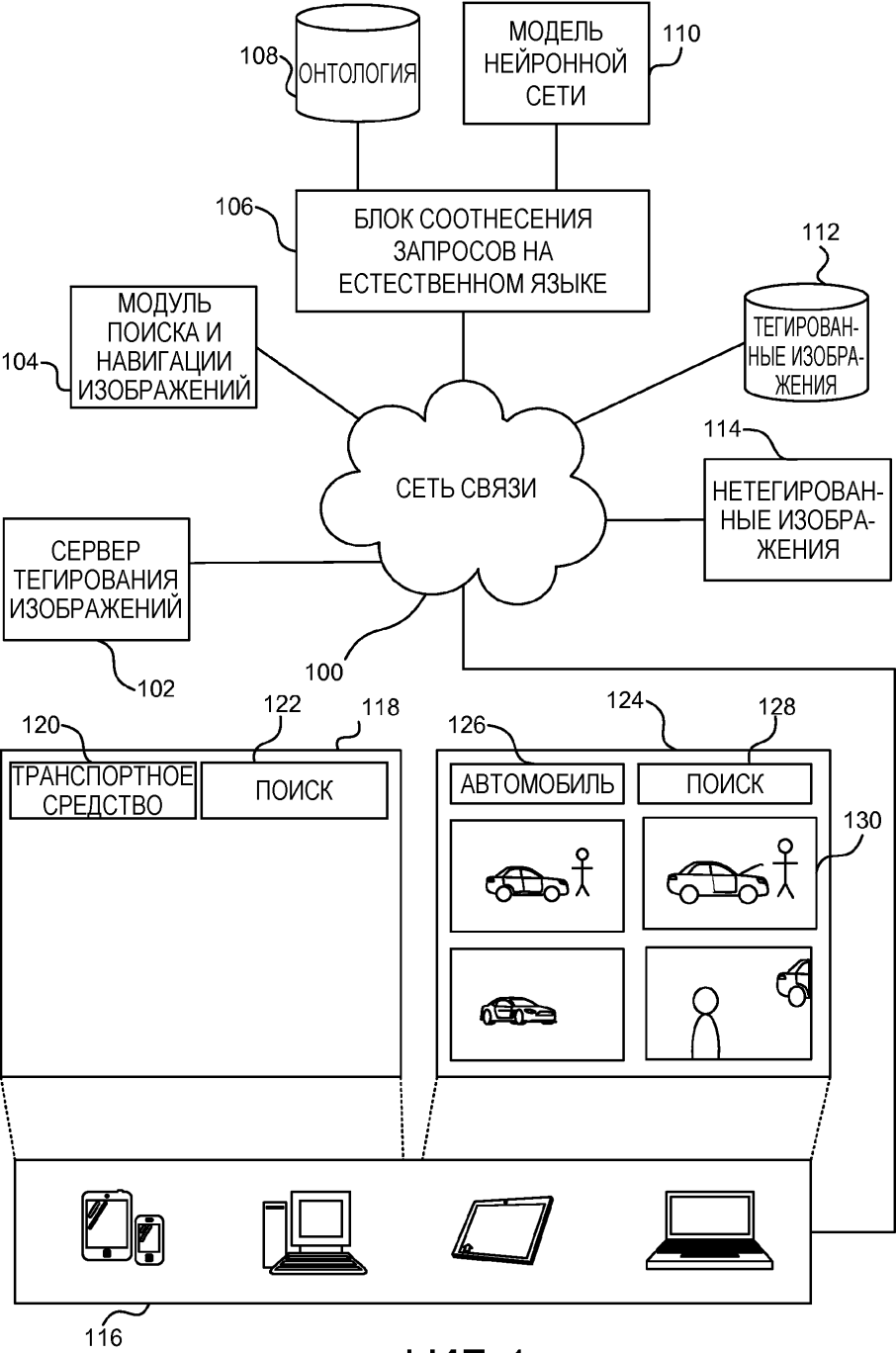
извлечение, в ответ на упомянутый поиск, из базы данных изображений одного или более изображений, тегированных выбранным по меньшей мере одним тегом
изображения; и

30 обеспечение представления извлеченных одного или более изображений в пользовательском интерфейсе.

35

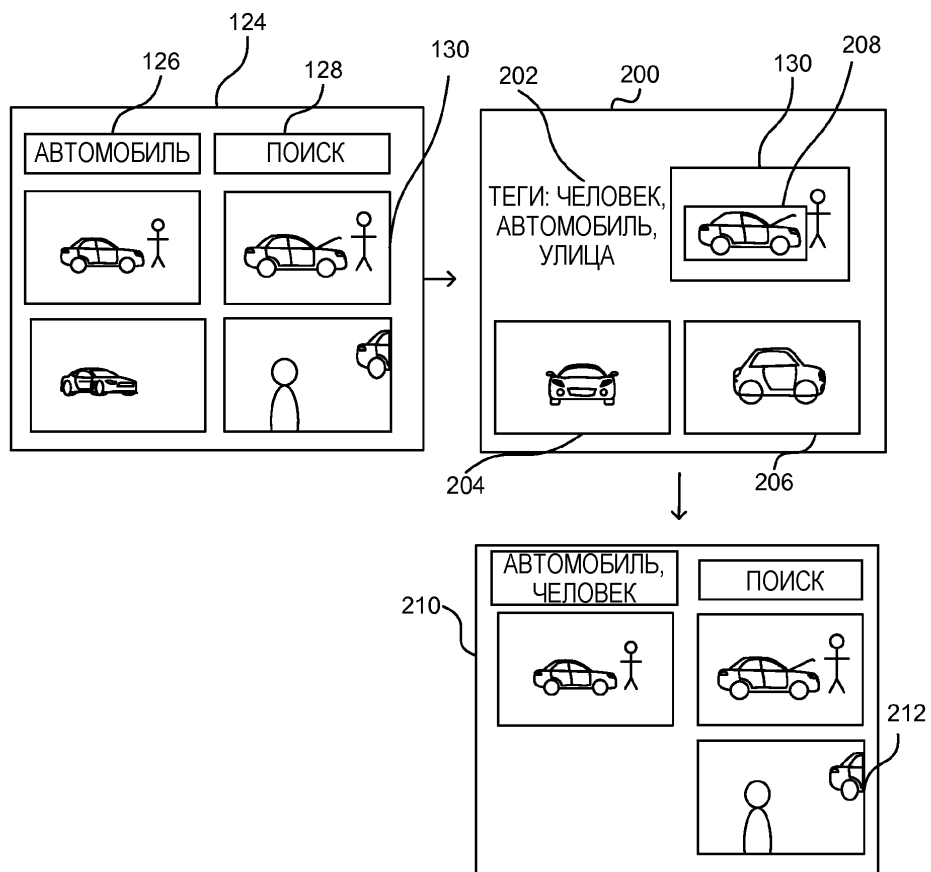
40

45



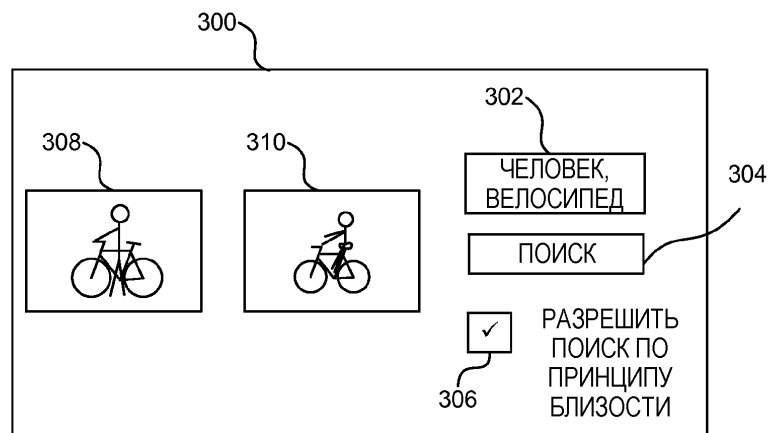
ФИГ. 1

2/9



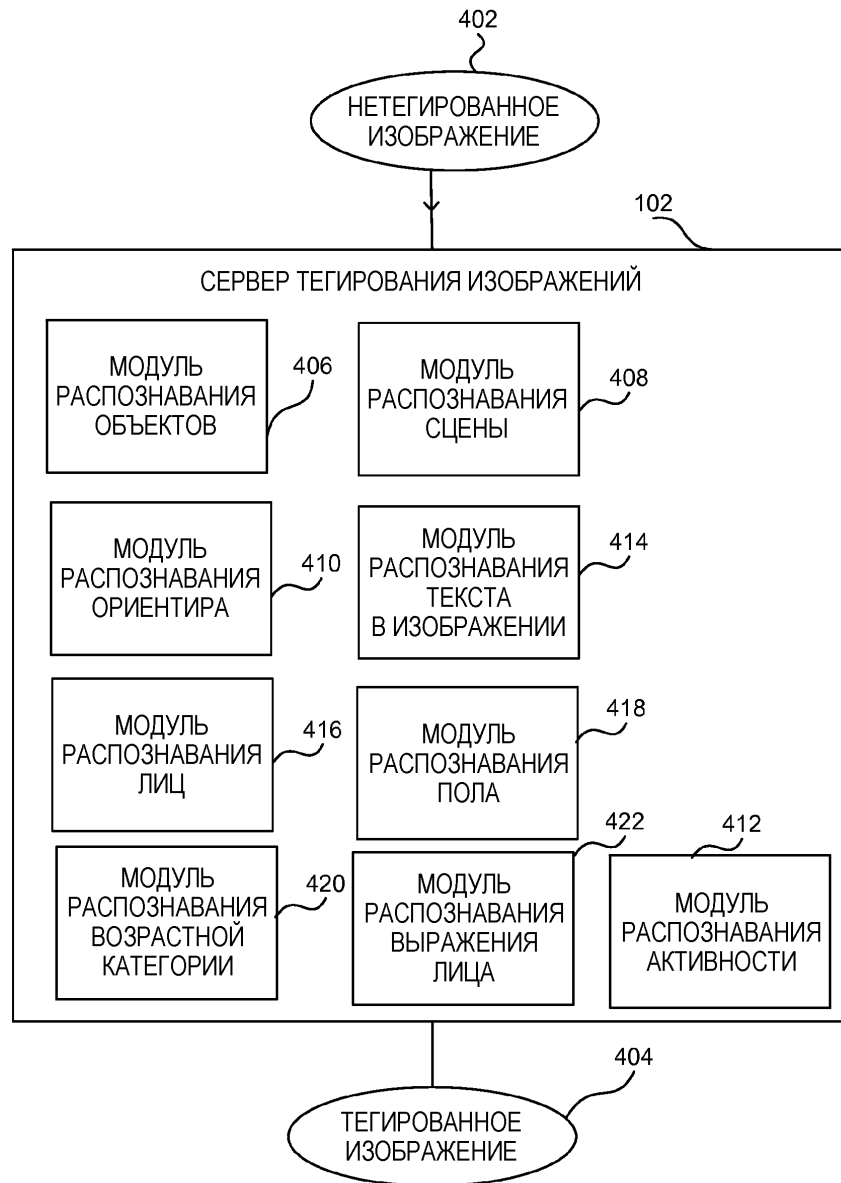
ФИГ. 2

3/9



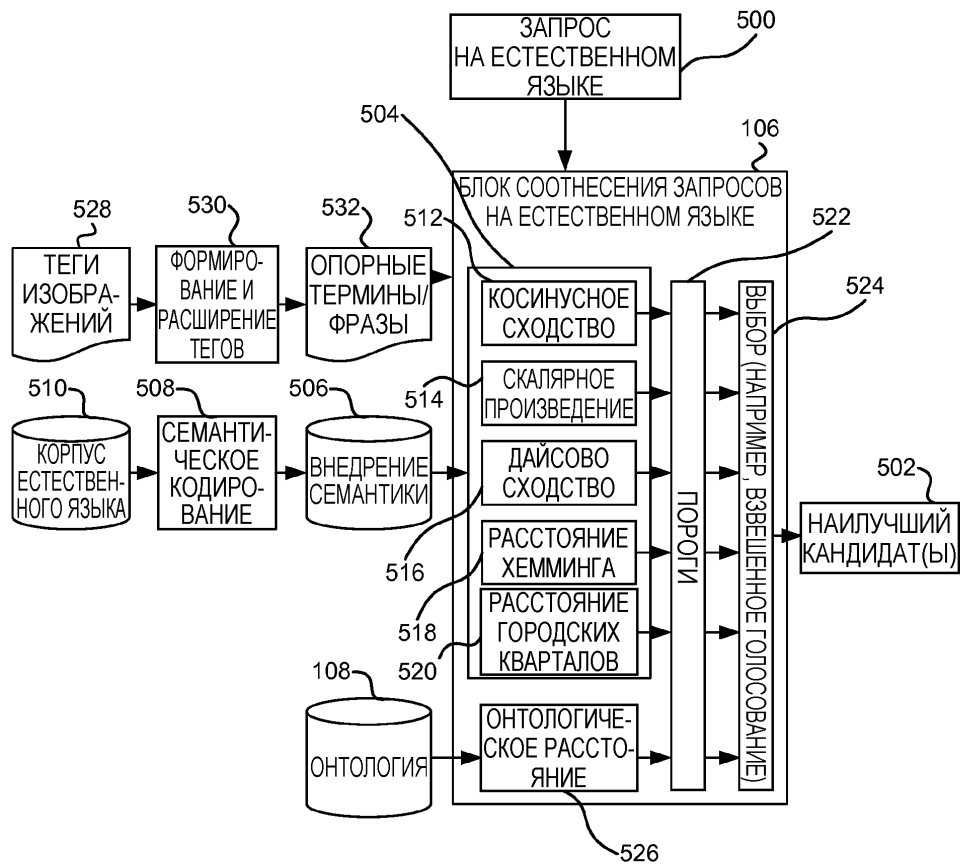
ФИГ. 3

4/9



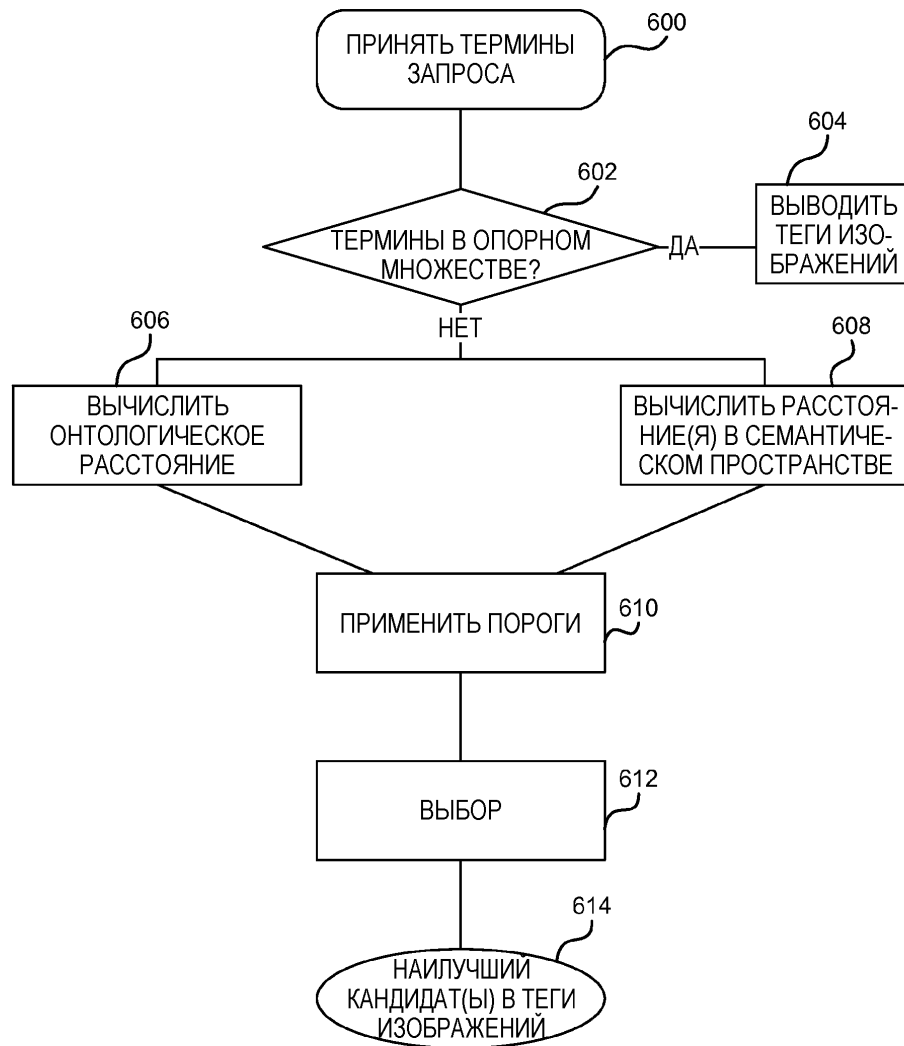
ФИГ. 4

5/9



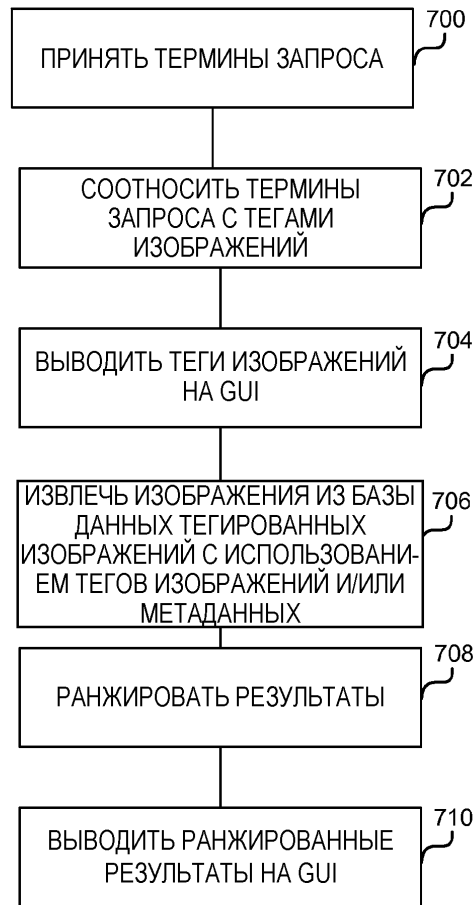
ФИГ. 5

6/9



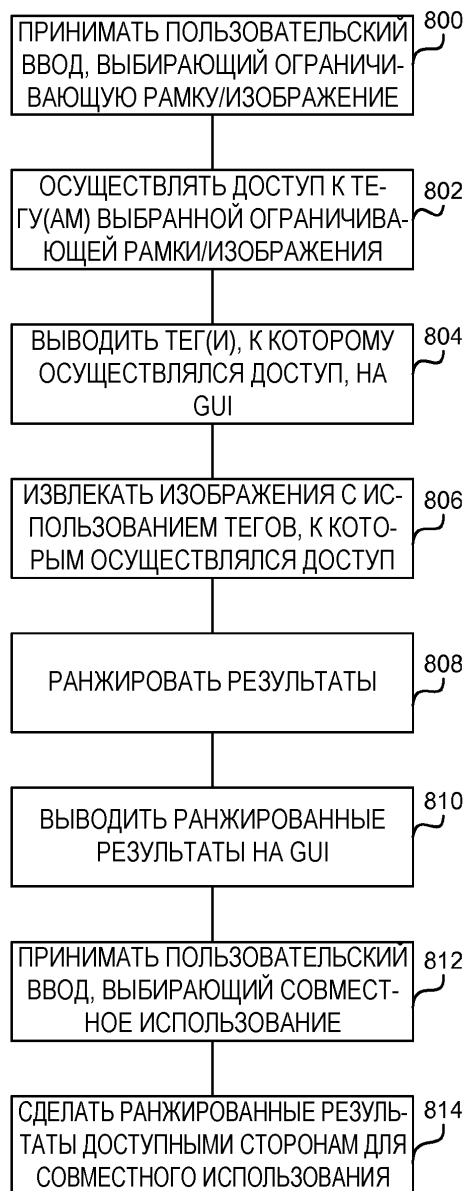
ФИГ. 6

7/9



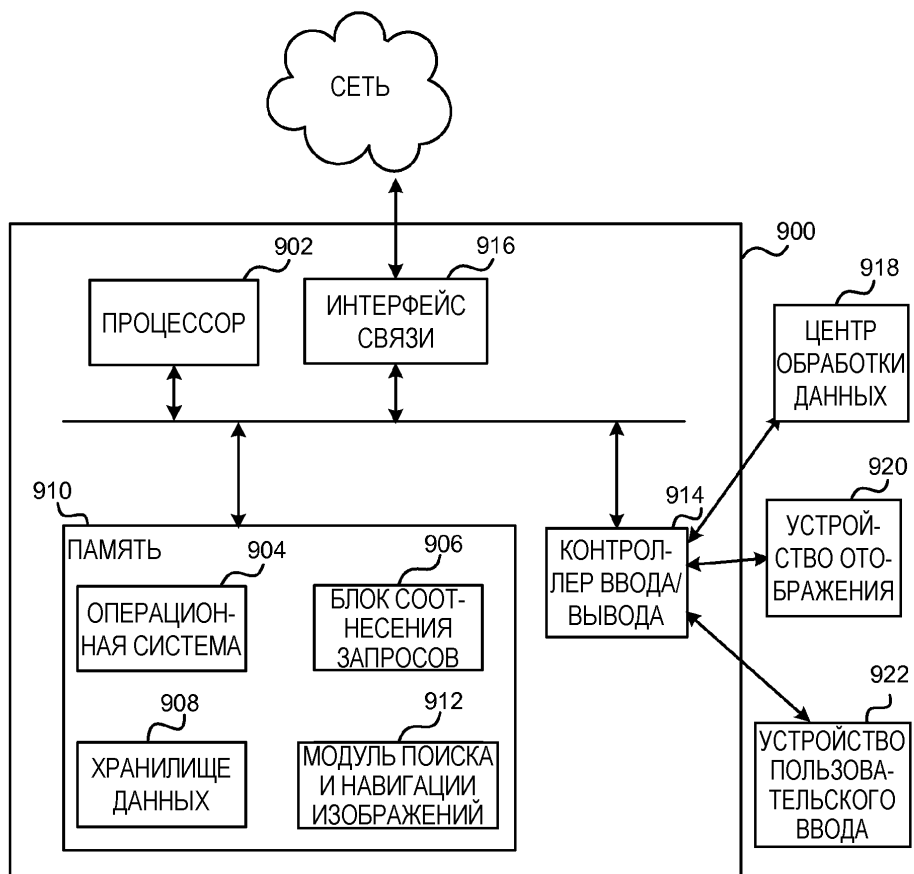
ФИГ. 7

8/9



ФИГ. 8

9/9



ФИГ. 9