(12) 按照专利合作条约所公布的国际申请

(19) 世界知识产权组织
国 际 局

(43) 国际公布日
2018 年 5 月 31 日 (31.05.2018)

WIPO | PCT

(10) 国际公布号
WO 2018/094952 A1

(51) 国际专利分类号:
*G06F 17/30* (2006.01)

(21) 国际申请号: PCT/CN2017/079624

(22) 国际申请日: 2017 年 4 月 6 日 (06.04.2017)

(25) 申请语言: 中文

(26) 公布语言: 中文

(30) 优先权:
201611036906.4 2016年11月22日 (22.11.2016) CN

(71) 申请人: 百度在线网络技术（北京）有限公司 (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.) [CN/CN]; 中国北京市海淀区上地十街10号百度大厦, Beijing 100085 (CN)。

(72) 发明人: 崔宝宏 (CUI, Baohong); 中国北京市海淀区上地十街10号百度大厦, Beijing 100085 (CN)。
王方舟 (WANG, Fangzhou); 中国北京市海淀区上地十街10号百度大厦, Beijing 100085 (CN)。

(74) 代理人: 北京汉昊知识产权代理事务所（普通合伙）(HANHOW INTELLECTUAL PROPERTY); 中国北京市东长安街1号东方广场西1区1111室, Beijing 100738 (CN)。

(81) 指定国（除另有指明，要求每一种可提供的国家保护）: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW。

(84) 指定国（除另有指明，要求每一种可提供的地区保护）: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), 欧亚 (AM, AZ, BY, KG, KZ, RU, TJ, TM), 欧洲 (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU,

(54) **Title:** CONTENT RECOMMENDATION METHOD AND APPARATUS

(54) **发明名称:** 一种内容推荐方法与装置

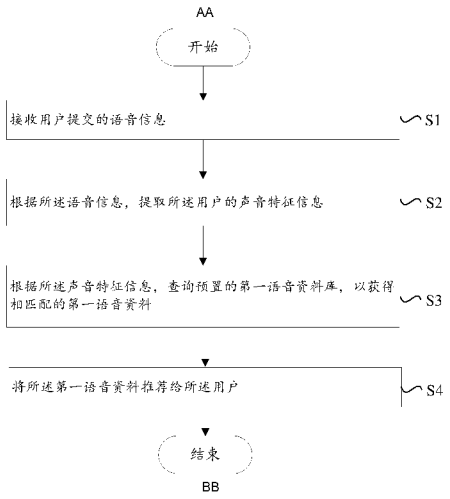



图 1

S1 Receiving voice information submitted by a user
S2 Extracting voice feature information of the user according to the voice information
S3 Querying a preset first voice database according to the voice feature information so as to obtain matching first voice data
S4 Recommending the first voice data to the user
AA Start
BB End

(57) **Abstract:** A content recommendation method and apparatus. The method comprises: a network device receiving voice information submitted by a user (S1); extracting voice feature information of the user according to the voice information (S2); querying a preset first voice database according to the voice feature information so as to obtain matching first voice data (S3); and recommending the first voice data to the user (S4). The present technical solution may recommend matching voice data such as songs and video clips to a user according to the voice feature information of the user. In current, gradually emerging interactive voice applications, it is possible to use interesting operation methods to increase the willingness of users to automatically use voice, which helps to cultivate the habit of using voice to carry out a search. For example, on the basis of the social and entertainment needs of the current voice audience, various voice feature information of a user, such as voice and audio, is automatically analyzed when the user searches for music or uses voice, and songs suitable for singing performance or a singer suitable for imitation are appropriately recommended to the user by means of big data analysis of voice.

WO 2018/094952 A1

[见续页]

IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG)。

**本国际公布**:

— 包括国际检索报告(条约第21条(3))。

**(57) 摘要**: 一种内容推荐方法与装置。其中，网络设备接收用户提交的语音信息（S1）；根据所述语音信息，提取所述用户的声音特征信息（S2）；根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料（S3）；将所述第一语音资料推荐给所述用户（S4）。本技术方案可以根据用户的声音特征信息为其推荐相匹配的语音资料，诸如歌曲、视频片段等。在如今逐步爆发的语音交互应用场景中，能够使用趣味性运营方法来增加用户自动使用语音的意愿，有利于培养用户使用语音进行搜索的习惯。例如，基于当前语音受众的社交及娱乐需求，当用户搜索音乐或者使用语音时自动分析用户的各项声音特征信息，诸如声色、音频等，通过语音大数据分析来适度推荐给用户适合演唱的歌曲或者适合模仿的歌手。

# 一种内容推荐方法与装置

本申请以一中国专利申请作为优先权申请，该中国专利申请的申请日为 2016 年 11 月 22 日，申请号为 201611036906.4，发明名称为“一种内容推荐方法与装置”。

## 技术领域

本发明涉及语音信息搜索技术领域，尤其涉及一种内容推荐技术。

## 背景技术

现有的各种语音搜索方案，均是基于对用户的语音内容的识别来进行。在识别出用户的语音内容后，搜索引擎为用户提供各种与该语音内容相应的内容。因此，这些语音搜索方案从其本质而言，仍是传统的内容搜索，只是用户的输入方式发生了变化，不再是文字输入，而可以采用语音输入。

## 发明内容

本发明的目的是提供一种内容推荐方法与装置。

根据本发明的一个方面，提供了一种内容推荐方法，其中，该方法包括以下步骤:

- 接收用户提交的语音信息;

- 根据所述语音信息，提取所述用户的声音特征信息;

- 根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料;

- 将所述第一语音资料推荐给所述用户。

根据本发明的另一个方面，还提供了一种内容推荐装置，其中，该装置包括:

- 用于接收用户提交的语音信息的装置;

- 用于根据所述语音信息，提取所述用户的声音特征信息的装置;

- 用于根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料的装置;

- 用于将所述第一语音资料推荐给所述用户的装置。

与现有技术相比，本发明可以根据用户的声音特征信息为其推荐相匹配的语音资料，诸如歌曲、视频片段等。本发明在如今逐步爆发的语音交互应用场景中，使用趣味性运营方法来增加用户自动使用语音的意愿，有利于培养用户使用语音进行搜索的习惯。例如，基于当前语音受众的社交及娱乐需求，当用户搜索音乐或者使用语音时自动分析用户的各项声音特征信息，诸如声色、音频等，本发明通过语音大数据分析来适度推荐给用户适合演唱的歌曲或者适合模仿的歌手。

## 附图说明

通过阅读参照以下附图所作的对非限制性实施例所作的详细描述，本发明的其它特征、目的和优点将会变得更明显:

图 1 示出根据本发明一个实施例的进行内容推荐的方法流程图;

图 2 示出根据本发明一个实施例的进行内容推荐的装置示意图。

附图中相同或相似的附图标记代表相同或相似的部件。

## 具体实施方式

在更加详细地讨论示例性实施例之前应当提到的是，一些示例性实施例被描述成作为流程图描绘的处理或方法。虽然流程图将各项操作描述成顺序的处理，但是其中的许多操作可以被并行地、并发地或者同时实施。此外，各项操作的顺序可以被重新安排。当其操作完成时所述处理可以被终止，但是还可以具有未包括在附图中的附加步骤。所述处理可以对应于方法、函数、规程、子例程、子程序等等。

在上下文中所称“计算机设备”，也称为“电脑”，是指可以通过运

行预定程序或指令来执行数值计算和/或逻辑计算等预定处理过程的智能电子设备，其可以包括处理器与存储器，由处理器执行在存储器中预存的程序指令来执行预定处理过程，或是由 ASIC、FPGA、DSP 等硬件执行预定处理过程，或是由上述二者组合来实现。计算机设备包括但不限于服务器、个人电脑（PC）、笔记本电脑、平板电脑、智能手机等。

所述计算机设备例如包括用户设备与网络设备。其中，所述用户设备包括但不限于个人电脑（PC）、笔记本电脑、移动终端等，所述移动终端包括但不限于智能手机、PDA 等；所述网络设备包括但不限于单个网络服务器、多个网络服务器组成的服务器组或基于云计算（Cloud Computing）的由大量计算机或网络服务器构成的云，其中，云计算是分布式计算的一种，由一群松散耦合的计算机集组成的一个超级虚拟计算机。其中，所述计算机设备可单独运行来实现本发明，也可接入网络并通过与网络中的其他计算机设备的交互操作来实现本发明。其中，所述计算机设备所处的网络包括但不限于互联网、广域网、城域网、局域网、VPN 网络等。

需要说明的是，所述用户设备、网络设备和网络等仅为举例，其他现有的或今后可能出现的计算机设备或网络如可适用于本发明，也应包含在本发明保护范围以内，并以引用方式包含于此。

本文后面所讨论的方法（其中一些通过流程图示出）可以通过硬件、软件、固件、中间件、微代码、硬件描述语言或者其任意组合来实施。当用软件、固件、中间件或微代码来实施时，用以实施必要任务的程序代码或代码段可以被存储在机器或计算机可读介质（比如存储介质）中。（一个或多个）处理器可以实施必要的任务。

这里所公开的具体结构和功能细节仅仅是代表性的，并且是用于描述本发明的示例性实施例的目的。但是本发明可以通过许多替换形式来具体实现，并且不应当被解释成仅仅受限于这里所阐述的实施例。

应当理解的是，虽然在这里可能使用了术语“第一”、“第二”等等

来描述各个单元，但是这些单元不应当受这些术语限制。使用这些术语仅仅是为了将一个单元与另一个单元进行区分。举例来说，在不背离示例性实施例的范围的情况下，第一单元可以被称为第二单元，并且类似地第二单元可以被称为第一单元。这里所使用的术语“和/或”包括其中一个或更多所列出的相关联项目的任意和所有组合。

应当理解的是，当一个单元被称为“连接”或“耦合”到另一单元时，其可以直接连接或耦合到所述另一单元，或者可以存在中间单元。与此相对，当一个单元被称为“直接连接”或“直接耦合”到另一单元时，则不存在中间单元。应当按照类似的方式来解释被用于描述单元之间的关系的其他词语（例如“处于...之间”相比于“直接处于...之间”，“与...邻近”相比于“与...直接邻近”等等）。

这里所使用的术语仅仅是为了描述具体实施例而不意图限制示例性实施例。除非上下文明确地另有所指，否则这里所使用的单数形式“一个”、“一项”还意图包括复数。还应当理解的是，这里所使用的术语“包括”和/或“包含”规定所陈述的特征、整数、步骤、操作、单元和/或组件的存在，而不排除存在或添加一个或更多其他特征、整数、步骤、操作、单元、组件和/或其组合。

还应当提到的是，在一些替换实现方式中，所提到的功能/动作可以按照不同于附图中标示的顺序发生。举例来说，取决于所涉及的功能/动作，相继示出的两幅图实际上可以基本上同时执行或者有时可以按照相反的顺序来执行。

本发明可由计算机设备实现。典型地，本发明可由网络设备实现，但本领域技术人员应能理解，本发明的方案同样可由用户设备实现，只要其具备本发明所要求的计算/处理能力。为便于说明，本说明书中以下多用网络设备的实现进行举例，但本领域技术人员应能理解，此等举例仅用于解释说明本发明之目的，而不应被理解为对本发明的任何限制。

下面结合附图对本发明作进一步详细描述。

图 1 示出根据本发明一个实施例的方法流程图，其中具体示出一种内容推荐过程。

如图 1 所示，在步骤 S1 中，网络设备接收用户提交的语音信息；在步骤 S2 中，网络设备根据所述语音信息，提取所述用户的声音特征信息；在步骤 S3 中，网络设备根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料；在步骤 S4 中，网络设备将所述第一语音资料推荐给所述用户。

具体地，本发明典型地适用于搜索场景或音乐交互类场景。

其中，推荐给用户的第一语音资料可以是任何通过声音特征信息来进行匹配的语音资料，包括但不限于歌曲、视频片段。

在此，本领域技术人员应能理解，“语音资料”的表述并不意在限定其表现形式，如各种音频资料，而是可以包括任何具有音频内容的资料，因此其不仅可以包括歌曲，也可以包括视频片段。

例如，当用户进行语音搜索时，网络设备根据用户的声音特征信息为其推荐第一语音资料，例如，适合该用户模仿的视频片段。此时，网络设备还需判断用户是否具有娱乐需求，如对用户提交的语音内容进行识别，以分析语音传递的内容是否具备视频、音乐等相关娱乐需求，例如通过预置的关键词表来识别娱乐需求，具体如该语音内容是否为一句歌词、一段台词等。

又如，用户当前使用音乐类 APP，当用户在该 APP 内提交一段语音信息时，网络设备可以根据该用户的声音特征信息为其推荐第一语音资料，例如，适合该用户唱的歌曲。

优选地，本发明的内容推荐方案可以具体表现为一“匹配”功能，该“匹配”功能可以为用户提供与其声音特征信息相匹配的第一语音资料。

例如，用户端可以向用户呈现一“匹配”功能按钮，如放置于语音搜索框的右侧或设置于音乐类 APP 的界面中的特定位置。当用户点击该

“匹配”功能按钮后，用户端即将用户提交的语音信息发送至网络设备，以由网络设备根据该用户的声音特征信息为其推荐相应的第一语音资料。

可替代地，在用户提交语音信息之后，“匹配”功能可以被直接提示给用户，以提醒用户可以对之前提交的语音信息做语音资料推荐。

在此，本领域技术人员应能理解，“匹配”功能按钮仅用于举例说明本发明之目的，其在实际应用中的表现形式可以不同，例如可以呈现为“主打歌”功能按钮、“大明星”功能按钮等，只要此等功能按钮意在为用户提供如本发明一般的内容推荐方案，其即应落入本发明的专利保护范围之内。

此外，本领域技术人员还应能理解，本发明中的“网络设备”并不限于一台计算机设备，其可以是多台计算机设备，典型地如多台服务器，这些服务器相互配合来实现本发明的内容推荐方案。

例如，接口服务器负责与用户端以及其他服务器的交互，其中，与用户端的交互如接收用户端发送的语音信息以及向用户端返回推荐的第一语音资料，与匹配服务器的交互如将语音信息提交至匹配服务器并从匹配服务器获得匹配的第一语音资料；匹配服务器负责对接口服务器提交的语音信息在第一语音资料库中进行匹配查询，以获得相匹配的第一语音资料。这些服务器相对于其他外部设备而言，被视为一个整体，并被通常整体视为“网络设备”。

仍返回图 1，在步骤 S1 中，网络设备接收用户提交的语音信息。

在此，本发明需要网络设备与用户端进行交互来实现。其中，用户端诸如各种在用户设备中运行的功能实体，具体如 PC 中的网页、手机中的 APP 等。

用户通过其用户设备提交一段语音信息，用户设备将该语音信息发送至网络设备，从而网络设备接收该用户提交的语音信息。

在步骤 S2 中，网络设备根据用户提交的语音信息，提取该用户的声音特征信息。

在此，用户的声音特征信息可以至少从以下两个维度来定义:

1）感知维度

感知维度意在从对声音的感知特征来进行定义。声音特征信息在该维度下具体诸如音调特征信息、音高特征信息、旋律特征信息、节奏特征信息、音频特征信息、语速特征信息等。

2）声学维度

声学维度意在从声音在声学角度定义的各种特征来进行表达。声音特征信息在该维度下具体诸如能量特征信息、过零率、LPC（线性预测系数）参数信息等。其中，过零率意指声音信号的符号变化的比率，如声音信号从正数变成负数或反向。

其中，网络设备可以采用任何已知或将来可行的技术对上述声音特征信息进行提取，本发明对此不作限定。

在步骤 S3 中，网络设备根据用户的声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料。

在此，网络设备可以预先建立第一语音资料库，其中存储有各语音资料的各项声音特征信息。可替代地，第一语音资料库也可以由其他设备来建立及维护，网络设备对其具有访问权限即可。

在一个示例性的应用场景中，第一语音资料库中例如存储有各明星（包括歌手和演员）的歌曲或参演视频片段，并且这些语音资料均已标注有各项声音特征信息。网络设备根据用户所识别出的声音特征信息在该第一语音资料库中进行匹配查询，以获得相匹配的歌曲或视频片段。

在步骤 S4 中，网络设备将所匹配的第一语音资料推荐给用户。

在此，网络设备将所匹配的如歌曲或视频片段等推荐给用户，从而本发明可以为用户提供其适合演唱或模仿的歌曲或视频配音，以为用户定制自己的“主打歌”或“成名歌”并帮助用户找到其适合模仿的歌手。这在商业应用中是有实际意义的，并显著提升了用户粘性及使用时长。

本发明通过引导用户通过说话、大声喊、唱两句、读两句诗等各类语音输入来自动匹配出歌手、诗人、名人等。进一步地，本发明还可以

继续接收用户对推荐歌曲的演唱，进而对用户自己的演唱相对与原唱来进行打分，并通过分享到好友、邀请好友一起来 PK 等提升趣味性帮助用户找到自己最适合唱的歌曲，打造用户自己的模仿成名曲，从而逐步引导用户形成使用语音搜索习惯以及与搜索引擎的二次交互，提供用户粘性及使用时长。据此，本发明还可以通过提供分享功能来从搜索引擎衍生至社交领域。

根据本发明的一个优选实施例，图 1 所示的内容推荐过程还包括一步骤 S5。在步骤 S5 中，网络设备根据用户语音信息的内容，查询预置的第二语音资料库，以获得相匹配的第二语音资料；随后，网络设备将所匹配的第二语音资料推荐给用户。

在此，网络设备对第二语音资料的推荐步骤可以与上述步骤 S4 的推荐合并，从而步骤 S5 发生在步骤 S4 之前，网络设备可以在步骤 S4 中向用户一并推荐第一语音资料和第二语音资料。

可替代地，网络设备对第二语音资料的推荐也可以独立于步骤 S4 中对第一语音资料的推荐。例如，网络设备在执行步骤 S3 和 S4 的操作时，可以在这两个步骤之前、之后或与这两个步骤同时来执行步骤 S5 对第二语音资料的获取以及步骤 S6 对第二语音资料的推荐。

在此，第二语音资料库中至少存储有语音资料的内容信息或内容摘要信息。据此，网络设备在识别出用户的语音信息中的内容后，在该第二语音资料库中进行匹配查询，以获得内容相同的第二语音资料，如内容相同但声音不同的语音资料，具体如使用不用方言来讲的同一内容。

例如，用户提交的语音信息为普通话的“你好”，网络设备识别出其中的内容信息为“你好”并在第二语音资料库中进行匹配查询，以获得各种方言的“你好”，诸如上海话、广东话的“你好”等。

优选地，第一语音资料库可以与第二语音资料库集成在一起。也即，本发明中仅使用一个语音资料库来进行第一语音资料和第二语音资料的匹配查询，从而该语音资料库中需至少存储有各语音资料的各项声音特征信息和内容信息。

图 2 示出根据本发明一个实施例的装置示意图，其中具体示出一种内容推荐装置。如图 2 所示，内容推荐装置 20 装置于网络设备 200 中，并具体包括接收装置 21、提取装置 22、匹配装置 23 和推荐装置 24。

其中，接收装置 21 接收用户提交的语音信息；提取装置 22 根据所述语音信息，提取所述用户的声音特征信息；匹配装置 23 根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料；推荐装置 24 将所述第一语音资料推荐给所述用户。

具体地，本发明典型地适用于搜索场景或音乐交互类场景。

其中，推荐给用户的第一语音资料可以是任何通过声音特征信息来进行匹配的语音资料，包括但不限于歌曲、视频片段。

在此，本领域技术人员应能理解，“语音资料”的表述并不意在限定其表现形式，如各种音频资料，而是可以包括任何具有音频内容的资料，因此其不仅可以包括歌曲，也可以包括视频片段。

例如，当用户进行语音搜索时，内容推荐装置 20 根据用户的声音特征信息为其推荐第一语音资料，例如，适合该用户模仿的视频片段。此时，内容推荐装置 20 或网络设备中的其他装置还需判断用户是否具有娱乐需求，如对用户提交的语音内容进行识别，以分析语音传递的内容是否具备视频、音乐等相关娱乐需求，例如通过预置的关键词表来识别娱乐需求，具体如该语音内容是否为一句歌词、一段台词等

又如，用户当前使用音乐类 APP，当用户在该 APP 内提交一段语音信息时，内容推荐装置 20 可以根据该用户的声音特征信息为其推荐第一语音资料，例如，适合该用户唱的歌曲。

优选地，本发明的内容推荐方案可以具体表现为一“匹配”功能，该“匹配”功能可以为用户提供与其声音特征信息相匹配的第一语音资料。

例如，用户端可以向用户呈现一“匹配”功能按钮，如放置于语音搜索框的右侧或设置于音乐类 APP 的界面中的特定位置。当用户点击该“匹配”功能按钮后，用户端即将用户提交的语音信息发送至网络设备，以由内容推荐装置 20 根据该用户的声音特征信息为其推荐相应的第一

语音资料。

可替代地，在用户提交语音信息之后，“匹配”功能可以被直接提示给用户，以提醒用户可以对之前提交的语音信息做语音资料推荐。

在此，本领域技术人员应能理解，“匹配”功能按钮仅用于举例说明本发明之目的，其在实际应用中的表现形式可以不同，例如可以呈现为“主打歌”功能按钮、“大明星”功能按钮等，只要此等功能按钮意在为用户提供如本发明一般的内容推荐方案，其即应落入本发明的专利保护范围之内。

此外，本领域技术人员还应能理解，本发明中的“网络设备”并不限于一台计算机设备，其可以是多台计算机设备，典型地如多台服务器，这些服务器相互配合来实现本发明的内容推荐方案。

例如，接口服务器负责与用户端以及其他服务器的交互，其中，与用户端的交互如接收用户端发送的语音信息以及向用户端返回推荐的第一语音资料，与匹配服务器的交互如将语音信息提交至匹配服务器并从匹配服务器获得匹配的第一语音资料；匹配服务器负责对接口服务器提交的语音信息在第一语音资料库中进行匹配查询，以获得相匹配的第一语音资料。这些服务器相对于其他外部设备而言，被视为一个整体，并被通常整体视为“网络设备”。

仍返回图 2，接收装置 21 接收用户提交的语音信息。

在此，本发明需要网络设备与用户端进行交互来实现。其中，用户端诸如各种在用户设备中运行的功能实体，具体如 PC 中的网页、手机中的 APP 等。

用户通过其用户设备提交一段语音信息，用户设备将该语音信息发送至网络设备，从而接收装置 21 接收该用户提交的语音信息。

随后，提取装置 22 根据用户提交的语音信息，提取该用户的声音特征信息。

在此，用户的声音特征信息可以至少从以下两个维度来定义：

1）感知维度

感知维度意在从对声音的感知特征来进行定义。声音特征信息在该维度下具体诸如音调特征信息、音高特征信息、旋律特征信息、节奏特征信息、音频特征信息、语速特征信息等。

2）声学维度

声学维度意在从声音在声学角度定义的各种特征来进行表达。声音特征信息在该维度下具体诸如能量特征信息、过零率、LPC（线性预测系数）参数信息等。其中，过零率意指声音信号的符号变化的比率，如声音信号从正数变成负数或反向。

其中，提取装置 22 可以采用任何已知或将来可行的技术对上述声音特征信息进行提取，本发明对此不作限定。

接着，匹配装置 23 根据用户的声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料。

在此，网络设备可以预先建立第一语音资料库，其中存储有各语音资料的各项声音特征信息。可替代地，第一语音资料库也可以由其他设备来建立及维护，匹配装置 23 对其具有访问权限即可。

在一个示例性的应用场景中，第一语音资料库中例如存储有各明星（包括歌手和演员）的歌曲或参演视频片段，并且这些语音资料均已标注有各项声音特征信息。匹配装置 23 根据用户所识别出的声音特征信息在该第一语音资料库中进行匹配查询，以获得相匹配的歌曲或视频片段。

随后，推荐装置 24 将所匹配的第一语音资料推荐给用户。

在此，推荐装置 24 将所匹配的如歌曲或视频片段等推荐给用户，从而本发明可以为用户提供其适合演唱或模仿的歌曲或视频配音，以为用户定制自己的“主打歌”或“成名歌”并帮助用户找到其适合模仿的歌手。。这在商业应用中是有实际意义的，并显著提升了用户粘性及使用时长。

本发明通过引导用户通过说话、大声喊、唱两句、读两句诗等各类语音输入来自动匹配出歌手、诗人、名人等。进一步地，本发明还可以继续接收用户对推荐歌曲的演唱，进而对用户自己的演唱相对与原唱来

进行打分，并通过分享到好友、邀请好友一起来 PK 等提升趣味性帮助用户找到自己最适合唱的歌曲，打造用户自己的模仿成名曲，从而逐步引导用户形成使用语音搜索习惯以及与搜索引擎的二次交互，提供用户粘性及使用时长。据此，本发明还可以通过提供分享功能来从搜索引擎衍生至社交领域。

根据本发明的一个优选实施例，图 2 所示的内容推荐装置 20 还包括另一匹配装置（以下称为第二匹配装置）。第二匹配装置根据用户语音信息的内容，查询预置的第二语音资料库，以获得相匹配的第二语音资料；随后，内容推荐装置 20 将所匹配的第二语音资料推荐给用户。

在此，内容推荐装置 20 对第二语音资料的推荐可以与上述对第一语音资料的推荐合并，从而第二匹配装置所执行的操作发生在推荐装置 24 所执行的操作之前，推荐装置 24 可以向用户一并推荐第一语音资料和第二语音资料。

可替代地，内容推荐装置 20 对第二语音资料的推荐也可以独立于推荐装置 24 对第一语音资料的推荐。例如，在匹配装置 23 和推荐装置 24 执行其各自的操作时，第二匹配装置可以在这两个装置的操作之前、之后或与这两个装置执行操作的同时，来执行对第二语音资料的获取，从而内容推荐装置 20 还需包括另一推荐装置（以下称为第二推荐装置）。来执行对第二语音资料的推荐。

优选地，第二匹配装置仍可以与匹配装置 23 集成在一起，以及第二推荐装置与推荐装置 24 集成在一起。此时，集成后的各装置分别执行两次匹配操作或两次推荐操作。

在此，第二语音资料库中至少存储有语音资料的内容信息或内容摘要信息。据此，在第二匹配装置或内容推荐装置 20 中的其他装置识别出用户的语音信息中的内容后，第二匹配装置在该第二语音资料库中进行匹配查询，以获得内容相同的第二语音资料，如内容相同但声音不同的语音资料，具体如使用不用方言来讲的同一内容。

例如，用户提交的语音信息为普通话的“你好”，第二匹配装置在第

二语音资料库中进行匹配查询，以获得各种方言的“你好”，诸如上海话、广东话的“你好”等。

优选地，第一语音资料库可以与第二语音资料库集成在一起。也即，本发明中仅使用一个语音资料库来进行第一语音资料和第二语音资料的匹配查询，从而该语音资料库中需至少存储有各语音资料的各项声音特征信息和内容信息。

需要注意的是，本发明可在软件和/或软件与硬件的组合体中被实施，例如，本发明的各个装置可采用专用集成电路（ASIC）或任何其他类似硬件设备来实现。在一个实施例中，本发明的软件程序可以通过处理器执行以实现上文所述步骤或功能。同样地，本发明的软件程序（包括相关的数据结构）可以被存储到计算机可读记录介质中，例如，RAM 存储器，磁或光驱动器或软磁盘及类似设备。另外，本发明的一些步骤或功能可采用硬件来实现，例如，作为与处理器配合从而执行各个步骤或功能的电路。

对于本领域技术人员而言，显然本发明不限于上述示范性实施例的细节，而且在不背离本发明的精神或基本特征的情况下，能够以其他的具体形式实现本发明。因此，无论从哪一点来看，均应将实施例看作是示范性的，而且是非限制性的，本发明的范围由所附权利要求而不是上述说明限定，因此旨在将落在权利要求的等同要件的含义和范围内的所有变化涵括在本发明内。不应将权利要求中的任何附图标记视为限制所涉及的权利要求。此外，显然“包括”一词不排除其他单元或步骤，单数不排除复数。系统权利要求中陈述的多个单元或装置也可以由一个单元或装置通过软件或者硬件来实现。第一，第二等词语用来表示名称，而并不表示任何特定的顺序。

# 权 利 要 求 书

1. 一种内容推荐方法，其中，该方法包括以下步骤:

- 接收用户提交的语音信息;

- 根据所述语音信息，提取所述用户的声音特征信息;

- 根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料;

- 将所述第一语音资料推荐给所述用户。

2. 根据权利要求 1 所述的方法，其中，该方法还包括:

- 根据所述语音信息的内容，查询预置的第二语音资料库，以获得相匹配的第二语音资料;

其中，所述推荐包括对所述第二语音资料的推荐。

3. 根据权利要求 2 所述的方法，其中，所述第一语音资料库与所述第二语音资料库集成在一起。

4. 根据权利要求 1 至 3 中任一项所述的方法，其中，所述声音特征信息基于以下至少任一维度来提取:

- 感知维度;

- 声学维度。

5. 根据权利要求 1 至 4 中任一项所述的方法，其中，所述第一语音资料包括以下至少任一项:

- 歌曲;

- 视频片段。

6. 根据权利要求 1 至 5 中任一项所述的方法，其中，所述接收用户提交的步骤用于搜索场景或音乐交互类场景。

7. 一种内容推荐装置，其中，该装置包括:

- 用于接收用户提交的语音信息的装置;

- 用于根据所述语音信息，提取所述用户的声音特征信息的装置;

- 用于根据所述声音特征信息，查询预置的第一语音资料库，以获

得相匹配的第一语音资料的装置;

- 用于将所述第一语音资料推荐给所述用户的装置。

8. 根据权利要求 7 所述的装置，其中，该装置还包括:

- 用于根据所述语音信息的内容，查询预置的第二语音资料库，以获得相匹配的第二语音资料的装置;

其中，所述推荐包括对所述第二语音资料的推荐。

9. 根据权利要求 8 所述的装置，其中，所述第一语音资料库与所述第二语音资料库集成在一起。

10. 根据权利要求 7 至 9 中任一项所述的装置，其中，所述声音特征信息基于以下至少任一维度来提取:

- 感知维度;

- 声学维度。

11. 根据权利要求 7 至 10 中任一项所述的装置，其中，所述第一语音资料包括以下至少任一项:

- 歌曲;

- 视频片段。

12. 根据权利要求 7 至 11 中任一项所述的装置，其中，所述接收用户提交的操作用于搜索场景或音乐交互类场景。

13. 一种计算机可读存储介质，所述计算机可读存储介质包括计算机指令，当所述计算机指令被执行时，如权利要求 1 至 6 中任一项所述的方法被执行。

14. 一种计算机程序产品，当所述计算机程序产品被执行时，如权利要求 1 至 6 中任一项所述的方法被执行。

15. 一种计算机设备，所述计算机设备包括存储器和处理器，所述存储器中存储有计算机指令，所述处理器被配置来通过执行所述计算机指令以执行如权利要求 1 至 6 中任一项所述的方法。

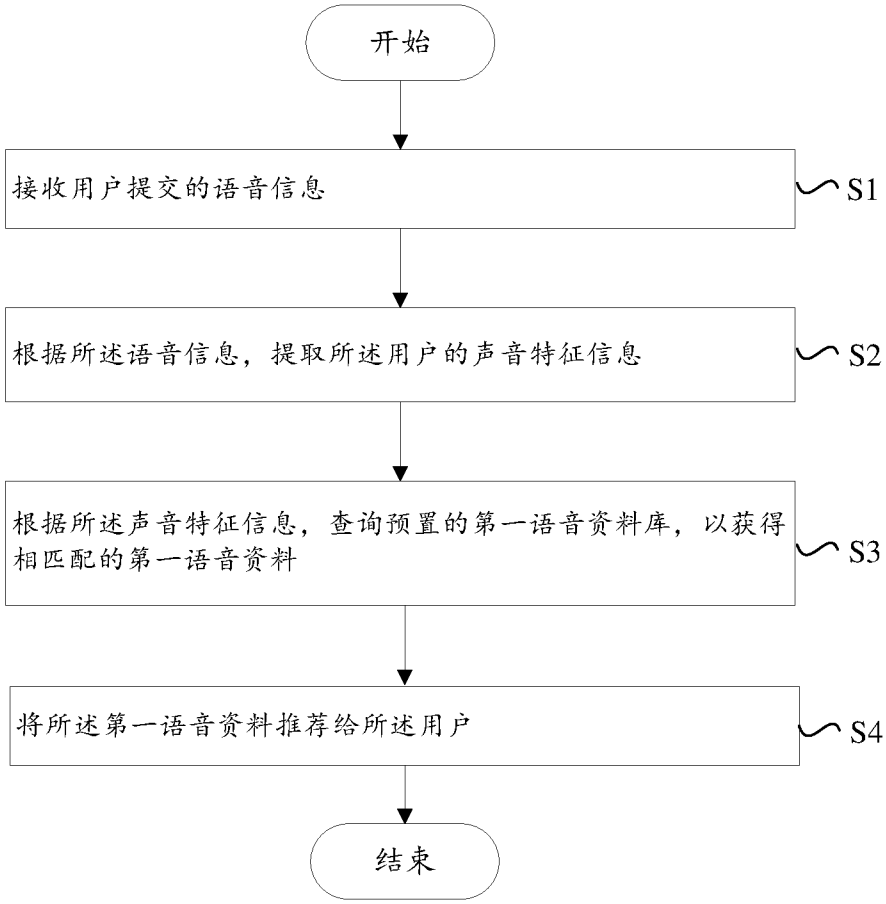

开始
接收用户提交的语音信息
S1
根据所述语音信息，提取所述用户的声音特征信息
S2
根据所述声音特征信息，查询预置的第一语音资料库，以获得相匹配的第一语音资料
S3
将所述第一语音资料推荐给所述用户
S4
结束

图 1

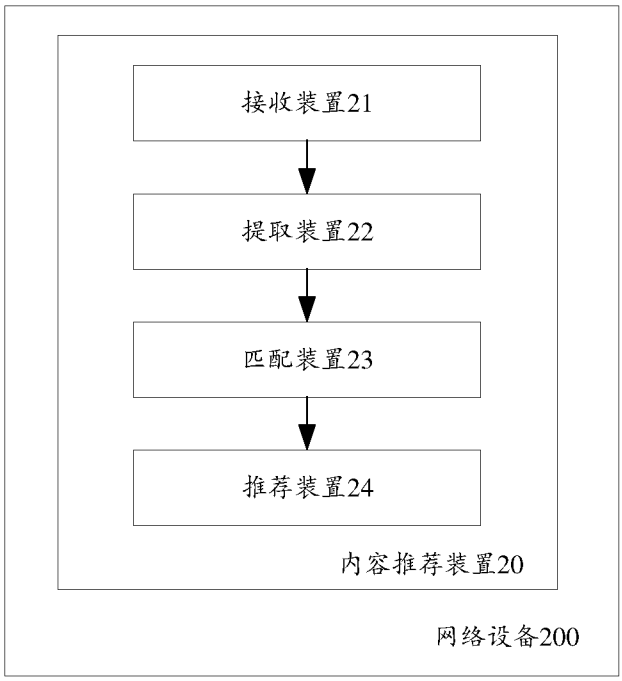



图 2

# INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2017/079624

**A. CLASSIFICATION OF SUBJECT MATTER**

G06F 17/30 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06F

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

CNPAT; CNKI; WPI; EPODOC: 语音, 声音, 识别, 模型, 比对, 匹配, 特征, 推荐, music, sound, character, module, song, voice, model, match+, similarity, recognition, audio, identify+

C. DOCUMENTS CONSIDERED TO BE RELEVANT

| Category* | Citation of document, with indication, where appropriate, of the relevant passages | Relevant to claim No. |
|---|---|---|
| X | CN 102404278 A (SHENGYUE INFORMATION TECHNOLOGY (SHANGHAI) CO., LTD.), 04 April 2012 (04.04.2012), description, paragraphs [0023]-[0038] | 1-15 |
| A | CN 102347060 A (HONGFUJIN PRECISION INDUSTRY (SHENZHEN) CO., LTD. et al.), 08 February 2012 (08.02.2012), entire document | 1-15 |
| A | CN 103685520 A (SHENZHEN TCL NEW TECHNOLOGY CO., LTD.), 26 March 2014 (26.03.2014), entire document | 1-15 |

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

| Date of the actual completion of the international search | Date of mailing of the international search report |
|---|---|
| 26 July 2017 | 10 August 2017 |

| Name and mailing address of the ISA | Authorized officer |
|---|---|
| State Intellectual Property Office of the P. R. China<br>No. 6, Xitucheng Road, Jimenqiao<br>Haidian District, Beijing 100088, China<br>Facsimile No. (86-10) 62019451 | LIU, Man<br>Telephone No. (86-10) 62414025 |

# INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2017/079624

| Patent Documents referred in the Report | Publication Date | Patent Family | Publication Date |
|---|---|---|---|
| CN 102404278 A | 04 April 2012 | None | |
| CN 102347060 A | 08 February 2012 | None | |
| CN 103685520 A | 26 March 2014 | None | |

# 国际检索报告

国际申请号
PCT/CN2017/079624

A. 主题的分类

G06F 17/30(2006.01)i

按照国际专利分类(IPC)或者同时按照国家分类和IPC两种分类

B. 检索领域

检索的最低限度文献(标明分类系统和分类号)

G06F

包含在检索领域中的除最低限度文献以外的检索文献

在国际检索时查阅的电子数据库(数据库的名称，和使用的检索词（如使用））

CNPAT;CNKI;WPI;EPODOC:语音，声音，识别，模型，比对，匹配，特征，推荐，music，sound，character，module，song，voice，model，match+，similarity，recognition，audio，identify+

C. 相关文件

| 类 型* | 引用文件，必要时，指明相关段落 | 相关的权利要求 |
|---|---|---|
| X | CN 102404278 A (盛乐信息技术上海有限公司) 2012年 4月 4日 (2012-04-04)<br>说明书第[0023]-[0038]段 | 1-15 |
| A | CN 102347060 A (鸿富锦精密工业深圳有限公司 等) 2012年 2月 8日 (2012-02-08)<br>全文 | 1-15 |
| A | CN 103685520 A (深圳TCL新技术有限公司) 2014年 3月 26日 (2014-03-26)<br>全文 | 1-15 |

☐ 其余文件在C栏的续页中列出。

☑ 见同族专利附件。

* 引用文件的具体类型：

“A” 认为不特别相关的表示了现有技术一般状态的文件

“E” 在国际申请日的当天或之后公布的在先申请或专利

“L” 可能对优先权要求构成怀疑的文件，或为确定另一篇引用文件的公布日而引用的或者因其他特殊理由而引用的文件（如具体说明的）

“O” 涉及口头公开、使用、展览或其他方式公开的文件

“P” 公布日先于国际申请日但迟于所要求的优先权日的文件

“T” 在申请日或优先权日之后公布，与申请不相抵触，但为了理解发明之理论或原理的在后文件

“X” 特别相关的文件，单独考虑该文件，认定要求保护的发明不是新颖的或不具有创造性

“Y” 特别相关的文件，当该文件与另一篇或者多篇该类文件结合并且这种结合对于本领域技术人员为显而易见时，要求保护的发明不具有创造性

“&” 同族专利的文件

| 国际检索实际完成的日期 | 国际检索报告邮寄日期 |
|---|---|
| 2017年 7月 26日 | 2017年 8月 10日 |

| ISA/CN的名称和邮寄地址 | 受权官员 |
|---|---|
| 中华人民共和国国家知识产权局(ISA/CN)<br>中国北京市海淀区蓟门桥西土城路6号 100088 | 刘曼 |
| 传真号 (86-10)62019451 | 电话号码 (86-10)62414025 |

# 国际检索报告
## 关于同族专利的信息

国际申请号
PCT/CN2017/079624

| 检索报告引用的专利文件 | 公布日（年/月/日） | 同族专利 | 公布日（年/月/日） |
|---|---|---|---|
| CN 102404278 A | 2012年 4月 4日 | 无 | |
| CN 102347060 A | 2012年 2月 8日 | 无 | |
| CN 103685520 A | 2014年 3月 26日 | 无 | |