

(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
2014 年 4 月 24 日 (24.04.2014)



(10) 国际公布号
WO 2014/059863 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2013/084263
- (22) 国际申请日: 2013 年 9 月 26 日 (26.09.2013)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201210398027.1 2012 年 10 月 18 日 (18.10.2012) CN
- (71) 申请人: 腾讯科技(深圳)有限公司 (TENCENT TECHNOLOGY (SHENZHEN) COMPANY LIMITED) [CN/CN]; 中国广东省深圳福田区振兴路赛格科技园 2 栋东 403 室, Guangdong 518000 (CN)。
- (72) 发明人: 刘刚 (LIU, Gang); 中国广东省深圳福田区振兴路赛格科技园 2 栋东 403 室, Guangdong 518000 (CN)。
- (74) 代理人: 深圳翼盛智成知识产权事务所(普通合伙) (ESSEN PATENT&TRADEMARK AGENCY); 中国广东省深圳福田区天安数码城数码时代大厦 A 座 1409 室, Guangdong 518040 (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, 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, 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, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO,

[见续页]

(54) Title: SUBTITLE QUERYING METHOD, ELECTRONIC DEVICE AND STORAGE MEDIUM

(54) 发明名称: 字幕查询方法、电子设备及存储介质

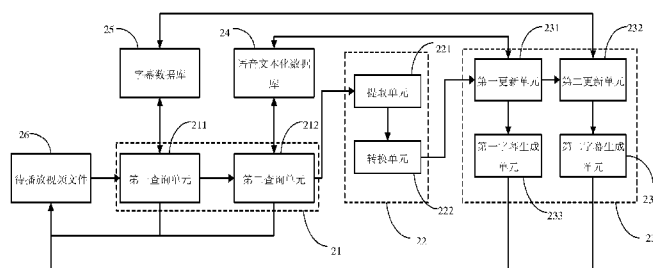


图 2 / FIG. 2

- | | |
|--------------------------|-------------------------------------|
| 211 FIRST QUERYING UNIT | 233 FIRST SUBTITLE GENERATING UNIT |
| 212 SECOND QUERYING UNIT | 234 SECOND SUBTITLE GENERATING UNIT |
| 221 EXTRACTING UNIT | 24 VOICE TEXTUALIZATION DATABASE |
| 222 CONVERTING UNIT | 26 TO-BE-PLAYED VIDEO FILE |
| 231 FIRST UPDATE UNIT | 25 SUBTITLE DATABASE |
| 232 SECOND UPDATE UNIT | |

(57) Abstract: The present invention provides a subtitle querying method, an electronic device and storage medium. The subtitle querying method comprises the steps of: obtaining feature information of a to-be-played video file, and performing subtitle querying on the to-be-played video file in a subtitle database according to the feature information, so as to generate a querying result; performing voice textualization processing on the to-be-played video file according to the querying result; and updating the subtitle database according to the textualization subtitle generated according to the voice textualization processing, and using the updated subtitle in the subtitle database as subtitle of the to-be-played video file. The present invention further provides an electronic device and a storage medium. The subtitle querying method, the electronic device and the storage medium according to the present invention are used to query the subtitle based on audio information identification, which can improve subtitle matching accuracy and reduce an error rate.

(57) 摘要:

[见续页]



RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, **本国际公布:**

CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, — 包括国际检索报告(条约第 21 条(3))。
TG)。

本发明提供一种字幕查询方法、电子设备及存储介质。该字幕查询方法包括步骤：获取待播放视频文件的特征信息，并根据特征信息，在字幕数据库中对待播放视频文件进行字幕查询，以产生查询结果；根据查询结果，对待播放视频文件进行语音文本化处理；以及根据语音文本化处理生成的文本化字幕，对字幕数据库进行更新，并将字幕数据库中进行更新的字幕作为待播放视频文件的字幕。本发明还提供一种电子设备及存储介质。本发明的字幕查询方法、电子设备及存储介质基于音频信息识别对字幕进行查询，可提高字幕匹配命中率以及降低误差率。

发明名称：字幕查询方法、电子设备及存储介质

技术领域

- [1] 本发明涉及互联网技术领域，特别是涉及一种字幕查询方法、电子设备及存储介质。

背景技术

- [2] 随着多媒体技术和互联网技术的迅速发展，视频作为一种媒体类型在人们的生活、教育以及娱乐等方面日益成为不可或缺的信息载体。在视频文件，尤其是外语视频文件中，字幕是非常重要的部分，因为有些用户的外语水平有限，在看外文电影时难以融入进剧情；而国语配音又很难达到原音的效果，因此使用字幕是解决这个问题最有效以及最便捷的方法。
- [3] 目前的主流播放器都提供将视频文件和在线字幕库进行匹配的功能，但匹配成功率较低或误差率较大，导致用户的观看字幕的体验很差。造成这种问题的主要原因是：没有足够丰富的在线字幕库供位于客户端的播放器使用，从而导致播放器的字幕匹配命中率不高或误差率较大。

对发明的公开

技术问题

- [4] 本发明的目的在于提供一种基于音频信息识别的字幕查询方法、电子设备及存储介质，可提高字幕匹配命中率以及降低误差率，以解决现有的视频播放器的字幕匹配命中率不高或误差率较大的技术问题。

问题的解决方案

技术解决方案

- [5] 为解决上述问题，本发明提供的技术方案如下：
- [6] 本发明提供一种字幕查询方法，其包括步骤：
- [7] 获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果；
- [8] 根据所述查询结果，对所述待播放视频文件进行语音文本化处理；以及

- [9] 根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕；
- [10] 其中所述语音文本化处理包括步骤：
- [11] 提取所述待播放视频文件的音频信息；以及
- [12] 将所述音频信息转换为文本化字幕。
- [13] 本发明还提供一种电子设备，其包括：
- [14] 一个或一个以上处理器；
- [15] 存储器；以及
- [16] 一个或一个以上程序，存储在所述存储器中，且经配置以由所述一个或一个以上处理器执行以提供一字幕查询方法，所述一个或一个以上程序按照功能划分，包括：
- [17] 查询模块，用于获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果；
- [18] 语音文本化模块，用于根据所述查询结果，对所述待播放视频文件进行语音文本化处理；以及
- [19] 更新模块，用于根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕；
- [20] 其中所述语音文本化模块包括：
- [21] 提取单元，用于提取所述待播放视频文件的音频信息；以及
- [22] 转换单元，用于将所述音频信息转换为文本化字幕。
- [23] 本发明还提供一种储介质，其内存储有处理器可执行指令，该处理器通过执行所述指令提供一种字幕查询方法，包括：
- [24] 获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果；
- [25] 根据所述查询结果，对所述待播放视频文件进行语音文本化处理；以及
- [26] 根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕；

- [27] 其中所述语音文本化处理包括步骤：
- [28] 提取所述待播放视频文件的音频信息；以及
- [29] 将所述音频信息转换为文本化字幕。

发明的有益效果

有益效果

- [30] 相较于现有技术，本发明的字幕查询方法、电子设备及存储介质基于音频信息识别对字幕进行查询，可提高字幕匹配命中率以及降低误差率，解决了现有的视频播放器的字幕匹配命中率不高或误差率较大的技术问题。

对附图的简要说明

附图说明

- [31]
- [32] 为了更清楚地说明本发明实施例或现有技术中的技术方案，下面将对实施例或现有技术描述中所需要使用的附图作简单地介绍，显而易见地，下面描述中的附图仅仅是本发明的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动性的前提下，还可以根据这些附图获得其他的附图。
- [33] 图1为本发明的电子设备的工作环境结构示意图；
- [34] 图2为本发明的电子设备的优选实施例的结构示意图；
- [35] 图3为本发明的字幕查询方法的优选实施例的流程图。

实施该发明的最佳实施例

- [36] 请参照图式，其中相同的组件符号代表相同的组件，本发明的原理是以实施在一适当的运算环境中来举例说明。以下的说明是基于所例示的本发明具体实施例，其不应被视为限制本发明未在此详述的其它具体实施例。
- [37] 在以下的说明中，本发明的具体实施例将参考由一部或多部计算机所执行之作业的步骤及符号来说明，除非另有述明。因此，其将可了解到这些步骤及操作，其中有数次提到为由计算机执行，包括了由代表了以一结构化型式中的数据之电子信号的计算机处理单元所操纵。此操纵转换该数据或将其维持在该计算机之内存系统中的位置处，其可重新配置或另外以本领域技术人员所熟知的方

式来改变该计算机之运作。该数据所维持的数据结构为该内存之实体位置，其具有由该数据格式所定义的特定特性。但是，本发明原理以上述文字来说明，其并不代表为一种限制，本领域技术人员将可了解到以下所述的多种步骤及操作亦可实施在硬件当中。

[38] 如本申请所使用的术语“组件”、“模块”、“系统”、“接口”等等一般地旨在指计算机相关实体：硬件、硬件和软件的组合、软件或执行中的软件。例如，组件可以是但不限于是运行在处理器上的进程、处理器、对象、可执行应用、执行的线程、程序和/或计算机。通过图示，运行在控制器上的应用和该控制器二者都可以是组件。一个或多个组件可以有在于执行的进程和/或线程内，并且组件可以位于一个计算机上和/或分布在两个或更多计算机之间。

[39] 而且，要求保护的主题可以被实现为使用标准编程和/或工程技术产生软件、固件、硬件或其任意组合以控制计算机实现所公开的主题的方法、装置或制造品。本文所使用的术语“制造品”旨在包含可从任意计算机可读设备、载体或介质访问的计算机程序。当然，本领域技术人员将认识到可以对该配置进行许多修改，而不脱离要求保护的主题的范围或精神。

[40] 图1和随后的讨论提供了对实现本发明所述的电子设备的工作环境的简短、概括的描述。图1的工作环境仅仅是适当的工作环境的一个实例并且不旨在建议关于工作环境的用途或功能的范围的任何限制。实例电子设备112包括但不限于个人计算机、服务器计算机、手持式或膝上型设备、移动设备(比如移动电话、个人数字助理(PDA)、媒体播放器等等)、多处理器系统、消费型电子设备、小型计算机、大型计算机、包括上述任意系统或设备的分布式计算环境，等等。

[41] 尽管没有要求，但是在“计算机可读指令”被一个或多个电子设备执行的通用背景下描述实施例。计算机可读指令可以经由计算机可读介质来分布(下文讨论)。计算机可读指令可以实现为程序模块，比如执行特定任务或实现特定抽象数据类型的功能、对象、应用编程接口(API)、数据结构等等。典型地，该计算机可读指令的功能可以在各种环境中随意组合或分布。

[42] 图1图示了包括本发明的字幕查询方法的一个或多个实施例的电子设备112的实例。在一种配置中，电子设备112包括至少一个处理单元116和存储器118。根据

电子设备的确切配置和类型，存储器118可以是易失性的(比如RAM)、非易失性的(比如ROM、闪存等)或二者的某种组合。该配置在图1中由虚线114图示。

[43] 在其他实施例中，电子设备112可以包括附加特征和/或功能。例如，设备112还可以包括附加的存储装置(例如可移除和/或不可移除的)，其包括但不限于磁存储装置、光存储装置等等。这种附加存储装置在图1中由存储装置120图示。在一个实施例中，用于实现本文所提供的一个或多个实施例的计算机可读指令可以在存储装置120中。存储装置120还可以存储用于实现操作系统、应用程序等的其他计算机可读指令。计算机可读指令可以载入存储器118中由例如处理单元116执行。

[44] 本文所使用的术语“计算机可读介质”包括计算机存储介质。计算机存储介质包括以用于存储诸如计算机可读指令或其他数据之类的信息的任何方法或技术实现的易失性和非易失性、可移除和不可移除介质。存储器118和存储装置120是计算机存储介质的实例。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字通用盘(DVD)或其他光存储装置、盒式磁带、磁带、磁盘存储装置或其他磁存储设备、或可以用于存储期望信息并可以被电子设备112访问的任何其他介质。任意这样的计算机存储介质可以是电子设备112的一部分。

[45] 电子设备112还可以包括允许电子设备112与其他设备通信的通信连接126。通信连接126可以包括但不限于调制解调器、网络接口卡(NIC)、集成网络接口、射频发射器/接收器、红外端口、USB连接或用于将电子设备112连接到其他电子设备的其他接口。通信连接126可以包括有线连接或无线连接。通信连接126可以发射和/或接收通信媒体。

[46] 术语“计算机可读介质”可以包括通信介质。通信介质典型地包含计算机可读指令或诸如载波或其他传输机构之类的“已调制数据信号”中的其他数据，并且包括任何信息递送介质。术语“已调制数据信号”可以包括这样的信号：该信号特性中的一个或多个按照将信息编码到信号中的方式来设置或改变。

[47] 电子设备112可以包括输入设备124，比如键盘、鼠标、笔、语音输入设备、触摸输入设备、红外相机、视频输入设备和/或任何其他输入设备。设备112中也

可以包括输出设备122，比如一个或多个显示器、扬声器、打印机和/或任何其他输出设备。输入设备124和输出设备122可以经由有线连接、无线连接或其任意组合连接到电子设备112。在一个实施例中，来自另一个电子设备的输入设备或输出设备可以被用作电子设备112的输入设备124或输出设备122。

[48] 电子设备112的组件可以通过各种互连（比如总线）连接。这样的互连可以包括外围组件互连(PCI)(比如快速PCI)、通用串行总线(USB)、火线(IEEE 1394)、光学总线结构等等。在另一个实施例中，电子设备112的组件可以通过网络互连。例如，存储器118可以由位于不同物理位置中的、通过网络互连的多个物理存储器单元构成。

[49] 本领域技术人员将认识到，用于存储计算机可读指令的存储设备可以跨越网络分布。例如，可经由网络128访问的电子设备130可以存储用于实现本发明所提供的一个或多个实施例的计算机可读指令。电子设备112可以访问电子设备130并且下载计算机可读指令的一部分或所有以供执行。可替代地，电子设备112可以按需要下载多条计算机可读指令，或者一些指令可以在电子设备112处执行并且一些指令可以在电子设备130处执行。

[50] 本文提供了实施例的各种操作。在一个实施例中，所述的一个或多个操作可以构成一个或多个计算机可读介质上存储的计算机可读指令，其在被电子设备执行时将使得计算设备执行所述操作。描述一些或所有操作的顺序不应当被解释为暗示这些操作必需是顺序相关的。本领域技术人员将理解具有本说明书的益处的可替代的排序。而且，应当理解，不是所有操作必需在本文所提供的每个实施例中存在。

[51] 而且，本文所使用的词语“优选的”意指用作实例、示例或例证。奉文描述为“优选的”任意方面或设计不必被解释为比其他方面或设计更有利。相反，词语“优选的”的使用旨在以具体方式提出概念。如本申请中所使用的术语“或”旨在意指包含的“或”而非排除的“或”。即，除非另外指定或从上下文中清楚，“X使用A或B”意指自然包括排列的任意一个。即，如果X使用A；X使用B；或X使用A和B二者，则“X使用A或B”在前述任一示例中得到满足。

[52] 而且，尽管已经相对于一个或多个实现方式示出并描述了本公开，但是本领域

技术人员基于对本说明书和附图的阅读和理解将会想到等价变型和修改。本公开包括所有这样的修改和变型，并且仅由所附权利要求的范围限制。特别地关于由上述组件（例如元件、资源等）执行的各种功能，用于描述这样的组件的术语旨在对应于执行所述组件的指定功能（例如其在功能上是等价的）的任意组件（除非另外指示），即使在结构上与执行本文所示的本公开的示范性实现方式中的功能的公开结构不等同。此外，尽管本公开的特定特征已经相对于若干实现方式中的仅一个被公开，但是这种特征可以与如可以对给定或特定应用而言是期望和有利的其他实现方式的一个或多个其他特征组合。而且，就术语“包括”、“具有”、“含有”或其变形被用在具体实施方式或权利要求中而言，这样的术语旨在以与术语“包含”相似的方式包括。

[53] 本发明提供一种电子设备，如图2所示，图2为本发明的电子设备的优选实施例的结构示意图。该电子设备包括一个或一个以上处理器；存储器以及一个或一个以上程序，存储在存储器中，且经配置以由一个或一个以上处理器执行以提供一字幕查询方法，所述一个或一个以上程序按照功能划分，包括查询模块21、语音文本化模块22、以及更新模块23。

[54] 其中查询模块21用于获取待播放视频文件26的特征信息，并根据特征信息，在字幕数据库25中对待播放视频文件26进行字幕查询，以产生查询结果；语音文本化模块22用于根据查询结果，对待播放视频文件26进行语音文本化处理；更新模块23用于根据语音文本化处理生成的文本化字幕，对字幕数据库25进行更新，并将字幕数据库25中进行更新的字幕作为待播放视频文件26的字幕。

[55] 其中字幕数据库25包括标准字幕数据库以及语音本地化数据库，标准字幕数据库包括字幕以及与字幕关联的视频文件的哈希码；语音文本化数据库包括文本化字幕以及与文本化字幕关联的视频文件的哈希码。

[56] 其中查询模块21包括第一查询单元211以及第二查询单元212。第一查询单元211用于获取待播放视频文件26的特征信息，并根据特征信息，在标准字幕数据库对待播放视频文件26进行字幕查询；具体为如标准字幕数据库中存在待播放视频文件26的哈希码，则返回与哈希码关联的字幕。第二查询单元212用于如在标准字幕数据库中没有查询到待播放视频文件26的字幕，则在语音文本化数据

库中对待播放视频文件26进行字幕查询；具体为如语音文本化数据库中存在待播放视频文件26的哈希码，则返回与哈希码关联的文本化字幕。

[57] 语音文本化模块22包括提取单元221以及转换单元222。提取单元221用于提取待播放视频文件26的音频信息；转换单元222用于将音频信息转换为文本化字幕。

[58] 更新模块23包括第一更新单元231、第二更新单元232、第一字幕生成单元233以及第二字幕生成单元234。第一更新单元231用于根据语音文本化处理生成的文本化字幕，对语音文本化数据库进行更新；第二更新单元232用于将语音文本化处理生成的文本化字幕与标准字幕数据库中的字幕进行比较，如标准字幕数据库中存在字幕与语音文本化处理生成的文本化字幕一致，则对标准字幕数据库进行更新，具体为将与语音文本化处理生成的文本化字幕对应的视频文件与标准字幕数据库中的字幕进行关联；第一字幕生成单元233用于将语音文本化数据库中进行更新的字幕作为待播放视频文件26的字幕；第二字幕生成单元234用于将标准字幕数据库中进行更新的字幕作为待播放视频文件26的字幕。

[59] 本发明的电子设备使用时，会先建立标准字幕数据库以及语音文本化数据库。随后查询模块21的第一查询单元211会先查询待播放视频文件26的哈希码，然后在标准字幕数据库中查询该哈希码，如标准字幕数据库中存在该待播放视频文件26的哈希码，则将与该哈希码关联的字幕返回给用户，完成字幕查询。否则查询模块21的第二查询单元212会在语音文本化数据库中查询该待播放视频文件26的哈希码，如语音文本化数据库中存在该待播放视频文件26的哈希码，则将与该哈希码关联的文本化字幕返回给用户，完成字幕查询。如在标准字幕数据库和语音文本化数据库中均没有查询到待播放视频文件26的字幕，则语音文本化模块22对该待播放视频文件26进行语音文本化处理，并同时更新模块23的第一更新单元231将转换后的文本化字幕存入语音文本化数据库供字幕查询使用，实现了语音文本化数据库的更新。这时第一字幕生成单元233可将语音文本化数据库中进行更新的字幕作为待播放视频文件26的字幕。

[60] 优选的，更新模块23的第二更新单元232可将语音文本化处理生成的文本化字幕与标准字幕数据库中的字幕进行比较，如标准字幕数据库中存在字幕与语音

文本化处理生成的文本化字幕一致，则第二更新单元232将与语音文本化处理生成的文本化字幕对应的视频文件与字幕进行关联，即将该视频文件的哈希码与该字幕进行关联，完成标准字幕数据库的更新，优化了标准字幕数据库中的字幕的匹配命中率。这时第二字幕生成单元234可将标准字幕数据库中进行更新的字幕作为待播放视频文件26的字幕。当然这里第二更新单元232和第二字幕生成单元234可根据用户的要求进行选择设置。

[61] 本发明的电子设备的具体工作原理与下述的字幕查询方法的优选实施例中描述的相同或相似，具体请参见上述的字幕查询方法的优选实施例的详细流程。

[62] 本优选实施例的电子设备设置在视频播放器中，进行视频播放器中播放的视频文件的字幕的查询。同时本优选实施例的电子设备也可以设置在字幕查询服务器中，该字幕查询服务器可以给用户提供视频文件的字幕的查询服务。

[63] 本发明还提供一种字幕查询方法，请参照图2和图3，图3为本发明的字幕查询方法的优选实施例的流程图，该字幕查询方法包括：

[64] 步骤S301，获取待播放视频文件的特征信息，并根据特征信息，在字幕数据库中对待播放视频文件进行字幕查询，以产生查询结果；

[65] 步骤S302，根据查询结果，对待播放视频文件进行语音文本化处理；

[66] 步骤S303，根据语音文本化处理生成的文本化字幕，对字幕数据库进行更新，并将字幕数据库中进行更新的字幕作为待播放视频文件的字幕；

[67] 本优选实施例的字幕查询方法结束于步骤S303。

[68] 其中语音文本化处理包括步骤：提取待播放视频文件26的音频信息以及将音频信息转换为文本化字幕，从而实现了待播放视频文件26的文本化。

[69] 本优选实施例的字幕查询方法对于匹配不到字幕的待播放视频文件，对其进行语音文本化处理，并根据处理后的文本化字幕更新字幕数据库25，保证了字幕的匹配，同时可提高下一次字幕查询的匹配命中率。

[70] 下面详细说明本发明的字幕查询方法的优选实施例的各步骤的详细流程。

[71] 在步骤S301之前，会先建立标准字幕数据库以及语音文本化数据库以构成字幕数据库25，其中标准字幕数据库为标准字幕格式的数据库，包括字幕以及与字幕关联的视频文件的哈希码。可通过人工输入建立或通过字幕网络爬虫下载字

幕网站上的标准字幕，并可定期通过字幕网络爬虫对标准字幕数据库进行更新。语音文本化数据库是由文本化字幕建立起来的数据库，包括文本化字幕以及与文本化字幕关联的视频文件的哈希码。其中文本化字幕是对视频文件进行语音识别而形成的文本，其中包括视频文件的文字对白以及相应的时间戳。初期的语音文本化数据库可较小，但随着字幕查询的不断进行，该语音文本化数据库可被逐渐扩充（主要是最新的视频文件的文本化字幕或一些稀少的视频文件的文本化字幕）。

[72] 在步骤S301中，首先查询模块21的第一查询单元211获取待播放视频文件26的特征信息，并根据特征信息，在标准字幕数据库中对所述待播放视频文件26进行字幕查询。这里的视频文件的特征信息是指文件名、文件大小、整个或部分文件的哈希码以及上述的组合等可以标识视频文件的特征值。在本实施例中特征信息以视频文件的哈希码为例进行说明，在本步骤中，查询模块21的第一查询单元211会先查询待播放视频文件26的哈希码，然后在标准字幕数据库中查询该哈希码，如标准字幕数据库中存在该待播放视频文件26的哈希码，则将与该哈希码关联的字幕返回给用户，完成字幕查询。否则判断标准字幕数据库中没有该待播放视频文件26的字幕。

[73] 如在标准字幕数据库中没有查询到待播放视频文件26的字幕，则查询模块21的第二查询单元212在语音文本化数据库中对待播放视频文件26进行字幕查询。在本步骤中，同样第二查询单元212在语音文本化数据库中查询该待播放视频文件26的哈希码，如语音文本化数据库中存在该待播放视频文件26的哈希码，则将与该哈希码关联的文本化字幕返回给用户，完成字幕查询。否则判断语音文本化数据库中没有该待播放视频文件26的字幕。

[74] 随后转到步骤S302。

[75] 在步骤S302中，如在标准字幕数据库和语音文本化数据库中均没有查询到待播放视频文件26的字幕，则语音文本化模块22对该待播放视频文件26进行语音文本化处理。首先语音文本化模块22的提取单元221提取待播放视频文件26的音频信息，即将待播放视频文件26中的语音信号从视频中分离出来。无论待播放视频文件26中包括一个或多个语音，通过现有的语音识别技术均可输出一路主要

的语音信号（音频信息）。随后语音文本化模块22的转换单元222将音频信息转换为文本化字幕，即将音频信息转换为对白文字及相应的时间戳。该对白文字可为中文、英文或其他语言，并且该转换应具有一定的容错率。

[76] 随后转到步骤S303。

[77] 在步骤S303中，更新模块23的第一更新单元231根据语音文本化处理生成的文本化字幕，对语音文本化数据库进行更新；更新模块23的第一字幕生成单元233将语音文本化数据库中进行更新的字幕作为待播放视频文件26的字幕。这里的对语音文本化数据库进行更新主要将语音文本化处理生成的文本化字幕及与该文本化字幕关联的视频文件的哈希码存储到语音文本化数据库中。

[78] 进一步，在本步骤中，更新模块23的第二更新单元232还可将语音文本化处理生成的文本化字幕与标准字幕数据库中的字幕进行比较，如标准字幕数据库中不存在字幕与语音文本化处理生成的文本化字幕一致，则更新模块23的第二更新单元232对标准字幕数据库进行更新，并通过更新模块23的第二字幕生成单元234将标准字幕数据库中进行更新的字幕作为待播放视频文件26的字幕，否则通过更新模块23的第一字幕生成单元233将语音文本化数据库中进行更新的字幕作为待播放视频文件26的字幕。

[79] 由于标准字幕数据库为标准字幕格式的数据库，该数据库中的字幕品质较高，因此在本步骤中可尽可能将语音文本化数据库中的文本化字幕与标准字幕数据库中的字幕建立关联，以便更多的视频文件可以直接使用标准字幕数据库中的高品质字幕。

[80] 字幕和文本化字幕均包括字符串（对白文字）以及相应的时间戳。具体比较时，取播放总时间大致相同的字幕和文本化字幕进行比较，任取文本化字幕的 n （ n 大于等于2）个时间点（如5分钟处、10分钟处或20分钟处等）的字符串，和字幕的相应时间点位置的字符串进行比较（这里的时间点选取应允许一定的误差，例如和该时间点附近的10个字符串分别进行比较等），如文本化字幕的每个被比较的字符串在字幕中均有相应的字符串匹配，则判断标准字幕数据库中存在字幕与语音文本化处理生成的文本化字幕一致，否则判断标准字幕数据库中没有字幕与语音文本化处理生成的文本化字幕一致。

- [81] 如标准字幕数据库中不存在字幕与语音文本化处理生成的文本化字幕一致，第二更新单元232则将与语音文本化处理生成的文本化字幕对应的视频文件与字幕进行关联，即将该视频文件的哈希码与该字幕进行关联并将该视频文件的哈希码存储到标准字幕数据库中，完成标准字幕数据库的更新。
- [82] 这样即完成了整个待播放视频文件26的字幕查询过程以及相应的字幕数据库25的更新过程。
- [83] 这里需要指出的是，步骤S303中的语音文本化处理生成的文本化字幕与标准字幕数据库中的字幕的比较步骤以及标准字幕数据库的更新步骤是为了优化标准字幕数据库中的字幕的匹配命中率，也可根据用户的要求不进行这两步骤的设置。
- [84] 同时本优选实施例中各步骤的顺序并不限制本发明的保护范围，如对最新的视频文件可直接从步骤S302开始，以更新语音文本化数据库，而不必在字幕数据库25中进行字幕查询。因此按照其他的步骤顺序执行本发明的上述步骤也能达到优化字幕匹配命中率的效果，也属于本发明的保护范围。
- [85] 本领域普通技术人员可以理解实现上述实施例方法中的全部或部分流程，是可以通过计算机程序来指令相关的硬件来完成，所述的程序可存储于一计算机可读取存储介质中，该程序在执行时，可包括如上述各方法的实施例的流程。其中，所述的存储介质可为磁碟、光盘、只读存储记忆体（Read-Only Memory，ROM）或随机存储记忆体（Random Access Memory，RAM）等。
- [86] 本发明的字幕查询方法、电子设备及存储介质基于音频信息识别对字幕进行查询，提高了字幕匹配命中率，降低了误差率。同时根据语音文本化处理生成的文本化字幕对标准字幕数据库和语音文本化数据库进行及时更新，优化了标准字幕数据库和语音文本化数据库中的字幕的匹配命中率。
- [87] 综上所述，虽然本发明已以优选实施例揭露如上，但上述优选实施例并非用以限制本发明，本领域的普通技术人员，在不脱离本发明的精神和范围内，均可作各种更动与润饰，因此本发明的保护范围以权利要求界定的范围为准。

发明实施例

本发明的实施方式

[88]

工业实用性

[89]

序列表自由内容

[90]

权利要求书

[权利要求 1]

一种字幕查询方法，其包括步骤：

获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果；

根据所述查询结果，对所述待播放视频文件进行语音文本化处理；以及

根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕；

其中所述语音文本化处理包括步骤：

提取所述待播放视频文件的音频信息；以及

将所述音频信息转换为文本化字幕。

[权利要求 2]

根据权利要求1所述的字幕查询方法，其中所述字幕数据库包括标准字幕数据库以及语音文本化数据库；

所述获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果的步骤包括：

获取待播放视频文件的特征信息，并根据所述特征信息，在所述标准字幕数据库中对所述待播放视频文件进行字幕查询；以及

如在所述标准字幕数据库中没有查询到所述待播放视频文件的字幕，则在所述语音文本化数据库中对所述待播放视频文件进行字幕查询，以产生查询结果。

[权利要求 3]

根据权利要求2所述的字幕查询方法，其中所述根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕的步骤包括：

根据所述语音文本化处理生成的所述文本化字幕，对所述语音文

本化数据库进行更新；并将所述语音文本化数据库中进行更新的字幕作为所述待播放视频文件的字幕。

[权利要求 4] 根据权利要求3所述的字幕查询方法，其中所述根据所述查询结果，对所述待播放视频文件进行语音文本化处理的步骤包括：
如在所述字幕数据库中没有查询到所述待播放视频文件的字幕，则对所述待播放视频文件进行语音文本化处理；
所述根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕的步骤包括：
将所述语音文本化处理生成的文本化字幕与所述标准字幕数据库中的字幕进行比较，如所述标准字幕数据库中存在字幕与所述语音文本化处理生成的文本化字幕一致，则对所述标准字幕数据库进行更新，并将所述标准字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕，否则将所述语音文本化数据库中进行更新的字幕作为所述待播放视频文件的字幕。

[权利要求 5] 根据权利要求4所述的字幕查询方法，其中，
如所述标准字幕数据库中存在字幕与所述语音文本化处理生成的文本化字幕一致，则将与所述语音文本化处理生成的文本化字幕对应的视频文件与所述字幕进行关联。

[权利要求 6] 根据权利要求2所述的字幕查询方法，其中所述视频文件的特征信息为视频文件的哈希码，所述标准字幕数据库包括字幕以及与所述字幕关联的视频文件的哈希码；所述语音文本化数据库包括文本化字幕以及与所述文本化字幕关联的视频文件的哈希码；
所述获取待播放视频文件的特征信息，并根据所述特征信息，在所述标准字幕数据库中对所述待播放视频文件进行字幕查询的步骤包括：如所述标准字幕数据库中存在所述待播放视频文件的哈希码，则返回与所述哈希码关联的字幕；否则使用所述语音文本化数据库对所述待播放视频文件进行字幕查询；

所述如在所述标准字幕数据库中查询到所述待播放视频文件的字幕，则在所述语音文本化数据库中对所述待播放视频文件进行字幕查询，以产生查询结果的步骤包括：

如所述语音文本化数据库中存在所述待播放视频文件的哈希码，则返回与所述哈希码关联的文本化字幕；否则对所述待播放视频文件进行语音文本化处理。

[权利要求 7]

一种电子设备，其包括：

一个或一个以上处理器；

存储器；以及

一个或一个以上程序，存储在所述存储器中，且经配置以由所述一个或一个以上处理器执行以提供一字幕查询方法，所述一个或一个以上程序按照功能划分，包括：

查询模块，用于获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果；

语音文本化模块，用于根据所述查询结果，对所述待播放视频文件进行语音文本化处理；以及

更新模块，用于根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕；

其中所述语音文本化模块包括：

提取单元，用于提取所述待播放视频文件的音频信息；以及

转换单元，用于将所述音频信息转换为文本化字幕。

[权利要求 8]

根据权利要求7所述的电子设备，其中所述字幕数据库包括标准字幕数据库以及语音文本化数据库；

所述查询模块包括：

第一查询单元，用于获取待播放视频文件的特征信息，并根据所述特征信息，在所述标准字幕数据库中对所述待播放视频文件进

行字幕查询；以及

第二查询单元，用于如在所述标准字幕数据库中沒有查询到所述待播放视频文件的字幕，则在所述语音文本化数据库中对所述待播放视频文件进行字幕查询，以产生查询结果。

[权利要求 9]

根据权利要求8所述的电子设备，其中所述更新模块包括：

第一更新单元，用于根据所述语音文本化处理生成的所述文本化字幕，对所述语音文本化数据库进行更新；以及

第一字幕生成单元，用于将所述语音文本化数据库中进行更新的字幕作为所述待播放视频文件的字幕。

[权利要求 10]

根据权利要求9所述的电子设备，其中所述语音文本化模块用于：如在所述字幕数据库中沒有查询到所述待播放视频文件的字幕，则对所述待播放视频文件进行语音文本化处理；

所述更新模块还包括：

第二更新单元，用于将所述语音文本化处理生成的文本化字幕与所述标准字幕数据库中的字幕进行比较，如所述标准字幕数据库中不存在字幕与所述语音文本化处理生成的文本化字幕一致，则对所述标准字幕数据库进行更新；以及

第二字幕生成单元，用于将所述标准字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕。

[权利要求 11]

根据权利要求10所述的电子设备，其中所述第二更新单元用于：如所述标准字幕数据库中不存在字幕与所述语音文本化处理生成的文本化字幕一致，则将与所述语音文本化处理生成的文本化字幕对应的视频文件与所述字幕进行关联。

[权利要求 12]

根据权利要求8所述的电子设备，其中所述视频文件的特征信息为视频文件的哈希码，所述标准字幕数据库包括字幕以及与所述字幕关联的视频文件的哈希码；所述语音文本化数据库包括文本化字幕以及与所述文本化字幕关联的视频文件的哈希码；

所述第一查询单元用于：如所述标准字幕数据库中不存在所述待播

放视频文件的哈希码，则返回与所述哈希码关联的字幕；

所述第二查询单元用于：如所述语音文本化数据库中存在所述待播放视频文件的哈希码，则返回与所述哈希码关联的文本化字幕。

[权利要求 13]

一种存储介质，其内存储有处理器可执行指令，该处理器通过执行所述指令提供一种字幕查询方法，包括：

获取待播放视频文件的特征信息，并根据所述特征信息，在字幕数据库中对所述待播放视频文件进行字幕查询，以产生查询结果

；

根据所述查询结果，对所述待播放视频文件进行语音文本化处理；以及

根据所述语音文本化处理生成的所述文本化字幕，对所述字幕数据库进行更新，并将所述字幕数据库中进行更新的字幕作为所述待播放视频文件的字幕；

其中所述语音文本化处理包括步骤：

提取所述待播放视频文件的音频信息；以及

将所述音频信息转换为文本化字幕。

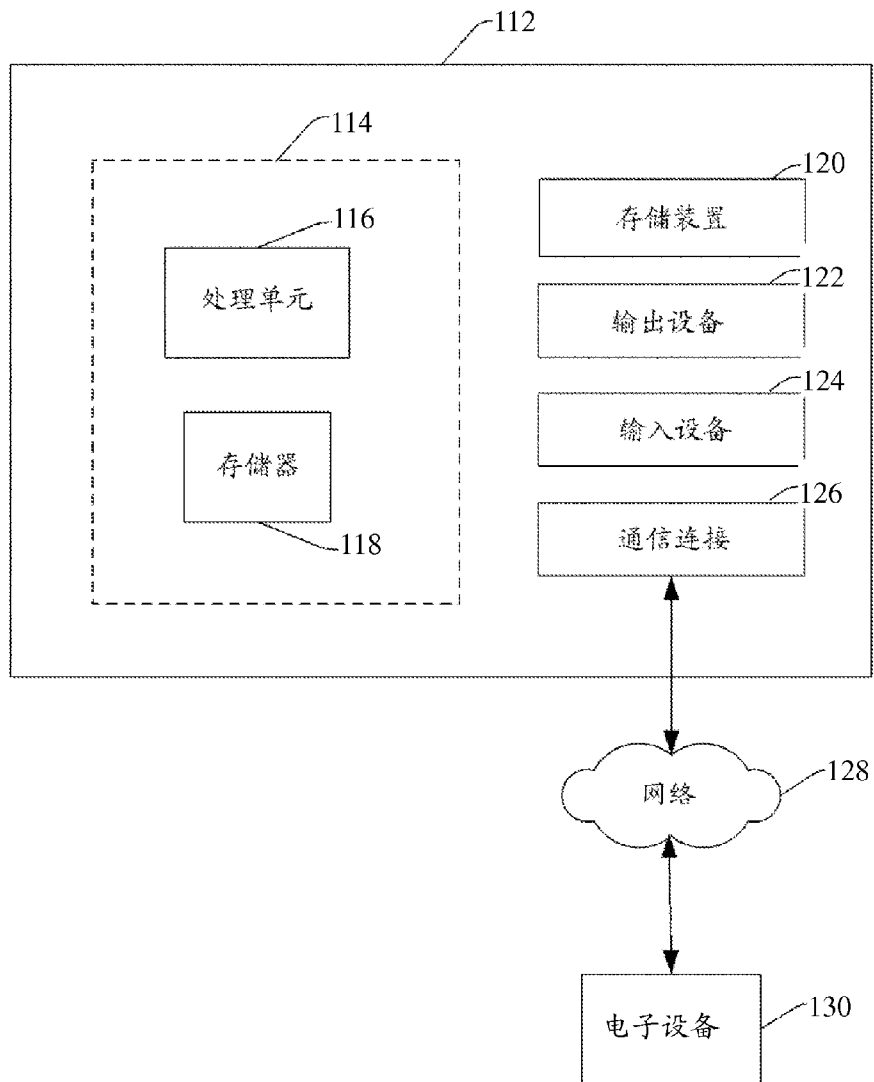


图 1

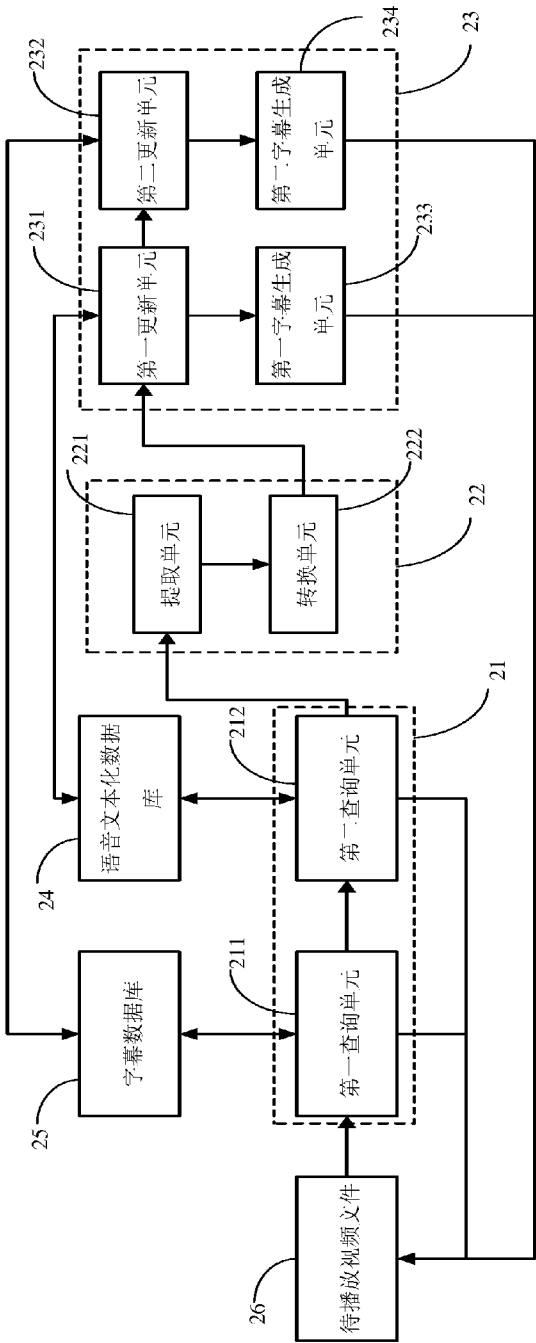


图 2

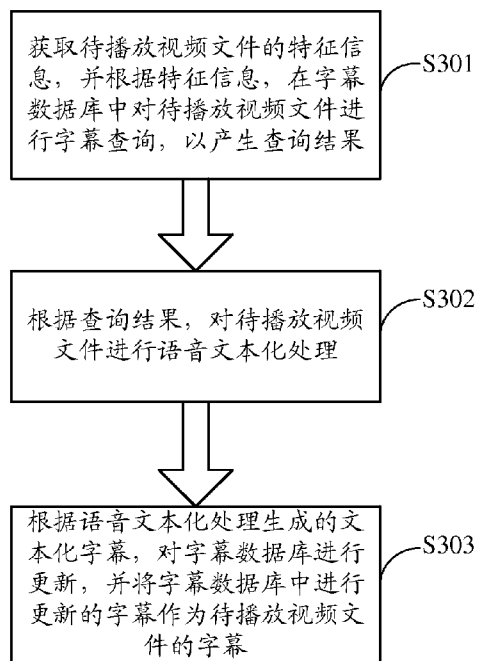


图 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/084263

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)

IPC: 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: database, video, search+, index+, recognition, speech, voice, subtitle

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 102087668 A (SKY UNION MEDIA (BEIJING) TECHNOLOGY CO., LTD.) 08 June 2011 (08.06.2011) see description, paragraphs [0008] and [0009]	1-13
Y	CN 102074235 A (SHANGHAI HUAQIN COMMUNICATION TECHNOLOGY CO., LTD.) 25 May 2011 (25.05.2011) see description, paragraph [0024]	1-13
A	US 5703655 A (U S WEST TECHNOLOGIES, INC.) 30 December 1997 (30.12.1997) the whole document	1-13

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:	“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
“A” document defining the general state of the art which is not considered to be of particular relevance	“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
“E” earlier application or patent but published on or after the international filing date	“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
“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)	“&” document member of the same patent family
“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	

Date of the actual completion of the international search 09 December 2013 (09.12.2013)	Date of mailing of the international search report 02 January 2014 (02.01.2014)
Name and mailing address of the ISA State Intellectual Property Office of the P. R. China No. 6, Xitucheng Road, Jimenqiao Haidian District, Beijing 100088, China Facsimile No. (86-10) 62019451	Authorized officer LIU, Man Telephone No. (86-10) 62414025

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/084263

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 102087668 A	08.06.2011	None	
CN 102074235 A	25.05.2011	None	
US 5703655 A	30.12.1997	None	

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)

IPC: 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: 数据库,语音,识别,字幕,歌词,视频,检索,查找,查询,搜索,video, search+, index+, database, recognition, speech, voice, subtitle

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
Y	CN 102087668 A (天擎华媒(北京)科技有限公司) 08.6 月 2011 (08.06.2011) 参见说明书第[0008],[0009]段	1-13
Y	CN 102074235 A (上海华勤通讯技术有限公司) 25.5 月 2011 (25.05.2011) 参见说明书第[0024]段	1-13
A	US 5703655 A (U S WEST TECHNOLOGIES, INC.) 30.12 月 1997 (30.12.1997) 全文	1-13

☐ 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

09.12 月 2013 (09.12.2013)

国际检索报告邮寄日期

02.1 月 2014 (02.01.2014)

ISA/CN 的名称和邮寄地址:

中华人民共和国国家知识产权局

中国北京市海淀区蓟门桥西土城路 6 号 100088

传真号: (86-10)62019451

受权官员

刘曼

电话号码: (86-10) 62414025

关于同族专利的信息

PCT/CN2013/084263

PCT/ISA/210 表(同族专利附件) (2009 年 7 月)