

(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
2017 年 10 月 5 日 (05.10.2017)



(10) 国际公布号
WO 2017/166430 A1

- (51) 国际专利分类号:
H04N 21/431 (2011.01) H04N 21/488 (2011.01)
H04N 21/435 (2011.01)
- (21) 国际申请号: PCT/CN2016/084872
- (22) 国际申请日: 2016 年 6 月 4 日 (04.06.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201610184443.X 2016 年 3 月 28 日 (28.03.2016) CN
- (71) 申请人: 深圳 TCL 新技术有限公司 (SHENZHEN TCL NEW TECHNOLOGY CO., LTD) [CN/CN]; 中国广东省深圳市南山区中山园路 1001 号 TCL 国际 E 城科技大厦 D4 栋 7 楼, Guangdong 518052 (CN)。
- (72) 发明人: 柯杰燕 (KE, Jieyan); 中国广东省深圳市南山区中山园路 1001 号 TCL 国际 E 城科技大厦 D4 栋 7 楼, Guangdong 518052 (CN)。
- (74) 代理人: 深圳市世纪恒程知识产权代理事务所 (CENFO INTELLECTUAL PROPERTY AGENCY); 中国广东省深圳市南山区南山大道 3838 号设计产业园金栋二层 210-212 (原南头城工业村 11 栋), Guangdong 518052 (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, IR, IS, JP, KE, KG, KN, KP, KR, 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,

[见续页]

(54) Title: CAPTION DISPLAY METHOD AND DEVICE

(54) 发明名称: 字幕显示方法及装置

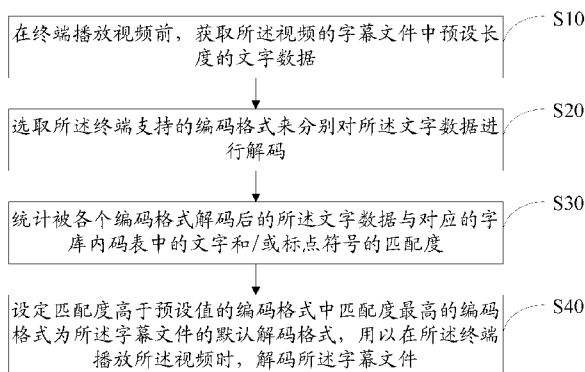


图 1

- S10 Acquire text data having a preset length in a caption file of a video before the video is played by a terminal
- S20 Select encoding formats supported by the terminal for decoding the text data respectively
- S30 Calculate matching degrees between the text data after being decoded in the encoding formats and the texts and/or the punctuation marks in a corresponding internal code table in a word stock
- S40 Set the encoding format having a highest matching degree among the encoding formats having the matching degrees being higher than a preset value as a default decoding format of the caption file to decode the caption file when the video is played by the terminal

(57) Abstract: Disclosed is a caption display method, comprising: acquiring text data having a preset length in a caption file of a video before the video is played by a terminal; selecting encoding formats supported by the terminal for decoding the text data respectively; calculating matching degrees between the text data after being decoded in the encoding formats and the texts and/or the punctuation marks in a corresponding internal code table in a word stock; and setting the encoding format having a highest matching degree among the encoding formats having the matching degrees being higher than a preset value as a default decoding format of the caption file to decode the caption file when the video is played by the terminal. Also disclosed is a caption display device. The present invention improves the accuracy and the convenience of caption display.

(57) 摘要: 本发明公开了一种字幕显示方法, 包括: 在终端播放视频前, 获取所述视频的 字幕文件中预设长度的文字数据; 选取所述终端支持的编码格式来分别对所述文字数据进行解码; 统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度; 及设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式, 用以在所述终端播放所述视频时, 解码所述字幕文件。本发明还公开了一种字幕显示装置。本发明提高了字幕显示的准确率及便捷性。



WO 2017/166430 A1



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,
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))。

字幕显示方法及装置

[1] 技术领域

[2] 本发明涉及终端技术领域，尤其涉及一种字幕显示方法及装置。

[3] 背景技术

[4] 目前，在终端的多媒体上播放视频时，用户一般打开字幕观看，特别是与母语不一样的视频。该终端以电视为例，由于字幕是独立于视频另外制作的，因此不同的制作方式将会造成文字的编码格式有所不同。此外，同一编码格式所对应的字库的内码表（codepage）可能不一样，内码表的设置或选择错误，导致系统无法解码出正常的文字。如此一来，用户在观看视频的过程中，就存在字幕显示出来是一些乱码或存在缺失，字幕显示的准确率大大降低。而对于没有手动切换字幕语言或编码格式功能的电视，导致用户无法观看字幕，非常不便捷，给用户带来不好的体验。

[5] 发明内容

[6] 本发明的主要目的在于提供一种字幕显示方法及装置，旨在提高字幕显示的准确率及便捷性。

[7] 为实现上述目的，本发明提供了一种字幕显示方法，包括：

[8] 在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；

[9] 选取所述终端支持的编码格式来分别对所述文字数据进行解码；

[10] 统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；及

[11] 设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件；

[12] 其中，若统计的所述匹配度均不高于所述预设值，则提示用户所述终端解码所述字幕文件的失败；

[13] 若匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种，则从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式

- 。
- [14] 可选地，该方法还包括：
- [15] 若统计的所述匹配度均为零，则重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。
- [16] 可选地，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：
- [17] 对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。
- [18] 本发明还提供了一种字幕显示方法，包括：
- [19] 在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；
- [20] 选取所述终端支持的编码格式来分别对所述文字数据进行解码；
- [21] 统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；及
- [22] 设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件。
- [23] 可选地，该方法还包括：
- [24] 若统计的所述匹配度均不高于所述预设值，则提示用户所述终端解码所述字幕文件的失败。
- [25] 可选地，该方法还包括：
- [26] 若匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种，则从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。
- 。
- [27] 可选地，该方法还包括：
- [28] 若统计的所述匹配度均为零，则重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。
- [29] 可选地，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：
- [30] 对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标

识。

- [31] 此外，为实现上述目的，本发明还提供了一种字幕显示装置，包括：
- [32] 获取模块，用于在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；
- [33] 解密模块，用于选取所述终端支持的编码格式来分别对所述文字数据进行解码；
- [34] 统计模块，用于统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；
- [35] 设定模块，用于设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件。
- [36] 可选地，所述字幕显示装置还包括：
- [37] 提示模块，用于在统计的所述匹配度均不高于所述预设值时，提示用户所述终端解码所述字幕文件的失败。
- [38] 可选地，所述字幕显示装置还包括：
- [39] 选取模块，用于在匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种时，从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。
- [40] 可选地，所述字幕显示装置还包括：
- [41] 处理模块，用于在统计的所述匹配度均为零时，重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。
- [42] 可选地，所述字幕显示装置还包括：
- [43] 标识模块，用于对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。
- [44] 本发明实施例终端在进行字幕显示时，首先通过终端支持的编码格式分别对预设长度的文字数据进行解码，设定各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，以在终端播放该视频时，解码字幕

文件，对字幕进行显示。使得终端可以自动地切换编码格式来寻找正确的编码格式，将字幕进行正常显示，避免字幕乱码或丢失，提高了字幕显示的准确率及便捷性。

[45] 附图说明

[46] 图1为本发明字幕显示方法一实施例的流程示意图；

[47] 图2为本发明字幕显示装置一实施例的功能模块示意图。

[48] 本发明目的的实现、功能特点及优点将结合实施例，参照附图做进一步说明。

[49] 具体实施方式

[50] 应当理解，此处所描述的具体实施例仅仅用以解释本发明，并不用于限定本发明。

[51] 如图1所示，示出了本发明一种字幕显示方法第一实施例。该实施例的字幕显示方法包括：

[52] 步骤S10、在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；

[53] 本实施例中，终端的类型可根据实际需要进行设置，例如，该终端包括电视、电脑等可进行视频播放的终端，以下将以该终端为电视进行详细说明。

[54] 当用户选择打开一个视频进行播放时，电视播放该视频之前，将会获取视频的相关信息，视频的相关信息包括视频文件、视频的字幕文件等。从视频的字幕文件中获取预设长度的文字数据，预设长度的文字数据看是一段连续的文字数据，也可以是不连续的多段文字数据，该预设长度可根据具体情况而灵活设置，例如，从字幕文件的中间选取10个字节长度的文字数据，或者从字幕文件的起始位置获取20个字节长度的文字数据。

[55] 步骤S20、选取所述终端支持的编码格式来分别对所述文字数据进行解码；

[56] 步骤S30、统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；

[57] 由于字幕一般只是显示文字和标点符号，因此，只需要将解码得到的文字和标点符号与对应的字库内码表中的文字和标点符号进行比较。电视在播放字幕前，先选取电视支持的一种编码格式进行解码，对获取得到的视频的字幕文件中

预设长度的文字数据进行解码。解码完成后，对该编码格式所解码出来的文字和/或标点符号的总数进行统计，以及对解码出来的文字和/或标点符号落在与该编码格式对应的字库内码表中的文字和/或标点符号相同的个数进行统计。将个数与总数之间比值作为匹配度，即由个数/总数，可得到该编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度。例如，当该编码格式解码出来的文字和/或标点符号的总数为10，解码出来的文字和/或标点符号落在与该编码格式对应的字库内码表中的文字和/或标点符号相同的个数为9，则该编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为90%。以此类推，继续用电视支持的下一编码格式进行解码判断，直到尝试完最后一种编码格式，最后得到各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度。

- [58] 步骤S40、设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件。
- [59] 由于错误编码格式导致字幕无法显示的概率要高于正确编码格式正常解码的概率，因此，对上述得到各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度后，电视将各个编码格式得到的匹配度高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，用以在终端播放所述视频时，解码字幕文件，对字幕进行显示。即电视进行视频播放的过程中，通过该默认解码格式解码字幕文件，显示解码得到的字幕。该预设值可根据具体情况而灵活设置，例如，预设值可设置为95%。
- [60] 以下进行举例说明，假设编码格式A解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为95%，编码格式B解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为100%，编码格式C解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为98%，则在视频播放的过程中，以编码格式B作为视频播放的默认编码格式对字幕文件中的文字数据进行解码，显示解码得到的字幕。

- [61] 本发明实施例电视在进行字幕显示时，首先通过电视支持的编码格式分别对预设长度的文字数据进行解码，设定各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，以在电视播放该视频时，解码字幕文件，对字幕进行显示。使得电视可以自动地切换编码格式来寻找正确的编码格式，将字幕进行正常显示，避免字幕乱码或丢失，提高了字幕显示的准确率及便捷性。
- [62] 进一步地，提出了本发明字幕显示方法第二实施例，该实施例中上述步骤S30之后包括：若统计的所述匹配度均不高于所述预设值，则提示用户所述终端解码所述字幕文件的失败。
- [63] 本实施例中，在将各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号进行比较的过程中，若各个编码格式对应的匹配度均小于或等于预设值，说明所有的编码格式对应的匹配度都比较低。这里的预设值与上述提到的预设值一致，可根据具体情况而灵活设置，例如，预设值可设置为95%。此时，电视判定没有找到合适的编码格式，输出解码结果相关的提示信息，例如，输出电视解码字幕文件的失败信息，以提示用户进行相应的处理。该失败信息可以是在电视的显示界面上显示各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度，也可以是通过语音播报的形式对各个编码格式解码结果的播报。用户通过提示信息了解各个编码格式的解码情况后，可以通过手动方式选择其中的一种的编码格式，电视根据接收到的设置指令将该编码格式作为视频播放的默认编码格式对字幕文件中的文字数据进行解码，并显示解码得到的字幕。
- [64] 本实施例当各个编码格式解码得到的匹配度均小于或等于预设值时，输出电视解码字幕文件的失败信息，提示用户进行相应处理。使得电视可以根据用户选取的编码格式将字幕进行显示，减少了字幕乱码现象，提高了字幕显示的准确率。
- [65] 进一步地，提出了本发明字幕显示方法第三实施例，该实施例中上述步骤S30之后包括：若匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种

，则从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。

[66] 本实施例中，在上述将各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号进行比较的过程中，当存在多个编码格式解码得到的文字/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度高于预设值的编码格式中匹配度最高的编码格式，例如，多于一种编码格式的匹配度是100%，这种情况则以电视的系统语言作为参考，此时，电视可根据系统语言选取对应的一种编码格式，作为视频播放的默认编码格式对字幕文件中的文字数据进行解码，并显示解码得到的字幕。这里的预设值与上述提到的预设值一致，可根据具体情况而灵活设置，例如，预设值可设置为95%。以下进行举例说明，假设电视的系统语言是希腊语，而CP1253编码格式解码得到的文字和/或标点符号落在希腊语的字库内码表CP1253的匹配度为100%，则选择该编码格式作为视频播放的默认编码格式对字幕文件中的文字数据进行解码。

[67] 本实施例中，当匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种时，电视根据系统语言选取对应的一种编码格式设定为字幕文件的默认解码格式，对字幕文件中的文字数据进行解码并显示字幕。使得电视可以自动地切换编码格式来寻找正确的默认编码格式，将字幕进行正常显示，避免字幕乱码或丢失，提高了字幕显示的准确率及便捷性。

[68] 进一步地，提出了本发明字幕显示方法第四实施例，该实施例中上述步骤S30之后包括：若统计的所述匹配度均为零，则重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。

[69] 本实施例中，当电视统计支持的所有编码格式统计解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度后，若各个编码格式对应的匹配度均为零，说明可能是由于电视获取得到的字幕文件中预设长度的文字数据恰好为非文字或标点符号，为了能够获取正确的默认编码格式进行解码，则电视重新获取视频的字幕文件中预设长度的新文字数据，按照上述方法根据电视支持的所有编码格式分别对新文字数据进行解码，统计被各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度。将

重新解码得到的匹配度中，设定匹配度高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，用以在电视播放该视频时，解码字幕文件。可以理解的是，当各个编码格式对应的匹配度均不高于预设值时，提示用户终端解码字幕文件的失败信息，并接收设置指令选取其中的一种的编码格式进行解码。或者，当匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种，则从中选取与电视的系统语言对应的编码格式设定为字幕文件的默认解码格式，即根据电视的系统语言选取对应的一种编码格式进行解码。

[70] 本实施例当各个编码格式对应的匹配度均为零时，电视重新获取该视频的字幕文件中预设长度的新文字数据进行解码比较，寻找正确的默认编码格式，将字幕进行正常显示，避免字幕乱码或丢失，不仅提升了用户体验，而且提高了字幕显示的准确率及便捷性。

[71] 进一步地，提出了本发明字幕显示方法第五实施例，该实施例中上述步骤S30之前包括：对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[72] 本实施例中，由于字幕一般只是显示文字与标点符号，为了使电视可快速统计被各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，因此，可预先将电视所支持的所有编码格式对应的字库内码表中的文字和标点符号标识出来。在电视播放视频进行字幕显示前，先获取视频的字幕文件中预设长度的文字数据，根据电视支持的所有编码格式分别对该文字数据进行解码，并根据预先标识各个编码格式对应的字库内码表中的文字和标点符号，统计各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中已标识的文字和/或标点符号的匹配度，电视将各个编码格式对应的匹配度进行比较后，选择一种编码格式作为默认编码格式对视频播放的字幕文件中的文字数据进行解码并显示字幕。

[73] 本实施例预先对电视支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识，使得电视可快速统计被各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，提高了电视进行字幕显示的快捷性。

[74] 对应地，如图2所示，提出本发明一种字幕显示装置第一实施例。该实施例的字幕显示装置包括：

[75] 获取模块100，用于在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；

[76] 本实施例中，终端的类型可根据实际需要进行设置，例如，该终端包括电视、电脑等可进行视频播放的终端，以下将以该终端为电视进行详细说明。

[77] 当用户选择打开一个视频进行播放时，电视播放该视频之前，获取模块100将会获取视频的相关信息，视频的相关信息包括视频文件、视频的字幕文件等。从视频的字幕文件中获取预设长度的文字数据，预设长度的文字数据看是一段连续的文字数据，也可以是不连续的多段文字数据，该预设长度可根据具体情况而灵活设置，例如，从字幕文件的中间选取10个字节长度的文字数据，或者从字幕文件的起始位置获取20个字节长度的文字数据。

[78] 解密模块200，用于选取所述终端支持的编码格式来分别对所述文字数据进行解码；

[79] 统计模块300，用于统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；

[80] 由于字幕一般只是显示文字和标点符号，因此，只需要将解码得到的文字和标点符号与对应的字库内码表中的文字和标点符号进行比较。电视在播放字幕前，解密模块200先选取电视支持的一种编码格式进行解码，对获取得到的视频的字幕文件中预设长度的文字数据进行解码。解码完成后，统计模块300对该编码格式所解码出来的文字和/或标点符号的总数进行统计，以及对解码出来的文字和/或标点符号落在与该编码格式对应的字库内码表中的文字和/或标点符号相同的个数进行统计。将个数与总数之间比值作为匹配度，即由个数/总数，可得到该编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度。例如，当该编码格式解码出来的文字和/或标点符号的总数为10，解码出来的文字和/或标点符号落在与该编码格式对应的字库内码表中的文字和/或标点符号相同的个数为9，则该编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为90%。依此类推，继续用

电视支持的下一一种编码格式进行解码判断，直到尝试完最后一种编码格式，最后得到各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度。

[81] 设定模块400，用于设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件。

[82] 由于错误编码格式导致字幕无法显示的概率要高于正确编码格式正常解码的概率，因此，对上述得到各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度后，设定模块400将各个编码格式得到的匹配度高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，用以在终端播放所述视频时，解码字幕文件，对字幕进行显示。即电视进行视频播放的过程中，通过该默认解码格式解码字幕文件，显示解码得到的字幕。该预设值可根据具体情况而灵活设置，例如，预设值可设置为95%。

[83] 以下进行举例说明，假设编码格式A解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为95%，编码格式B解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为100%，编码格式C解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度为98%，则在视频播放的过程中，以编码格式B作为视频播放的默认编码格式对字幕文件中的文字数据进行解码，显示解码得到的字幕。

[84] 本发明实施例电视在进行字幕显示时，首先通过电视支持的编码格式分别对预设长度的文字数据进行解码，设定各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，以在电视播放该视频时，解码字幕文件，对字幕进行显示。使得电视可以自动地切换编码格式来寻找正确的编码格式，将字幕进行正常显示，避免字幕乱码或丢失，提高了字幕显示的准确率及便捷性。

[85] 进一步地，提出了本发明字幕显示装置第二实施例，该实施例中上述字幕显示装置还包括：提示模块，用于在统计的所述匹配度均不高于所述预设值时，提

示用户所述终端解码所述字幕文件的失败。

- [86] 本实施例中，在将各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号进行比较的过程中，若各个编码格式对应的匹配度均小于或等于预设值，说明所有的编码格式对应的匹配度都比较低。这里的预设值与上述提到的预设值一致，可根据具体情况而灵活设置，例如，预设值可设置为95%。此时，电视判定没有找到合适的编码格式，由提示模块输出解码结果相关的提示信息，例如，提示模块输出电视解码字幕文件的失败信息，以提示用户进行相应的处理。该失败信息可以是在电视的显示界面上显示各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度，也可以是通过语音播报的形式对各个编码格式解码结果的播报。用户通过提示信息了解各个编码格式的解码情况后，可以通过手动方式选择其中的一种的编码格式，电视根据接收到的设置指令将该编码格式作为视频播放的默认编码格式对字幕文件中的文字数据进行解码，并显示解码得到的字幕。
- [87] 本实施例当各个编码格式解码得到的匹配度均小于或等于预设值时，输出电视解码字幕文件的失败信息，提示用户进行相应处理。使得电视可以根据用户选取的编码格式将字幕进行显示，减少了字幕乱码现象，提高了字幕显示的准确率。
- [88] 进一步地，提出了本发明字幕显示装置第三实施例，该实施例中上述字幕显示装置还包括：选取模块，用于在匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种时，从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。
- [89] 本实施例中，在上述将各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号进行比较的过程中，当存在多个编码格式解码得到的文字/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度高于预设值的编码格式中匹配度最高的编码格式，例如，多于一种编码格式的匹配度是100%，这种情况则以电视的系统语言作为参考，此时，选取模块可根据系统语言选取对应的一种编码格式，作为视频播放的默认编码格式对字幕文件中的文字数据进行解码，并显示解码得到的字幕。这里的预设值与上述提到

的预设值一致，可根据具体情况而灵活设置，例如，预设值可设置为95%。以下进行举例说明，假设电视的系统语言是希腊语，而CP1253编码格式解码得到的文字和/或标点符号落在希腊语的字库内码表CP1253的匹配度为100%，则选择该编码格式作为视频播放的默认编码格式对字幕文件中的文字数据进行解码。

[90] 本实施例中，当匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种时，电视根据系统语言选取对应的一种编码格式设定为字幕文件的默认解码格式，对字幕文件中的文字数据进行解码并显示字幕。使得电视可以自动地切换编码格式来寻找正确的默认编码格式，将字幕进行正常显示，避免字幕乱码或丢失，提高了字幕显示的准确率及便捷性。

[91] 进一步地，提出了本发明字幕显示装置第四实施例，该实施例中上述字幕显示装置还包括：

[92] 处理模块，用于在统计的所述匹配度均为零时，重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。

[93] 本实施例中，当电视统计支持的所有编码格式统计解码得到的文字和/或标点符号与对应的字库内码表中的文字和/或标点符号的匹配度后，若各个编码格式对应的匹配度均为零，说明可能是由于电视获取得到的字幕文件中预设长度的文字数据恰好为非文字或标点符号，为了能够获取正确的默认编码格式进行解码，则处理模块重新获取视频的字幕文件中预设长度的新文字数据，按照上述方法根据电视支持的所有编码格式分别对新文字数据进行解码，统计被各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度。将重新解码得到的匹配度中，设定匹配度高于预设值的编码格式中匹配度最高的编码格式为字幕文件的默认解码格式，用以在电视播放该视频时，解码字幕文件。可以理解的是，当各个编码格式对应的匹配度均不高于预设值时，提示用户终端解码字幕文件的失败信息，并接收设置指令选取其中的一种的编码格式进行解码。或者，当匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种，则从中选取与电视的系统语言对应的编码格式设定为字幕文件的默认解码格式，即根据电视的系统语言选取对应的一种编码格式进行解码。

[94] 本实施例当各个编码格式对应的匹配度均为零时，电视重新获取该视频的字幕

文件中预设长度的新文字数据进行解码比较，寻找正确的默认编码格式，将字幕进行正常显示，避免字幕乱码或丢失，不仅提升了用户体验，而且提高了字幕显示的准确率及便捷性。

[95] 进一步地，提出了本发明字幕显示装置第五实施例，该实施例中上述字幕显示装置还包括：

[96] 标识模块，用于对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[97] 本实施例中，由于字幕一般只是显示文字与标点符号，为了使电视可快速统计被各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，因此，标识模块可预先将电视所支持的所有编码格式对应的字库内码表中的文字和标点符号标识出来。在电视播放视频进行字幕显示前，先获取视频的字幕文件中预设长度的文字数据，根据电视支持的所有编码格式分别对该文字数据进行解码，并根据预先标识各个编码格式对应的字库内码表中的文字和标点符号，统计各个编码格式解码得到的文字和/或标点符号与对应的字库内码表中已标识的文字和/或标点符号的匹配度，电视将各个编码格式对应的匹配度进行比较后，选择一种编码格式作为默认编码格式对视频播放的字幕文件中的文字数据进行解码并显示字幕。

[98] 本实施例预先对电视支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识，使得电视可快速统计被各个编码格式解码后的文字数据与对应的字库内码表中的文字和/或标点符号的匹配度，提高了电视进行字幕显示的快捷性。

[99] 以上仅为本发明的优选实施例，并非因此限制本发明的专利范围，凡是利用本发明说明书及附图内容所作的等效结构或等效流程变换，或直接或间接运用在其他相关的技术领域，均同理包括在本发明的专利保护范围内。

权利要求书

- [权利要求 1] 一种字幕显示方法，其特征在于，所述字幕显示方法包括以下步骤：
- 在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；
- 选取所述终端支持的编码格式来分别对所述文字数据进行解码；
- 统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；及
- 设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件；
- 其中，若统计的所述匹配度均不高于所述预设值，则提示用户所述终端解码所述字幕文件的失败；
- 若匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种，则从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。
- [权利要求 2] 权利要求1所述的字幕显示方法，其特征在于，该方法还包括：
- 若统计的所述匹配度均为零，则重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。
- [权利要求 3] 权利要求1所述的字幕显示方法，其特征在于，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：
- 对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。
- [权利要求 4] 权利要求2所述的字幕显示方法，其特征在于，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：
- 对所述终端支持的所有编码格式对应的字库内码表中的文字和标

点符号进行标识。

- [权利要求 5] 一种字幕显示方法，其特征在于，所述字幕显示方法包括以下步骤：
在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；
选取所述终端支持的编码格式来分别对所述文字数据进行解码；
统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；及
设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件。
- [权利要求 6] 权利要求5所述的字幕显示方法，其特征在于，该方法还包括：
若统计的所述匹配度均不高于所述预设值，则提示用户所述终端解码所述字幕文件的失败。
- [权利要求 7] 权利要求5所述的字幕显示方法，其特征在于，该方法还包括：
若匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种，则从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。
- [权利要求 8] 权利要求5所述的字幕显示方法，其特征在于，该方法还包括：
若统计的所述匹配度均为零，则重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。
- [权利要求 9] 权利要求5所述的字幕显示方法，其特征在于，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：
对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。
- [权利要求 10] 权利要求6所述的字幕显示方法，其特征在于，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或

标点符号的匹配度之前还包括：

对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[权利要求 11]

权利要求7所述的字幕显示方法，其特征在于，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：

对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[权利要求 12]

权利要求8所述的字幕显示方法，其特征在于，所述统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度之前还包括：

对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[权利要求 13]

一种字幕显示装置，其特征在于，所述字幕显示装置包括：

获取模块，用于在终端播放视频前，获取所述视频的字幕文件中预设长度的文字数据；

解密模块，用于选取所述终端支持的编码格式来分别对所述文字数据进行解码；

统计模块，用于统计被各个编码格式解码后的所述文字数据与对应的字库内码表中的文字和/或标点符号的匹配度；

设定模块，用于设定匹配度高于预设值的编码格式中匹配度最高的编码格式为所述字幕文件的默认解码格式，用以在所述终端播放所述视频时，解码所述字幕文件。

[权利要求 14]

权利要求13所述的字幕显示装置，其特征在于，所述字幕显示装置还包括：

提示模块，用于在统计的所述匹配度均不高于所述预设值时，提示用户所述终端解码所述字幕文件的失败。

[权利要求 15]

权利要求13所述的字幕显示装置，其特征在于，所述字幕显示装

置还包括：

选取模块，用于在匹配度高于预设值的编码格式中匹配度最高的编码格式不止一种时，从中选取与所述终端的系统语言对应的编码格式设定为所述字幕文件的默认解码格式。

[权利要求 16]

权利要求13所述的字幕显示装置，其特征在于，所述字幕显示装置还包括：

处理模块，用于在统计的所述匹配度均为零时，重新从所述视频的字幕文件中获取所述预设长度的文字数据来进行解码及后续相关步骤。

[权利要求 17]

权利要求13所述的字幕显示装置，其特征在于，所述字幕显示装置还包括：

标识模块，用于对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[权利要求 18]

权利要求14所述的字幕显示装置，其特征在于，所述字幕显示装置还包括：

标识模块，用于对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[权利要求 19]

权利要求15所述的字幕显示装置，其特征在于，所述字幕显示装置还包括：

标识模块，用于对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

[权利要求 20]

权利要求16所述的字幕显示装置，其特征在于，所述字幕显示装置还包括：

标识模块，用于对所述终端支持的所有编码格式对应的字库内码表中的文字和标点符号进行标识。

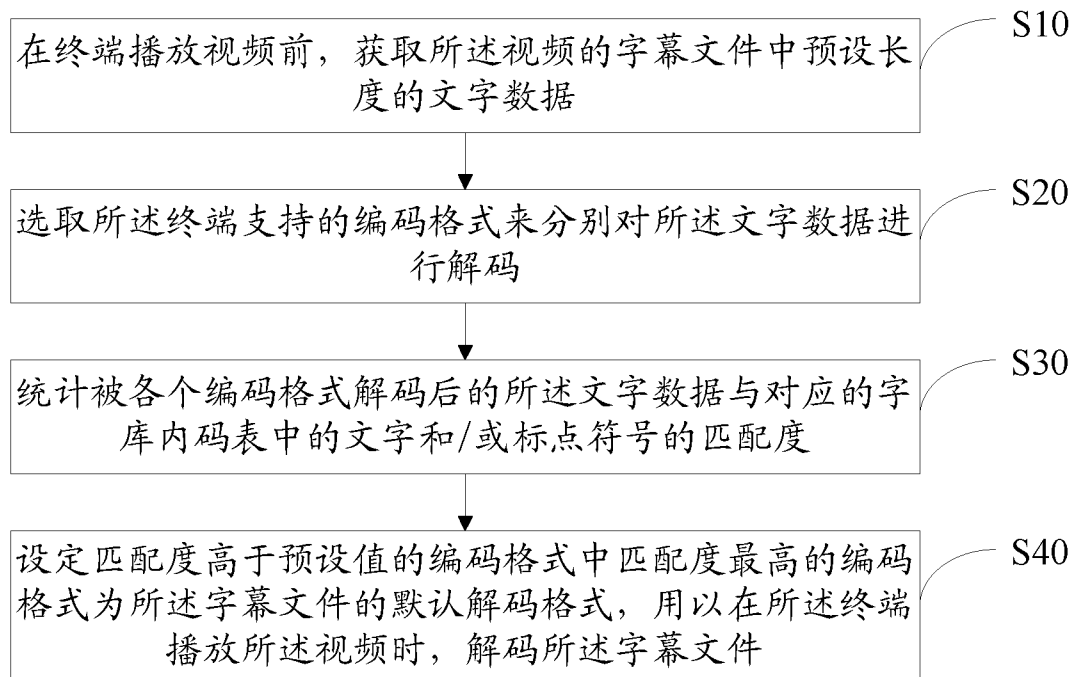


图 1

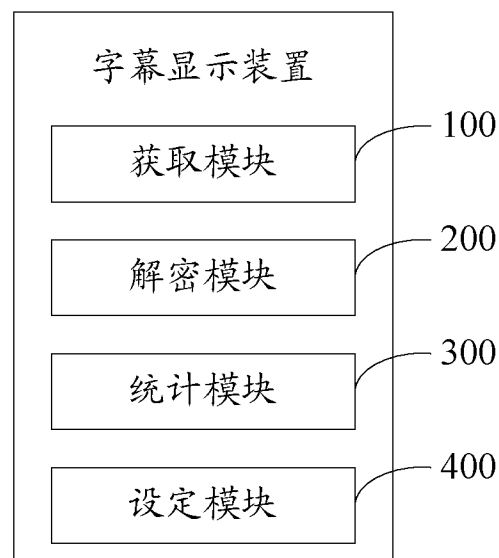


图 2

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/084872

A. CLASSIFICATION OF SUBJECT MATTER

H04N 21/431 (2011.01) i; H04N 21/435 (2011.01) i; H04N 21/488 (2011.01) 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 G11B G06K H04N

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, IEEE: character, file, document, word stock, messy code, coincide, comparison, Internal code, table, fail, decod+, encod+, text, string+, title, subtitle, caption, determing, recogniz+, detect+, code, format, type, match

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CN 104516862 A (FOUNDER GROUP et al.) 15 April 2015 (15.04.2015) claims 1, 5, 9 and 12, and description, paragraphs [0065]-[0080], and figure 1	1-20
Y	CN 102194503 A (TENCENT TECHNOLOGY (SHENZHEN) CO., LTD.) 21 September 2011 (21.09.2011) claims 1 and 4, and description, paragraphs [0036]-[0039], [0072] and [0074], and figures 1-3	1-20
A	CN 102595082 A (SHENZHEN SKYWORTH-RGB ELECTRONICS CO., LTD.) 18 July 2012 (18.07.2012) the whole document	1-20
A	CN 104750666 A (MAINBO EDUCATION TECHNOLOGY CO., LTD.) 01 July 2015 (01.07.2015) the whole document	1-20
A	CN 104156373 A (ACER INCORPORATED) 19 November 2014 (19.11.2014) the whole document	1-20

☒ 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 29 November 2016	Date of mailing of the international search report 21 December 2016
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 HU, Yaqin Telephone No. (86-10) 61648233

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2016/084872

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2013077855 A1 (ISYS SEARCH SOFTWARE PTY LTD.) 28 March 2013 (28.03.2013) the whole document	1-20
A	US 2006025988 A1 (XU, MING et al.) 02 February 2006 (02.02.2006) the whole document	1-20

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2016/084872

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
CN 104516862 A	15 April 2015	EP 3051428 A1	03 August 2016
		WO 2015043072 A1	02 April 2015
		US 2016239467 A1	18 August 2016
CN 102194503 A	21 September 2011	CN 102194503 B	27 August 2014
CN 102595082 A	18 July 2012	None	
CN 104750666 A	01 July 2015	None	
CN 104156373 A	19 November 2014	None	
US 2013077855 A1	28 March 2013	AU 2012201539 B2	16 June 2016
		US 9122898 B2	01 September 2015
		AU 2012201539 A1	06 December 2012
US 2006025988 A1	02 February 2006	US 7865355 B2	04 January 2011

国际检索报告

国际申请号

PCT/CN2016/084872

A. 主题的分类

H04N 21/431(2011.01)i; H04N 21/435(2011.01)i; H04N 21/488(2011.01)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 G11B G06K H04N

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;EPDOC;IEEE:字, 字幕, 文字, 文件, 文档, 编码, 字库, 格式, 解码, 识别, 乱码, 匹配, 一致, 对比, 比对, 内码, 表, 失败, decod+, encod+, text, string+, title, subtitle, caption, determing, recogniz+, detect+, code, format, type, match

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 104516862 A (北大方正集团有限公司 等) 2015年 4月 15日 (2015 - 04 - 15) 权利要求1、5、9、12, 说明书第[0065]-[0080]段, 附图1	1-20
Y	CN 102194503 A (腾讯科技深圳有限公司) 2011年 9月 21日 (2011 - 09 - 21) 权利要求1、4, 说明书第[0036]-[0039]、[0072]、[0074]段, 附图1-3	1-20
A	CN 102595082 A (深圳创维—RGB电子有限公司) 2012年 7月 18日 (2012 - 07 - 18) 全文	1-20
A	CN 104750666 A (明博教育科技有限公司) 2015年 7月 1日 (2015 - 07 - 01) 全文	1-20
A	CN 104156373 A (宏碁股份有限公司) 2014年 11月 19日 (2014 - 11 - 19) 全文	1-20
A	US 2013077855 A1 (ISYS SEARCH SOFTWARE PTY LTD.) 2013年 3月 28日 (2013 - 03 - 28) 全文	1-20

☒ 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

2016年 11月 29日

国际检索报告邮寄日期

2016年 12月 21日

ISA/CN的名称和邮寄地址

中华人民共和国国家知识产权局(ISA/CN)
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

胡雅琴

电话号码 (86-10)61648233

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
A	US 2006025988 A1 (XU, MING 等) 2006年 2月 2日 (2006 - 02 - 02) 全文	1-20

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/084872

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	104516862	A	2015年 4月 15日	EP	3051428	A1	2016年 8月 3日
				WO	2015043072	A1	2015年 4月 2日
				US	2016239467	A1	2016年 8月 18日
CN	102194503	A	2011年 9月 21日	CN	102194503	B	2014年 8月 27日
CN	102595082	A	2012年 7月 18日	无			
CN	104750666	A	2015年 7月 1日	无			
CN	104156373	A	2014年 11月 19日	无			
US	2013077855	A1	2013年 3月 28日	AU	2012201539	B2	2016年 6月 16日
				US	9122898	B2	2015年 9月 1日
				AU	2012201539	A1	2012年 12月 6日
US	2006025988	A1	2006年 2月 2日	US	7865355	B2	2011年 1月 4日

表 PCT/ISA/210 (同族专利附件) (2009年7月)