

(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6月8日 (08.06.2017)



(10) 国际公布号
WO 2017/092280 A1

- (51) 国际专利分类号:
G06F 17/30 (2006.01)
- (21) 国际申请号: PCT/CN2016/085839
- (22) 国际申请日: 2016年6月15日 (15.06.2016)
- (25) 申请语言: 中文
- (26) 公布语言: 中文
- (30) 优先权:
201510861520.6 2015年11月30日 (30.11.2015) CN
- (71) 申请人: 乐视控股(北京)有限公司 (LE HOLDINGS (BEIJING) CO., LTD.) [CN/CN]; 中国北京市朝阳区姚家园路105号3号楼10层1102, Beijing 100025 (CN)。乐视致新电子科技(天津)有限公司 (LE SHI ZHI XIN ELECTRONIC TECHNOLOGY (TIANJIN) LIMITED) [CN/CN]; 中国天津市滨海新区生态城动漫中路126号动漫大厦B1区2层201-427, Tianjin 300467 (CN)。
- (72) 发明人: 王振宇 (WANG, Zhenyu); 中国天津市滨海新区生态城动漫中路126号动漫大厦B1区2层201-427, Tianjin 300467 (CN)。
- (74) 代理人: 北京三聚阳光知识产权代理有限公司 (SUNSHINE INTELLECTUAL PROPERTY INTERNATIONAL CO., LTD); 中国北京市海淀区海淀南路甲21号中关村知识产权大厦A座5层503, Beijing 100080 (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,

[见续页]

(54) Title: MULTIMEDIA PHOTO GENERATION METHOD, APPARATUS AND DEVICE, AND MOBILE PHONE

(54) 发明名称: 一种多媒体照片生成方法、装置、设备及手机

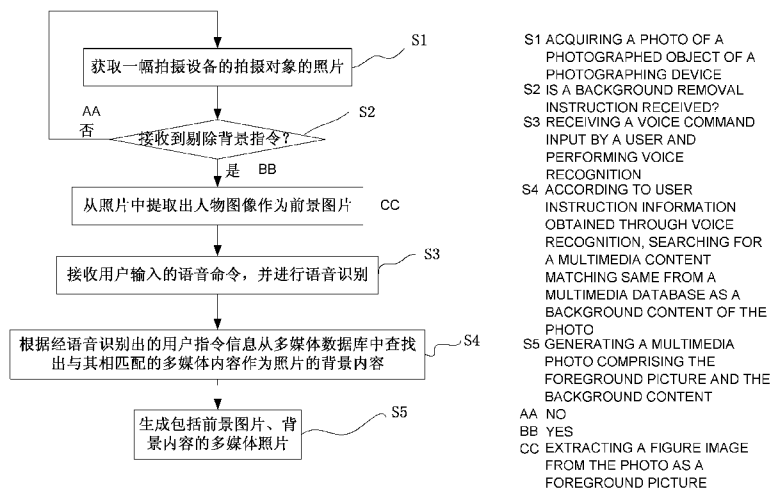


图1

(57) Abstract: Provided are a multimedia photo generation method, apparatus and device, and a mobile phone. The multimedia photo generation method comprises: firstly acquiring a photo of a photographed object of a photographing device; after an instruction for removing a photo background is received, extracting a figure image from the photo as a foreground picture; performing voice recognition after receiving a voice command input by the user; then, according to user instruction information obtained by means of voice recognition, searching for a multimedia content matching the user instruction information from a multimedia database as a background content of the photo; and generating a multimedia photo comprising the foreground picture and the background content. When the user wants to replace the photo background, the figure image can be automatically extracted from the photo as the foreground picture, an original background with a poor effect is removed, and a picture and/or music matching the user instruction information can be automatically searched from the multimedia database. Search efficiency is improved, optimization processing is simple, and the degree of user experience is high.

(57) 摘要:

[见续页]



WO 2017/092280 A1



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, 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)。

根据细则 4.17 的声明:

- 关于申请人有权要求在先申请的优先权(细则 4.17(iii))
- 发明人资格(细则 4.17(iv))

本国际公布:

- 包括国际检索报告(条约第 21 条(3))。

本发明提供了一种多媒体照片生成方法、装置、设备及手机，其中多媒体照片生成方法先获取一幅拍摄设备的拍摄对象的照片；并在接收到剔除照片背景的指令后从照片中提取出人物图像作为前景图片；在接收用户输入的语音命令后进行语音识别，然后根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为照片的背景内容，生成包括前景图片、背景内容的多媒体照片。在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，并从多媒体数据库中自动查找出与该用户指令信息相匹配的图片和/或音乐，提高了查找效率，优化处理简便，用户体验度高。

一种多媒体照片生成方法、装置、设备及手机

技术领域

本发明涉及多媒体技术领域，具体涉及一种多媒体照片生成方法、装置、设备及手机。

背景技术

智能手机、平板电脑等具有拍照功能的设备因具有便于携带、能快速传播信息等功能，已成为很好的社交化工具，随着普及率的不断提升，其添加的附属功能也越来越多，比如可以将拍摄的照片进行编辑后通过智能手机、平板电脑等内置的社交软件（比如微信、微博、QQ 聊天工具等）发送到社交平台进行分享。

目前如果用户想要对照片的背景进行修改，需要通过修图软件对照片进行分析、修复、美化、合成等处理，但由于修图软件专业性强，非常难于操作，对于非专业用户，需要学习和熟悉相应的软件操作方法才能达到用户想要的效果；即使是专业用户，在海量的图片中挑选与自己需求相匹配的图片，也需要耗费大量时间，自然也无法快速生成与自己需求相匹配的多媒体图片了，体验度较差。

发明内容

因此，本发明要解决的技术问题在于克服现有技术中用户对照片的背景进行优化处理时操作异常复杂，体验度差的缺陷，从而提供一种多媒体照片生成方法、装置、设备及手机。

为此，本发明的技术方案如下所述：

本发明提供了一种多媒体照片生成方法，用于拍摄设备，包括如下步骤：

获取一幅所述拍摄设备的拍摄对象的照片；

接收到剔除照片背景的指令后从所述照片中提取出人物图像作为前景图片；

接收用户输入的语音命令，并进行语音识别；

根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容；

生成包括所述前景图片、所述背景内容的多媒体照片。

本发明所述的多媒体照片生成方法，所述根据所述用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容的步骤包括：

以所述用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐，所述多媒体数据库中存储有与图片、音乐对应的关键字标引；

将与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

本发明所述的多媒体照片生成方法，所述以所述用户指令信息为关键

字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐的步骤之后还包括：

若接收到重新查找图片和/或音乐的指令，以所述用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令；

将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

本发明所述的多媒体照片生成方法，通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令。

本发明还提供了一种多媒体照片生成装置，用于拍摄设备，包括：

原始照片获取单元，用于获取一幅所述拍摄设备的拍摄对象的照片；

前景图片提取单元，用于在接收到剔除照片背景的指令后从所述照片中提取出人物图像作为前景图片；

指令识别单元，用于接收用户输入的语音命令，并进行语音识别；

匹配单元，用于根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容；

生成单元，用于生成包括所述前景图片、所述背景内容的多媒体照片。

本发明所述的多媒体照片生成装置，所述匹配单元以所述用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐，所述多媒体数据库中存储有与图片、音乐对应的关键字标引；并在没有接收到重新查找图片和/或音乐的指令时将与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

本发明所述的多媒体照片生成装置，所述匹配单元还用于在接收到重新查找图片和/或音乐的指令，以所述用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令；并将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

本发明所述的多媒体照片生成装置，所述指令识别单元还用于通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令。

本发明还提供了一种多媒体照片生成设备，包括：拍照装置、多媒体照片生成装置、显示装置以及播放装置；

所述拍照装置，用于拍摄所述拍摄设备的拍摄对象的照片；

所述多媒体照片生成装置，与所述拍照装置连接，用于接收所述拍照装置发送的照片，并在接收到剔除照片背景的指令后从所述照片中提取出人物图像作为前景图片，并在接收用户输入的语音命令后进行语音识别，

然后根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容，生成包括所述前景图片、所述背景内容的多媒体照片；

所述显示装置，与所述拍照装置和所述多媒体照片生成装置连接，用于显示所述照片以及所述多媒体照片；

所述播放装置，与所述多媒体照片生成装置连接，用于播放所述背景内容。

本发明还提供了一种手机，包括上述多媒体照片生成装置或上述多媒体照片生成设备。

本发明技术方案，具有如下优点：

本发明提供了一种多媒体照片生成方法及装置，先获取一幅拍摄设备的拍摄对象的照片；并在接收到剔除照片背景的指令后从照片中提取出人物图像作为前景图片；在接收用户输入的语音命令后进行语音识别，然后根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为照片的背景内容，生成包括前景图片、背景内容的多媒体照片。在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，无需用户通过专业的修图软件对照片背景进行优化，只需发出剔除照片背景的指令即可将背景更换为用户满意的图片，优化处理简便，用户体验度高。并且通过对用户的声音进行语音识别来获取用户指令信息，即可从多媒体数据库中自动查找出与该用户指令信息相匹配的多媒体内容，无需用户手动从海量

图片和/或音乐中选取出自己需求相匹配的多媒体内容，提高了查找的效率。之后生成包括前景图片、背景内容的多媒体照片，实现了视听的完美组合。

本发明所述多媒体照片生成方法及装置，若用户对查找出的图片和/或音乐不满意，会在收到重新查找图片和/或音乐的指令后，以用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令；之后将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。为生成用户满意度高的包含前景图片、背景内容的多媒体照片奠定了基础，用户体验度更好。

本发明所述多媒体照片生成方法及装置，通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令，无需用户手动输入指令，即可通过对用户发出的声音进行识别来接收到相应的指令并依据指令进行相应的操作，解放了用户的双手，能够快速响应用户的需求，提升了用户的体验度。

附图说明

为了更清楚地说明本发明具体实施方式或现有技术中的技术方案，下面将对具体实施方式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，下面描述中的附图是本发明的一些实施方式，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获

得其他的附图。

图1为本发明实施例1中多媒体照片生成方法的一个具体实例的流程图；

图2为本发明实施例1中多媒体照片生成方法中从多媒体数据库中查找出与该用户指令信息相匹配的图片和/或音乐的一个具体实例的流程图；

图3为本发明实施例2中多媒体照片生成装置的一个具体实例的结构框图；

图4为本发明实施例3中多媒体照片生成设备的一个具体实例的结构框图。

附图标记：

1-原始照片获取单元；2-前景图片提取单元；3-指令识别单元；4-匹配单元；5-生成单元；11-拍照装置；12-多媒体照片生成装置；13-显示装置；14-播放装置。

具体实施方式

下面将结合附图对本发明的技术方案进行清楚、完整地描述，显然，所描述的实施例是本发明一部分实施例，而不是全部的实施例。基于本发明中的实施例，本领域普通技术人员在没有做出创造性劳动前提下所获得的所有其他实施例，都属于本发明保护的范围。

在本发明的描述中，需要说明的是，术语“中心”、“上”、“下”、“左”、“右”、“竖直”、“水平”、“内”、“外”等指示的方位或位置关系为基于附

图所示的方位或位置关系，仅是为了便于描述本发明和简化描述，而不是指示或暗示所指的装置或元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。此外，术语“第一”、“第二”、“第三”仅用于描述目的，而不能理解为指示或暗示相对重要性。

在本发明的描述中，需要说明的是，除非另有明确的规定和限定，术语“安装”、“相连”、“连接”应做广义理解，例如，可以是固定连接，也可以是可拆卸连接，或一体地连接；可以是机械连接，也可以是电连接；可以是直接相连，也可以通过中间媒介间接相连，还可以是两个元件内部的连通，可以是无线连接，也可以是有线连接。对于本领域的普通技术人员而言，可以根据具体情况理解上述术语在本发明中的具体含义。

此外，下面所描述的本发明不同实施方式中所涉及的技术特征只要彼此之间未构成冲突就可以相互结合。

实施例 1

本实施例提供了一种多媒体照片生成方法，用于拍摄设备，拍摄设备包括手机、pad、笔记本电脑等具有拍摄功能的电子设备，如图 1 所示，包括如下步骤：

S1. 获取一幅拍摄设备的拍摄对象的照片。

S2. 接收到剔除照片背景的指令后从照片中提取出人物图像作为前景图片。具体地，可以通过轮廓分析法确定照片中的人物图像并提取出来作为前景图片，当然也可以选用现有技术中任何一种能够实现照片中人物图像提取的方法来实现上述操作。

S3. 接收用户输入的语音命令，并进行语音识别。具体地，可以选用现有技术中任何一款具有语音识别功能的软件来实现语音识别的功能。

S4. 根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为照片的背景内容。多媒体内容包括文本、图像、声音、动画和视频等，这些多媒体内容均可作为照片的背景内容与照片合成；本发明优选实施例以图片和/或音乐为例，优选的背景内容包括背景图片、背景音乐等。具体地，用户在拍照时有可能会对拍照的环境不满意，比如用户在家中拍照，但想要将照片背景更换成一些名胜古迹或者风景优美的地方，比如想要将照片背景更换成埃菲尔铁塔或者日本的富士山等，用户就可以发出包含埃菲尔铁塔或富士山等内容的声音，通过对用户的声音进行语音识别就可以获取用户指令信息为将包含埃菲尔铁塔或者富士山内容的图片作为照片的新的背景图片；又或者用户想要在照片中添加一些自己喜爱的音乐，比如用户想要添加一首自己喜爱的流行歌曲的名字，就可以直接说出该流行歌曲的歌名，通过语音识别就可以获取用户指令信息为将该流行歌曲作为照片的背景音乐，根据上述用户指令信息就可以自动从多媒体数据库中查找出与该用户指令信息相匹配的图片和/或音乐作为照片的背景图片和/或背景音乐了，确保了作为背景图片和背景音乐的图片和音乐符合用户的真实需求，提升了用户的体验度。

S5. 生成包括前景图片、背景内容的多媒体照片。

本实施例中的多媒体照片生成方法，在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，无需用户通过专业的修图软件对照片背景进行优化，只需发出剔除

照片背景的指令即可将背景更换为用户满意的图片，优化处理简便，用户体验度高。并且通过对用户的声音进行语音识别来获取用户指令信息，即可从多媒体数据库中自动查找出与该用户指令信息相匹配的多媒体内容，无需用户手动从海量图片和/或音乐中选取出自己需求相匹配的多媒体内容，提高了查找的效率。之后生成包括前景图片、背景内容的多媒体照片，实现了视听的完美组合。

优选地，如图 2 所示，步骤 S4 包括：

S41. 以用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐，多媒体数据库中存储有与图片、音乐对应的关键字标引。具体地，音乐和图片存入多媒体数据库时都会带有关键字标引，以便于后期的检索，常见的关键字标引方式为 TIF 格式的信息，通过读取图片、音乐的 TIF 信息就可以了解图片、音乐所展示的内容了，非常便捷。与用户指令信息关键字越贴近的关键字标引对应的图片和/或音乐，则说明其所展示的内容与用户的真实需求越接近，匹配度也越高，通过查找与该用户指令信息关键字相同或相近的关键字标引对应的图片和/或音乐，能够确保将与用户需求匹配度最高的图片和/或音乐推送给用户。

S42. 将与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

S43. 若接收到重新查找图片和/或音乐的指令，以用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关

键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令。之后返回步骤 S42，将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

本实施例中的多媒体照片生成方法，确保了最终选取的作为背景内容的图片和音乐都是用户满意的，为生成用户满意度高的包含前景图片、背景内容的多媒体照片奠定了基础，用户体验度更好。

优选地，本实施例中的多媒体照片生成方法，通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令。无需用户手动输入指令，即可通过对用户发出的声音进行识别来接收到相应的指令并依据指令进行相应的操作，解放了用户的双手，能够快速响应用户的需求，提升了用户的体验度。

实施例 2

本实施例提供了一种多媒体照片生成装置，用于拍摄设备，如图 3 所示，包括：

原始照片获取单元 1，用于获取一幅拍摄设备的拍摄对象的照片。

前景图片提取单元 2，用于在接收到剔除照片背景的指令后从照片中提取出人物图像作为前景图片。

指令识别单元 3，用于接收用户输入的语音命令，并进行语音识别。

匹配单元 4，用于根据经语音识别出的用户指令信息从多媒体数据库中

查找出与该用户指令信息相匹配的多媒体内容作为照片的背景内容。

生成单元 5，用于生成包括前景图片、背景内容的多媒体照片。

本实施例中的多媒体照片生成装置，在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，无需用户通过专业的修图软件对照片背景进行优化，只需发出剔除照片背景的指令即可将背景更换为用户满意的图片，优化处理简便，用户体验度高。并且通过对用户的声音进行语音识别来获取用户指令信息，即可从多媒体数据库中自动查找出与该用户指令信息相匹配的多媒体内容，无需用户手动从海量图片和/或音乐中选取出自已需求相匹配的多媒体内容，提高了查找的效率。之后生成包括前景图片、背景内容的多媒体照片，实现了视听的完美组合。

优选地，匹配单元 4 以用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐，多媒体数据库中存储有与图片、音乐对应的关键字标引；并在没有接收到重新查找图片和/或音乐的指令时将与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

优选地，匹配单元 4 还用于在接收到重新查找图片和/或音乐的指令，以用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令；并将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指

令信息相匹配的多媒体内容。

本实施例中的多媒体照片生成装置，确保了最终选取的作为背景内容的图片和音乐都是用户满意的，为生成用户满意度高的包含前景图片、背景内容的多媒体照片奠定了基础，用户体验度更好。

优选地，本实施例中的多媒体照片生成装置，指令识别单元 3 还用于通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令。无需用户手动输入指令，即可通过对用户发出的声音进行识别来接收到相应的指令并依据指令进行相应的操作，解放了用户的双手，能够快速响应用户的需求，提升了用户的体验度。

实施例 3

本实施例提供了一种多媒体照片生成设备，包括：拍照装置 11、多媒体照片生成装置 12、显示装置 13 以及播放装置 14。

拍照装置 11，用于拍摄拍摄设备的拍摄对象的照片。具体地，拍照装置 11 可以包括摄像头、闪光灯等部件。

多媒体照片生成装置 12，与拍照装置 11 连接，用于接收拍照装置 11 发送的照片，并在接收到剔除照片背景的指令后从照片中提取出人物图像作为前景图片，并在接收用户输入的语音命令后进行语音识别，然后根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为照片的背景内容，生成包括前景图片、背景内容的多媒体照片。具体地，多媒体照片生成装置 12 可以为控制器，内置能够

执行实施例 1 中的步骤的程序。实现了照片的自动优化处理。多媒体照片生成装置 12 既可以访问拍摄设备内存的多媒体数据库，也可以访问拍摄设备以外的存储设备存储的多媒体数据库，为查找符合用户需求的图片和/或音乐提供了技术上的支持。

显示装置 13，与拍照装置 11 和多媒体照片生成装置 12 连接，用于显示照片以及多媒体照片。具体地，显示装置 13 可以为显示屏，通过将拍摄的照片呈现给用户，能够使用户及时判断是否需要剔除背景，通过将查找的图片呈现给用户，能够使用户及时判断是否需要重新匹配图片，并能使用户第一时间欣赏到优化生成的多媒体照片。

播放装置 14，与多媒体照片生成装置 12 连接，用于播放背景内容。具体地，播放装置 14 可以为微型电子喇叭，通过将查找出的音乐播放给用户，能够使用户及时判断是否需要重新匹配音乐，并能使用户第一时间欣赏到多媒体照片中的背景内容中的音乐。

本实施例中的多媒体照片生成设备，在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，无需用户通过专业的修图软件对照片背景进行优化，只需发出剔除照片背景的指令即可将背景更换为用户满意的图片，优化处理简便，用户体验度高。并且通过对用户的语音进行语音识别来获取用户指令信息，即可从多媒体数据库中自动查找出与该用户指令信息相匹配的图片和/或音乐，无需用户手动从海量图片和音乐中选取出自己需求相匹配的图片和音乐，提高了查找的效率。之后生成包括前景图片、背景图片和背景音乐的多媒体照片，实现了视听的完美组合。并且无需用户手动输入指令，即

可通过对用户发出的声音进行识别来接收到相应的指令并依据指令进行相应的操作，解放了用户的双手，能够快速响应用户的需求，提升了用户的体验度。

实施例 4

本实施例提供了一种手机，包括实施例 2 中的多媒体照片生成装置或实施例 3 中的多媒体照片生成设备。

本实施例中的手机，在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，无需用户通过专业的修图软件对照片背景进行优化，只需发出剔除照片背景的指令即可将背景更换为用户满意的图片，优化处理简便，用户体验度高。并且通过对用户的声音进行语音识别来获取用户指令信息，即可从多媒体数据库中自动查找出与该用户指令信息相匹配的多媒体内容，无需用户手动从海量图片和/或音乐中选取出自己需求相匹配的图片和/或音乐，提高了查找的效率。之后生成包括前景图片、背景内容的多媒体照片，实现了视听的完美组合。并且无需用户手动输入指令，即可通过对用户发出的声音进行识别来接收到相应的指令并依据指令进行相应的操作，解放了用户的双手，能够快速响应用户的需求，提升了用户的体验度。

本领域内的技术人员应明白，本发明的实施例可提供为方法、系统、或计算机程序产品。因此，本发明可采用完全硬件实施例、完全软件实施例、或结合软件和硬件方面的实施例的形式。而且，本发明可采用在一个或多个其中包含有计算机可用程序代码的计算机可用存储介质（包括但不

限于磁盘存储器、CD-ROM、光学存储器等）上实施的计算机程序产品的形式。

本发明是参照根据本发明实施例的方法、设备（系统）、和计算机程序产品的流程图和 / 或方框图来描述的。应理解可由计算机程序指令实现流程图和 / 或方框图中的每一流程和 / 或方框、以及流程图和 / 或方框图中的流程和 / 或方框的结合。可提供这些计算机程序指令到通用计算机、专用计算机、嵌入式处理机或其他可编程数据处理设备的处理器以产生一个机器，使得通过计算机或其他可编程数据处理设备的处理器执行的指令产生用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的装置。

这些计算机程序指令也可存储在能引导计算机或其他可编程数据处理设备以特定方式工作的计算机可读存储器中，使得存储在该计算机可读存储器中的指令产生包括指令装置的制造品，该指令装置实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能。

这些计算机程序指令也可装载到计算机或其他可编程数据处理设备上，使得在计算机或其他可编程设备上执行一系列操作步骤以产生计算机实现的处理，从而在计算机或其他可编程设备上执行的指令提供用于实现在流程图一个流程或多个流程和 / 或方框图一个方框或多个方框中指定的功能的步骤。

工业实用性

本发明实施例中的多媒体照片生成方法、装置、设备及手机适合于多

媒体技术领域，特别地，适合于具有拍照功能的多媒体设备用户对照片的背景进行优化处理。在用户想要更换照片背景时，能自动从照片中提取出人物图像作为前景图片并将原有的效果差的背景予以剔除，无需用户通过专业的修图软件对照片背景进行优化，只需发出剔除照片背景的指令即可将背景更换为用户满意的图片，优化处理简便，用户体验度高；并且通过对用户的声音进行语音识别来获取用户指令信息，即可从多媒体数据库中自动查找出与该用户指令信息相匹配的多媒体内容，无需用户手动从海量图片和/或音乐中选取出自已需求相匹配的多媒体内容，提高了查找的效率；之后生成包括前景图片、背景内容的多媒体照片，实现了视听的完美组合，具有有益效果。因此，本发明的技术方案具有工业实用性。

权 利 要 求 书

1. 一种多媒体照片生成方法，用于拍摄设备，其特征在于，包括如下步骤：

获取一幅所述拍摄设备的拍摄对象的照片；

接收到剔除照片背景的指令后从所述照片中提取出人物图像作为前景图片；

接收用户输入的语音命令，并进行语音识别；

根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容；

生成包括所述前景图片、所述背景内容的多媒体照片。

2. 根据权利要求 1 所述的多媒体照片生成方法，其特征在于，所述根据所述用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容的步骤包括：

以所述用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐，所述多媒体数据库中存储有与图片、音乐对应的关键字标引；

将与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

3. 根据权利要求 2 所述的多媒体照片生成方法，其特征在于，所述以

所述用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐的步骤之后还包括：

若接收到重新查找图片和/或音乐的指令，以所述用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令；

将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

4. 根据权利要求 1-3 任一项所述的多媒体照片生成方法，其特征在于，通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令。

5. 一种多媒体照片生成装置，用于拍摄设备，其特征在于，包括：

原始照片获取单元（1），用于获取一幅所述拍摄设备的拍摄对象的照片；

前景图片提取单元（2），用于在接收到剔除照片背景的指令后从所述照片中提取出人物图像作为前景图片；

指令识别单元（3），用于接收用户输入的语音命令，并进行语音识别；

匹配单元（4），用于根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内

容；

生成单元（5），用于生成包括所述前景图片、所述背景内容的多媒体照片。

6. 根据权利要求 5 所述的多媒体照片生成装置，其特征在于，所述匹配单元（4）以所述用户指令信息为关键字，从多媒体数据库中查找出与该关键字相同或相近的关键字标引对应的图片和/或音乐，所述多媒体数据库中存储有与图片、音乐对应的关键字标引；并在没有接收到重新查找图片和/或音乐的指令时将与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

7. 根据权利要求 6 所述的多媒体照片生成装置，其特征在于，所述匹配单元（4）还用于在接收到重新查找图片和/或音乐的指令，以所述用户指令信息为关键字从多媒体数据库中查找除之前已查找出的图片和/或音乐以外的与该关键字相同或相近的关键字标引对应的图片和/或音乐，直至不再接收到重新查找图片和/或音乐的指令；并将最终查找出的与该关键字相同或相近的关键字标引对应的图片和/或音乐作为与该用户指令信息相匹配的多媒体内容。

8. 根据权利要求 5-7 任一项所述的多媒体照片生成装置，其特征在于，所述指令识别单元（3）还用于通过对用户的声音进行识别来判断是否接收到剔除照片背景的指令以及是否接收到重新查找图片和/或音乐的指令。

9. 一种多媒体照片生成设备，其特征在于，包括：拍照装置（11）、多媒体照片生成装置（12）、显示装置（13）以及播放装置（14）；

所述拍照装置（11），用于拍摄所述拍摄设备的拍摄对象的照片；

所述多媒体照片生成装置（12），与所述拍照装置（11）连接，用于接收所述拍照装置（11）发送的照片，并在接收到剔除照片背景的指令后从所述照片中提取出人物图像作为前景图片，并在接收用户输入的语音命令后进行语音识别，然后根据经语音识别出的用户指令信息从多媒体数据库中查找出与该用户指令信息相匹配的多媒体内容作为所述照片的背景内容，生成包括所述前景图片、所述背景内容的多媒体照片；

所述显示装置（13），与所述拍照装置（11）和所述多媒体照片生成装置（12）连接，用于显示所述照片以及所述多媒体照片；

所述播放装置（14），与所述多媒体照片生成装置（12）连接，用于播放所述背景内容。

10. 一种手机，其特征在于，包括权利要求 5-8 任意一项所述的多媒体照片生成装置或权利要求 9 所述的多媒体照片生成设备。

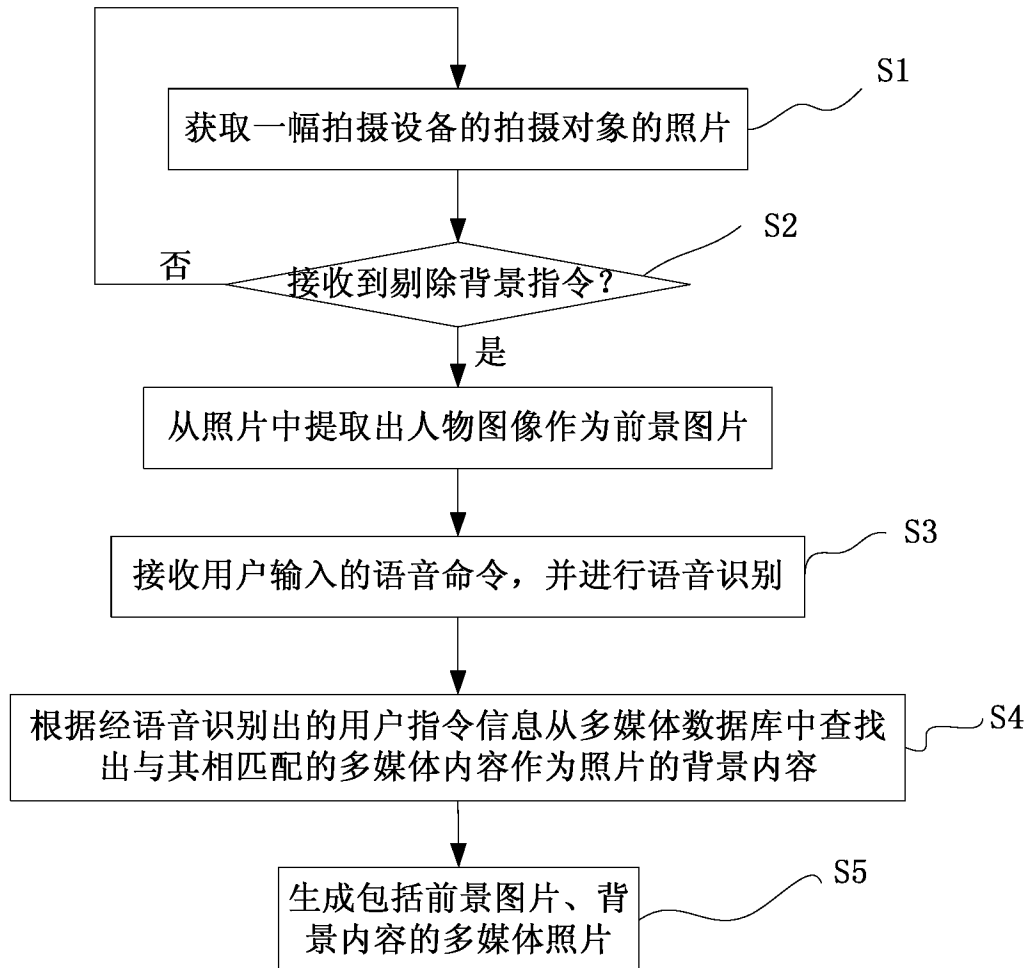


图 1

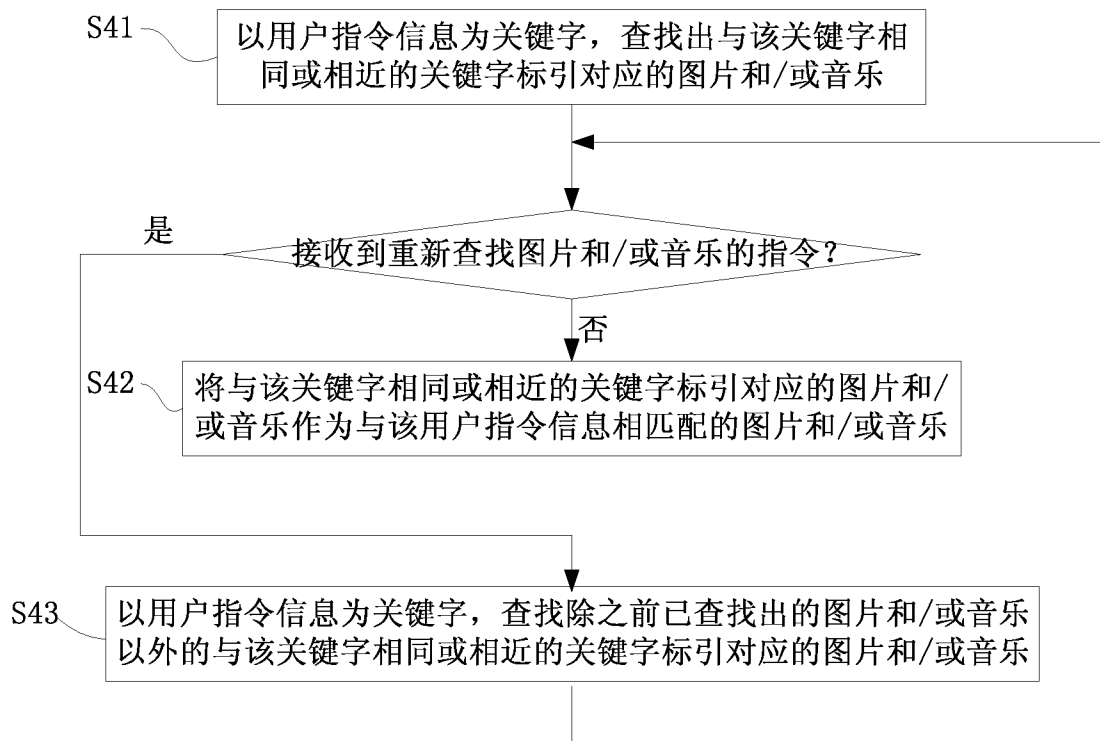


图 2

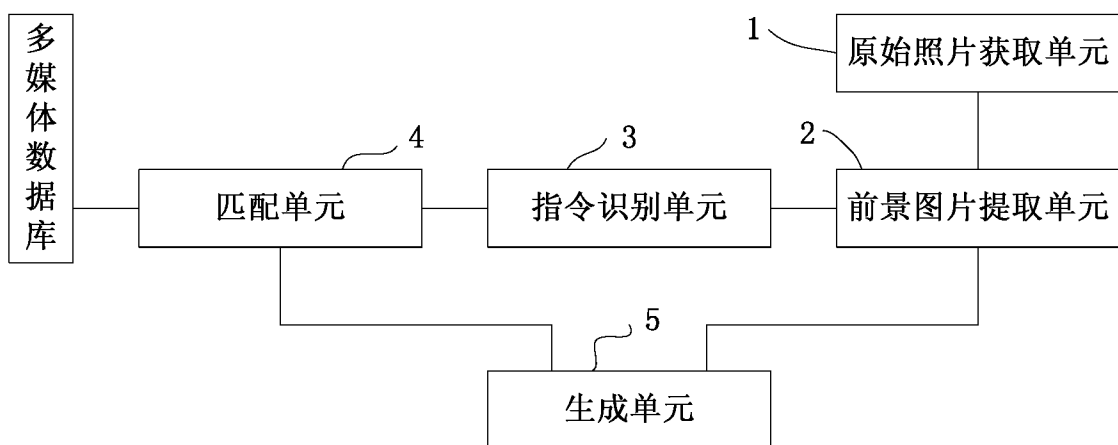


图 3

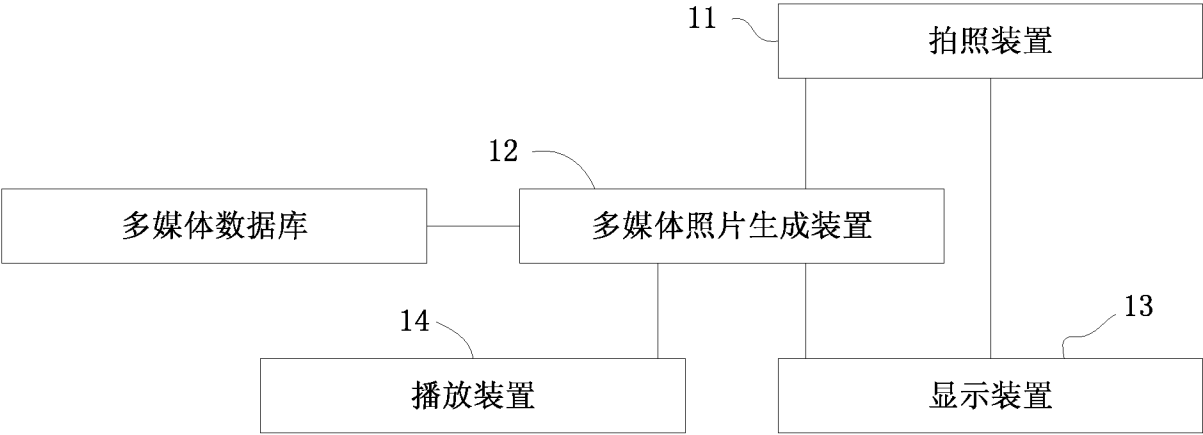


图 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2016/085839**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; H04M; G06K; G06T; 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, WPI, EPODOC, CNKI: combination, instruction, picture, photo+, image, synthesis+, synthesiz+, multimedia, portrait, person, foreground, background, separate, voice, audio, sound, identifi+, database, keyword, index+, 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 105100491 A (NUBIA TECHNOLOGY CO., LTD.), 25 November 2015 (25.11.2015), description, paragraphs [0082]-[0163], and figure 4	1-10
Y	CN 101853253 A (SAMSUNG ELECTRONICS CO., LTD. et al.), 06 October 2010 (06.10.2010), description, paragraphs [0086]-[0093], and figures 6-7	1-10
A	CN 104144097 A (BAIDU ONLINE NETWORK TECHNOLOGY (BEIJING) CO., LTD.), 12 November 2014 (12.11.2014), the whole document	1-10
A	CN 101216841 A (NANJING SEEKPAI INFORMATION TECHNOLOGY CO., LTD.), 09 July 2008 (09.07.2008), the whole document	1-10

☐ 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 30 August 2016 (30.08.2016)	Date of mailing of the international search report 14 September 2016 (14.09.2016)
Name and mailing address of the ISA/CN: 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 ZHAO, Xiaoning Telephone No.: (86-10) 62412071

International application No.
PCT/CN2016/085839

Form PCT/ISA/210 (patent family annex) (July 2009)

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; H04M; G06K; G06T; 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, WPI, EPDOC, CNKI: 图像, 图片, 照片, 相片, 合成, 多媒体, 组合, 人物, 人像, 前景, 背景, 分离, 语音, 指令, 识别, 数据库, 关键词, 关键字, 标引, 索引, 匹配, 音频, picture, photo+, image, synthesis+, synthesiz+, multimedia, portrait, person, foreground, background, separate, voice, audio, sound, identifi+, database, keyword, index+, 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 105100491 A (努比亚技术有限公司) 2015年 11月 25日 (2015 - 11 - 25) 说明书第[0082]-[0163]段及附图4	1-10
Y	CN 101853253 A (三星电子株式会社 等) 2010年 10月 6日 (2010 - 10 - 06) 说明书第[0086]-[0093]段及附图6-7	1-10
A	CN 104144097 A (百度在线网络技术北京有限公司) 2014年 11月 12日 (2014 - 11 - 12) 全文	1-10
A	CN 101216841 A (南京搜拍信息技术有限公司) 2008年 7月 9日 (2008 - 07 - 09) 全文	1-10
☐ 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 8月 30日		国际检索报告邮寄日期 2016年 9月 14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		授权官员 赵小宁 电话号码 (86-10) 62412071

国际检索报告
关于同族专利的信息

国际申请号

PCT/CN2016/085839

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	105100491	A	2015年 11月 25日	无	
CN	101853253	A	2010年 10月 6日	无	
CN	104144097	A	2014年 11月 12日	无	
CN	101216841	A	2008年 7月 9日	CN 100578508	C 2010年 1月 6日