(19)中华人民共和国国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 111462548 A
(43)申请公布日 2020.07.28

(21)申请号 201910054314.2

(22)申请日 2019.01.21

(71)申请人 北京字节跳动网络技术有限公司
地址 100080 北京市石景山区实兴大街30号院3号楼2层B-0035房间

(72)发明人 不公告发明人

(74)专利代理机构 北京品源专利代理有限公司 11332
代理人 孟金喆

(51)Int.Cl.
G09B 5/06(2006.01)

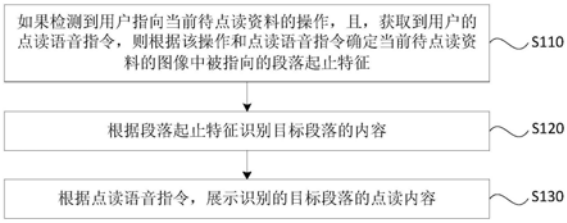
权利要求书2页 说明书10页 附图4页

(54)发明名称

一种段落点读方法、装置、设备和可读介质

(57)摘要

本公开实施例公开了一种段落点读方法、装置、设备和可读介质。其中，该方法包括：如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征；根据所述段落起止特征识别目标段落的内容；根据所述点读语音指令，展示识别的目标段落的点读内容。本公开实施例提供的技术方案，提高了段落点读的便利性和实时性，通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，降低了段落点读的局限性，提高了点读内容识别的准确性，增强了用户的使用体验。



CN 111462548 A

1.一种段落点读方法,其特征在于,包括:

如果检测到用户指向当前待点读资料的操作,且,获取到用户的点读语音指令,则根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征;

根据所述段落起止特征识别目标段落的内容;

根据所述点读语音指令,展示识别的目标段落的点读内容。

2.根据权利要求1所述的方法,其特征在于,根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征,包括:

根据所述操作获取当前待点读资料的图像中用户手指的关键特征点;

根据所述关键特征点确定手指对应的被指向位置;

确定与所述被指向位置关联的段落起始特征和段落终止特征。

3.根据权利要求1所述的方法,其特征在于,根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征,包括:

将当前待点读资料的图像输入预先构建的神经网络模型中,得到所述操作对应的被指向位置;

确定与所述被指向位置关联的段落起始特征和段落终止特征。

4.根据权利要求2或3所述的方法,其特征在于,根据所述段落起止特征识别目标段落的内容,包括:

根据所述段落起始特征和所述段落终止特征确定对应的目标段落;

采用光学字符识别算法识别所述目标段落的内容。

5.根据权利要求1所述的方法,其特征在于,根据所述点读语音指令,展示识别的目标段落的点读内容,包括:

采用自然语言处理算法识别所述点读语音指令,得到指令关键词;

根据所述指令关键词确定识别的目标段落内容的点读结果和对应的点读方式;

根据对应的点读方式展示所述点读结果。

6.一种段落点读装置,其特征在于,包括:

段落特征确定模块,用于如果检测到用户指向当前待点读资料的操作,且,获取到用户的点读语音指令,则根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征;

段落识别模块,用于根据所述段落起止特征识别目标段落的内容;

段落点读模块,用于根据所述点读语音指令,展示识别的目标段落的点读内容。

7.根据权利要求6所述的装置,其特征在于,所述段落特征确定模块包括:

特征点获取单元,用于根据所述操作获取当前待点读资料的图像中用户手指的关键特征点;

第一指向确定单元,用于根据所述关键特征点确定手指对应的被指向位置;

第一特征确定单元,用于确定与所述被指向位置关联的段落起始特征和段落终止特征。

8.根据权利要求6所述的装置,其特征在于,所述段落特征确定模块包括:

第二指向确定单元,用于将当前待点读资料的图像输入预先构建的神经网络模型中,得到所述操作对应的被指向位置;

第二特征确定单元，用于确定与所述被指向位置关联的段落起始特征和段落终止特征。

9.一种设备，其特征在于，所述设备包括：

一个或多个处理器；

存储器，用于存储一个或多个程序；

当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如权利要求1-5中任一所述的段落点读方法。

10.一种可读介质，其上存储有计算机程序，其特征在于，该程序被处理器执行时实现如权利要求1-5中任一所述的段落点读方法。

一种段落点读方法、装置、设备和可读介质

技术领域

[0001] 本公开实施例涉及计算机处理技术，尤其涉及一种段落点读方法、装置、设备和可读介质。

背景技术

[0002] 在现有的点读设备中，一般会为每个点读设备配置一个专用的点读笔，通过该点读笔逐行点击书本上的某一段落所在位置区域，从而获取该位置区域的文字信息进行识别，如将点读笔按压扫描书本中的某一段落的位置区域，根据该按压操作获取该位置区域内各位置点的坐标，从而将该位置区域内各位置点上的文字信息进行输出，得到该段落的整体信息。

[0003] 此时要求用户通过配置的点读笔，才能实现书本上的段落信息点读，在点读笔丢失时，无法使用该点读设备来识别书本中的段落信息，存在一定的点读局限性。

发明内容

[0004] 有鉴于此，本公开实施例提供了一种段落点读方法、装置、设备和可读介质，解决了现有技术中必须通过点读笔点读书本段落信息的问题，降低了段落点读的局限性，提高了段落点读的便利性。

[0005] 第一方面，本公开实施例提供了一种段落点读方法，该方法包括：

[0006] 如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征；

[0007] 根据所述段落起止特征识别目标段落的内容；

[0008] 根据所述点读语音指令，展示识别的目标段落的点读内容。

[0009] 进一步的，根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征，包括：

[0010] 根据所述操作获取当前待点读资料的图像中用户手指的关键特征点；

[0011] 根据所述关键特征点确定手指对应的被指向位置；

[0012] 确定与所述被指向位置关联的段落起始特征和段落终止特征。

[0013] 进一步的，根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征，包括：

[0014] 将当前待点读资料的图像输入预先构建的神经网络模型中，得到所述操作对应的被指向位置；

[0015] 确定与所述被指向位置关联的段落起始特征和段落终止特征。

[0016] 进一步的，根据所述段落起止特征识别目标段落的内容，包括：

[0017] 根据所述段落起始特征和所述段落终止特征确定对应的目标段落；

[0018] 采用光学字符识别算法识别所述目标段落的内容。

[0019] 进一步的，根据所述点读语音指令，展示识别的目标段落的点读内容，包括：

[0020] 采用自然语言处理算法识别所述点读语音指令，得到指令关键词；

[0021] 根据所述指令关键词确定识别的目标段落内容的点读结果和对应的点读方式；

[0022] 根据对应的点读方式展示所述点读结果。

[0023] 第二方面，本公开实施例提供了一种段落点读装置，该装置包括：

[0024] 段落特征确定模块，用于如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据所述操作和所述点读语音指令确定当前待点读资料的图像中被指向的段落起止特征；

[0025] 段落识别模块，用于根据所述段落起止特征识别目标段落的内容；

[0026] 段落点读模块，用于根据所述点读语音指令，展示识别的目标段落的点读内容。

[0027] 进一步的，所述段落特征确定模块，包括：

[0028] 特征点获取单元，用于根据所述操作获取当前待点读资料的图像中用户手指的关键特征点；

[0029] 第一指向确定单元，用于根据所述关键特征点确定手指对应的被指向位置；

[0030] 第一特征确定单元，用于确定与所述被指向位置关联的段落起始特征和段落终止特征。

[0031] 进一步的，所述段落特征确定模块，还包括：

[0032] 第二指向确定单元，用于将当前待点读资料的图像输入预先构建的神经网络模型中，得到所述操作对应的被指向位置；

[0033] 第二特征确定单元，用于确定与所述被指向位置关联的段落起始特征和段落终止特征。

[0034] 进一步的，所述段落识别模块，包括：

[0035] 目标段落确定单元，用于根据所述段落起始特征和所述段落终止特征确定对应的目标段落；

[0036] 目标段落识别单元，用于采用光学字符识别算法识别所述目标段落的内容。

[0037] 进一步的，所述段落点读模块，包括：

[0038] 关键词确定单元，用于采用自然语言处理算法识别所述点读语音指令，得到指令关键词；

[0039] 点读模式确定单元，用于根据所述指令关键词确定识别的目标段落内容的点读结果和对应的点读方式；

[0040] 段落点读单元，用于点读根据对应的点读方式展示所述点读结果。

[0041] 第三方面，本公开实施例还提供了一种设备，该设备包括：

[0042] 一个或多个处理器；

[0043] 存储器，用于存储一个或多个程序；

[0044] 当所述一个或多个程序被所述一个或多个处理器执行，使得所述一个或多个处理器实现如本公开任意实施例所述的段落点读方法。

[0045] 第四方面，本公开实施例提供了一种可读介质，其上存储有计算机程序，该程序被处理器执行时实现如本公开任意实施例所述的段落点读方法。

[0046] 本公开实施例提供的一种段落点读方法、装置、设备和可读介质，在检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令时，确定当前待点读资料的图像

中被指向的段落起止特征，并识别该段落起止特征确定的目标段落内的内容，进而根据点读语音指令，展示该目标段落的点读内容，提高了段落点读的便利性和实时性，通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，降低了段落点读的局限性，提高了点读内容识别的准确性，增强了用户的使用体验。

附图说明

[0047] 为了更清楚地说明本公开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做一简单地介绍，显而易见地，下面描述中的附图是本公开的一些实施例，对于本领域普通技术人员来讲，在不付出创造性劳动的前提下，还可以根据这些附图获得其他的附图。

[0048] 图1A示出了本公开实施例提供的一种段落点读方法的流程图；

[0049] 图1B示出了本公开实施例提供的段落点读过程的原理示意图；

[0050] 图2示出了本公开实施例提供的一种段落点读过程中确定被指向的段落起止特征的原理示意图；

[0051] 图3示出了本公开实施例提供的另一种段落点读过程中确定被指向的段落起止特征的原理示意图；

[0052] 图4示出了本公开实施例提供的一种段落点读装置的结构示意图；

[0053] 图5示出了本公开实施例提供的一种设备的结构示意图。

具体实施方式

[0054] 为使本公开的目的、技术方案和优点更加清楚，以下将参照本公开实施例中的附图，通过实施方式清楚、完整地描述本公开的技术方案，显然，所描述的实施例是本公开一部分实施例，而不是全部的实施例。基于本公开中的实施例，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，都属于本公开保护的范围。

[0055] 图1A示出了本公开实施例提供的一种段落点读方法的流程图，本公开实施例可适用于对书本上的资料内容进行点读的过程中。本实施例提供的一种段落点读方法可以由本公开实施例提供的段落点读装置来执行，该装置可以通过软件和/或硬件的方式来实现，并集成在执行本方法的设备中，在本实施例中执行本方法的设备可以是用于点读的智能终端。

[0056] 具体的，如图1A所示，本公开实施例中提供的段落点读方法可以包括如下步骤：

[0057] S110，如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据该操作和点读语音指令确定当前待点读资料的图像中被指向的段落起止特征。

[0058] 其中，本实施例中段落点读方法主要应用在对书本资料中的段落内容进行点读的点读设备中，该点读设备无需配置专用的点读笔来辅助实现点读功能，用户在使用该点读设备对书本资料中的段落内容进行点读时，首先会开启该点读设备上预先配置的摄像器件，用于在用户将待点读资料放置于点读设备的可采集区域时，能够实时将待点读资料预览在该点读设备的屏幕中，以供用户在屏幕上观看该待点读资料；同时用户会发出对应的点读语音，点读设备通过预先配置的语音采集器来获取并分析该点读语音，判断用户本次的点读意图，从而实现后续的点读功能。其中，由于本实施例中的点读设备主要用于学生的

学习辅助过程中，因此待点读资料可以是各类教学领域内的相关资料，例如各类书本、课件或者习题等。

[0059] 同时，用户在需要对书本资料中的某一段落内容进行点读时，会将用户手指指向待点读资料中该段落内容所在的位置，以明确本次需要点读的待点读资料中包含的具体段落内容，本实施例中直接通过用户手指指向对应的点读位置，而无需采用专门配置的点读笔来按压书本中具体的点读位置，此时用户指向当前待点读资料的操作是指点读设备在屏幕上显示的待点读资料的图像中用户通过手指指向待点读资料中的点读位置的操作；点读语音指令是用于指示当前需要对用户手指指向的待点读资料中的相应段落内容进行点读的触发指令。

[0060] 可选的，用户在需要使用本实施例中的点读设备来点读书本资料中的段落内容时，首先会开启该点读设备中配置的摄像器件，并将待点读资料对应放置于该点读设备的图像采集区域，以使待点读资料在该点读设备的屏幕上预览；在用户需要点读该待点读资料中的具体段落内容时，首先会将手指指向图像采集区域中待点读资料中的段落内容所在的特定位置，并发出对应的点读语音，此时点读设备可以在屏幕上检测到用户指向待点读资料的操作，并通过预先配置的语音采集器会获取到用户的点读语音指令，表示用户当前需要对待点读资料中的某一段落内容进行点读，因此首先获取当前待点读资料的图像，后续根据用户指向待点读资料的具体操作对应的指向位置以及点读语音指令中包含的段落点读信息，在当前待点读资料的图像中判断被指向的本次需要点读的具体段落内容所在的段落起止特征。其中，该段落起止特征是指能够明确表示段落开始和结束的特征标识，可以是相邻的存在空格两个字长度的特征区域，如某一段落首行开始空格两个字，该段落结束转换到下一段落时，下一段落首行又会空格两个字，此时空格两个字可以作为本实施例中的段落起止特征。

[0061] S120，根据段落起止特征识别目标段落的内容。

[0062] 其中，目标段落是指待点读资料中本次需要点读的具体段落。

[0063] 具体的，在根据用户指向待点读资料的操作和点读语音指令确定当前待点读资料的图像中被指向的段落起止特征时，可以在待点读资料中确定出由该段落起止特征之间的文字信息组成的段落内容，作为目标段落；从而采用相应的字符识别技术识别目标段落内包含的文字内容，得到本次点读对应的具体内容。

[0064] S130，根据点读语音指令，展示识别的目标段落的点读内容。

[0065] 可选的，在根据段落起止特征识别目标段落的内容后，可以对点读语音指令进行分析，判断用户本次点读的点读意图，如判断点读语音指令中针对的是翻译、朗读还是查找段落答案等，进而根据点读语音指令中携带的不同点读意图，可以进一步在后台服务器或者云服务端中查找与识别的目标段落的内容对应的符合本次点读意图的点读结果，进而展示该目标段落对应的点读结果；示例性的，若点读语音指令为“请翻译一下文言文中的这一段落”，此时可以根据识别的目标段落内包含的该文言文的内容，查找该段文言文内容对应的翻译结果，并在屏幕上展示给对应的用户，实现相应的点读功能。

[0066] 本公开实施例提供的技术方案，在检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令时，确定当前待点读资料的图像中被指向的段落起止特征，并识别该段落起止特征确定的目标段落内的内容，进而根据点读语音指令，展示该目标段落的

点读内容，提高了段落点读的便利性和实时性，通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，降低了段落点读的局限性，提高了点读内容识别的准确性，增强了用户的使用体验。

[0067] 在上述实施例提供的技术方案的基础上，对于本公开实施例提供的段落点读方法中在目标段落内识别的内容的点读结果进行展示的情况进行进一步说明。示例性的，如图1B所示，本实施例中根据点读语音指令，展示识别的目标段落的点读内容，具体可以包括：采用自然语言处理算法识别点读语音指令，得到指令关键词；根据指令关键词确定识别的目标段落内容的点读结果和对应的点读方式；根据对应的点读方式展示点读结果。

[0068] 可选的，本实施例在解析点读语音指令时，可以采用自然语言处理(Natural Language Processing，NLP)算法来识别，得到点读语音指令中包含的指令关键词，该指令关键词可以明确用户本次点读的点读意图，如点读语音指令为“请翻译一下这个段落的内容”时，得到的指令关键词为“翻译”和“段落”，则可以确定本次点读是对段落进行翻译；此时根据得到的指令关键词可以在后台服务器或者云服务端中查找与本次识别的目标段落内包含的内容对应的点读结果，如指令关键词为“翻译”时，可以查找出目标段落内包含内容的翻译结果；同时根据指令关键词确定对应的点读方式，如指令关键词为“翻译”和“段落”，并没有明确指出本次的点读方式，此时将默认的点读方式作为本次的点读方式，如直接将翻译结果显示在屏幕上；如果点读语音指令为“请读一下这个段落内的内容”时，可以确定本次的点读方式为语音播放。进一步的，在确定本次点读对应的点读结果和点读方式时，点读设备可以采用该点读方式来展示对应的点读结果；本实施例中全程通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，提高了点读的便利性和实时性。

[0069] 本公开实施例中对于目标段落的段落起止特征的具体确定主要分为两种情况，1)通过设定的指向规则判断对应的段落起止特征；2)通过模型训练的方法判断对应的段落起止特征；以下分别对这两种情况进行详细的解释说明。

[0070] 图2示出了本公开实施例提供的一种段落点读过程中确定被指向的段落起止特征的原理示意图，本实施例是在上述提供的各个可选方案的基础上进行优化。具体的，本实施例对于通过设定的指向规则判断对应的段落起止特征的具体确定过程进行详细的介绍。

[0071] 可选的，该方法具体可以包括如下步骤：

[0072] S210，如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据该操作获取当前待点读资料的图像中用户手指的关键特征点。

[0073] 具体的，在用户需要对待点读资料中的某一段落内容进行点读时，会将手指指向待点读资料中该段落所在的位置，以明确本次需要点读的待点读资料中包含的具体段落内容，并发出对应的点读语音，此时点读设备会检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，说明当前需要对待点读资料中的相应段落内容进行点读，此时首先获取当前待点读资料的图像，如图2所示，并根据该图像中用户指向当前待点读资料的操作确定图像中用户手指的关键特征点21，该关键特征点21是指用户手指上的主要关节点，以便后续确定手指在该图像中的指向位置。

[0074] S220，根据关键特征点确定手指对应的被指向位置。

[0075] 可选的，在获取到当前待点读资料的图像中用户手指的关键特征点21时，可以根

据该关键特征点21所在的位置，确定出用户手指指尖的位置，进而确定用户手指在待点读资料的图像中对应的被指向位置22，该被指向位置22可以为本次需要点读的段落内容中包括的任一文字所在的位置。

[0076] S230，确定与被指向位置关联的段落起始特征和段落终止特征。

[0077] 具体的，在确定用户手指的被指向位置22后，可以从该被指向位置22开始对当前待点读资料的图像逐行进行识别，从而确定出与该被指向位置22关联的段落起始特征23和段落终止特征24；示例性的，可以在当前待点读资料的图像中识别出与该被指向位置最相近的具有空格两个字的两个特征区域，此时确定的两个特征区域则分别为该被指向位置当前所在段落的段落起始特征23和段落终止特征24。

[0078] 此外，由于用户可以在通过手指在待点读资料上画出对应的点读区域，根据该画出的点读区域判断目标段落的段落起始特征和段落终止特征，因此本实施例中还可以在检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令时，连续获取多帧当前待点读资料的图像，并分析每一帧中用户手指的被指向位置，进而根据连续多帧中的被指向位置，确定用户通过手指画出的点读区域，首先分析该点读区域内是否存在对应的段落起始特征和段落终止特征，若存在则直接将其作为本实施例中的段落起始特征和段落终止特征，若不存在则在该点读区域外识别与该点读区域距离最近的段落起始特征和段落终止特征，作为本实施例中的段落起始特征和段落终止特征，后续对该段落起始特征和段落终止特征之间的文字信息进行识别，从而实现相应的点读功能。

[0079] S240，根据段落起始特征和段落终止特征确定对应的目标段落。

[0080] 可选的，在得到当前待点读资料的图像中用户手指指向的段落起始特征23和段落终止特征24时，直接将当前待点读资料的图像中位于段落起始特征23和段落终止特征24之间的文字内容作为目标段落25，后续对该目标段落25内包含的内容进行识别。

[0081] S250，采用光学字符识别算法识别目标段落的内容。

[0082] 具体的，在确定当前待点读资料的图像中的目标段落时，为了实现相应的点读功能，可以通过字符识别技术来识别目标段落内包含的内容；示例性的，本实施例中采用光学字符识别(Optical Character Recognition，OCR)算法识别该目标段落的内容。其中，OCR算法是指电子设备(例如扫描仪或数码相机)检查纸上打印的或者屏幕显示的字符，通过检测字符暗、亮的模式确定字符形状，然后用字符识别方法将形状翻译成计算机文字的过程，以对文字进行自动识别；本实施例中通过OCR算法检测目标段落内包含的各个字符的形状，从而确定该目标段落内的内容信息。

[0083] S260，根据点读语音指令，展示识别的目标段落的点读内容。

[0084] 本公开实施例提供的技术方案，通过设定的指向规则确定当前待点读资料的图像中被指向的段落起止特征，并识别该段落起止特征确定的目标段落内的内容，进而根据点读语音指令，展示该目标段落的点读内容，提高了段落点读的便利性和实时性，通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，降低了段落点读的局限性，提高了点读内容识别的准确性，增强了用户的使用体验。

[0085] 图3示出了本公开实施例提供的另一种段落点读过程中确定被指向的段落起止特征的原理示意图，本实施例是在上述提供的各个可选方案的基础上进行优化。具体的，本实施例对于通过模型训练的方法判断被指向的段落起止特征的具体确定过程进行详细的介

绍。

[0086] 可选的，该方法具体可以包括如下步骤：

[0087] S310，如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则将当前待点读资料的图像输入预先构建的神经网络模型中，得到该操作对应的被指向位置。

[0088] 可选的，本实施例中的神经网络模型是通过大量历史图像预先训练生成的，该历史图像中包含有用户指向待点读资料的操作信息，并在历史图像中对用户手指实际指向的位置进行标记，进而训练出一种能够准确识别出任一携带有用户指向待点读资料的操作信息的图像中包含的用户手指的被指向位置的神经网络模型。本实施例在检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令时，首先获取当前待点读资料的图像，并直接将该当前待点读资料的图像输入到该预先构建的神经网络模型中，该神经网络模型通过相应训练好的参数对该图像中的用户指向当前待点读资料的操作信息进行相应的分析，从而得到该操作对应的被指向位置31。

[0089] S320，确定与被指向位置关联的段落起始特征和段落终止特征。

[0090] 具体的，在确定用户手指的被指向位置31后，可以从该被指向位置31开始对当前待点读资料的图像逐行进行识别，从而确定出与该被指向位置31关联的段落起始特征32和段落终止特征33；示例性的，可以在当前待点读资料的图像中识别出与该被指向位置最相近的具有空格两个字的两个特征区域，此时确定的两个特征区域则分别为该被指向位置当前所在段落的段落起始特征32和段落终止特征33。

[0091] 此外，由于用户可以在通过手指在待点读资料上画出对应的点读区域，根据该画出的点读区域判断目标段落的段落起始特征和段落终止特征，因此本实施例中还可以在检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令时，连续获取多帧当前待点读资料的图像，并分析每一帧中用户手指的被指向位置，进而根据连续多帧中的被指向位置，确定用户通过手指画出的点读区域，首先分析该点读区域内是否存在对应的段落起始特征和段落终止特征，若存在则直接将其作为本实施例中的段落起始特征和段落终止特征，若不存在则在该点读区域外识别与该点读区域距离最近的段落起始特征和段落终止特征，作为本实施例中的段落起始特征和段落终止特征，后续对该段落起始特征和段落终止特征之间的文字信息进行识别，从而实现相应的点读功能。

[0092] S330，根据段落起始特征和段落终止特征确定对应的目标段落。

[0093] 可选的，在得到当前待点读资料的图像中用户手指指向的段落起始特征32和段落终止特征33时，直接将当前待点读资料的图像中位于段落起始特征32和段落终止特征33之间的文字内容作为目标段落34，后续对该目标段落34内包含的内容进行识别。

[0094] S340，采用光学字符识别算法识别目标段落的内容。

[0095] S350，根据点读语音指令，展示识别的目标段落的点读内容。

[0096] 本公开实施例提供的技术方案，通过预先训练好的神经网络模型确定当前待点读资料的图像中被指向的段落起止特征，并识别该段落起止特征确定的目标段落内的内容，进而根据点读语音指令，展示该目标段落的点读内容，提高了段落点读的便利性和实时性，通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，降低了段落点读的局限性，提高了点读内容识别的准确性，增强了用户的使用体验。

[0097] 图4示出了本公开实施例提供的一种段落点读装置的结构示意图，本公开实施例可适用于对书本上的资料内容进行点读的过程中，该装置可以通过软件和/或硬件来实现，并集成在执行本方法的设备中。如图4所示，本公开实施例中的段落点读装置，可以包括：

[0098] 段落特征确定模块410，用于如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据该操作和点读语音指令确定当前待点读资料的图像中被指向的段落起止特征；

[0099] 段落识别模块420，用于根据段落起止特征识别目标段落的内容；

[0100] 段落点读模块430，用于根据点读语音指令，展示识别的目标段落的点读内容。

[0101] 本公开实施例提供的技术方案，在检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令时，确定当前待点读资料的图像中被指向的段落起止特征，并识别该段落起止特征确定的目标段落内的内容，进而根据点读语音指令，展示该目标段落的点读内容，提高了段落点读的便利性和实时性，通过用户手指的直接操作实现点读，而无需通过专门配置的点读笔辅助实现对应的点读功能，降低了段落点读的局限性，提高了点读内容识别的准确性，增强了用户的使用体验。

[0102] 进一步的，上述段落特征确定模块410，可以包括：

[0103] 特征点获取单元，用于根据该操作获取当前待点读资料的图像中用户手指的关键特征点；

[0104] 第一指向确定单元，用于根据关键特征点确定手指对应的被指向位置；

[0105] 第一特征确定单元，用于确定与被指向位置关联的段落起始特征和段落终止特征。

[0106] 进一步的，上述段落特征确定模块410，还可以包括：

[0107] 第二指向确定单元，用于将当前待点读资料的图像输入预先构建的神经网络模型中，得到该操作对应的被指向位置；

[0108] 第二特征确定单元，用于确定与被指向位置关联的段落起始特征和段落终止特征。

[0109] 进一步的，上述段落识别模块420，可以包括：

[0110] 目标段落确定单元，用于根据段落起始特征和段落终止特征确定对应的目标段落；

[0111] 目标段落识别单元，用于采用光学字符识别算法识别目标段落的内容。

[0112] 进一步的，上述段落点读模块430，可以包括：

[0113] 关键词确定单元，用于采用自然语言处理算法识别点读语音指令，得到指令关键词；

[0114] 点读模式确定单元，用于根据指令关键词确定识别的目标段落内容的点读结果和对应的点读方式；

[0115] 段落点读单元，用于点读根据对应的点读方式展示点读结果。

[0116] 本公开实施例提供的段落点读装置，与上述实施例提供的段落点读方法属于同一发明构思，未在本公开实施例中详尽描述的技术细节可参见上述实施例，并且本公开实施例与上述实施例具有相同的有益效果。

[0117] 下面参考图5，其示出了适于用来实现本公开实施例的设备500的结构示意图。本

公开实施例中的设备可以包括但不限于诸如移动电话、笔记本电脑、数字广播接收器、PDA(个人数字助理)、PAD(平板电脑)、PMP(便携式多媒体播放器)、车载终端(例如车载导航终端)等等的移动终端以及诸如数字TV、台式计算机等等的固定终端。图5示出的设备仅仅是一个示例,不应对本公开实施例的功能和使用范围带来任何限制。

[0118] 如图5所示,设备500可以包括处理装置(例如中央处理器、图形处理器等)501,其可以根据存储在只读存储器(ROM)502中的程序或者从存储装置508加载到随机访问存储器(RAM)503中的程序而执行各种适当的动作和处理。在RAM 503中,还存储有设备500操作所需的各种程序和数据。处理装置501、ROM 502以及RAM 503通过总线504彼此相连。输入/输出(I/O)接口505也连接至总线504。

[0119] 通常,以下装置可以连接至I/O接口505:包括例如触摸屏、触摸板、键盘、鼠标、摄像头、麦克风、加速度计、陀螺仪等的输入装置506;包括例如液晶显示器(LCD)、扬声器、振动器等的输出装置507;包括例如磁带、硬盘等的存储装置508;以及通信装置509。通信装置509可以允许设备500与其他设备进行无线或有线通信以交换数据。虽然图5示出了具有各种装置的设备500,但是应理解的是,并不要求实施或具备所有示出的装置。可以替代地实施或具备更多或更少的装置。

[0120] 特别地,根据本公开的实施例,上文参考流程图描述的过程可以被实现为计算机软件程序。例如,本公开的实施例包括一种计算机程序产品,其包括承载在计算机可读介质上的计算机程序,该计算机程序包含用于执行流程图所示的方法的程序代码。在这样的实施例中,该计算机程序可以通过通信装置509从网络上被下载和安装,或者从存储装置508被安装,或者从ROM 502被安装。在该计算机程序被处理装置501执行时,执行本公开实施例的方法中限定的上述功能。

[0121] 需要说明的是,本公开上述的计算机可读介质可以是计算机可读信号介质或者计算机可读存储介质或者是上述两者的任意组合。计算机可读存储介质例如可以是——但不限于——电、磁、光、电磁、红外线、或半导体的系统、装置或器件,或者任意以上的组合。计算机可读存储介质的更具体的例子可以包括但不限于:具有一个或多个导线的电连接、便携式计算机磁盘、硬盘、随机访问存储器(RAM)、只读存储器(ROM)、可擦式可编程只读存储器(EPROM或闪存)、光纤、便携式紧凑磁盘只读存储器(CD-ROM)、光存储器件、磁存储器件、或者上述的任意合适的组合。在本公开中,计算机可读存储介质可以是任何包含或存储程序的有形介质,该程序可以被指令执行系统、装置或者器件使用或者与其结合使用。而在本公开中,计算机可读信号介质可以包括在基带中或者作为载波一部分传播的数据信号,其中承载了计算机可读的程序代码。这种传播的数据信号可以采用多种形式,包括但不限于电磁信号、光信号或上述的任意合适的组合。计算机可读信号介质还可以是计算机可读存储介质以外的任何计算机可读介质,该计算机可读信号介质可以发送、传播或者传输用于由指令执行系统、装置或者器件使用或者与其结合使用的程序。计算机可读介质上包含的程序代码可以用任何适当的介质传输,包括但不限于:电线、光缆、RF(射频)等等,或者上述的任意合适的组合。

[0122] 上述计算机可读介质可以是上述设备中所包含的;也可以是单独存在,而未装配入该设备中。

[0123] 上述计算机可读介质承载有一个或者多个程序,当上述一个或者多个程序被该设

备执行时，使得该设备：如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据该操作和点读语音指令确定当前待点读资料的图像中被指向的段落起止特征；根据段落起止特征识别目标段落的内容；根据点读语音指令，展示识别的目标段落的点读内容。

[0124] 可以以一种或多种程序设计语言或其组合来编写用于执行本公开的操作的计算机程序代码，上述程序设计语言包括面向对象的程序设计语言—诸如Java、Smalltalk、C++，还包括常规的过程式程序设计语言—诸如“C”语言或类似的程序设计语言。程序代码可以完全地在用户计算机上执行、部分地在用户计算机上执行、作为一个独立的软件包执行、部分在用户计算机上部分在远程计算机上执行、或者完全在远程计算机或服务器上执行。在涉及远程计算机的情形中，远程计算机可以通过任意种类的网络——包括局域网(LAN)或广域网(WAN)—连接到用户计算机，或者，可以连接到外部计算机(例如利用因特网服务提供商来通过因特网连接)。

[0125] 附图中的流程图和框图，图示了按照本公开各种实施例的系统、方法和计算机程序产品的可能实现的体系架构、功能和操作。在这点上，流程图或框图中的每个方框可以代表一个模块、程序段、或代码的一部分，该模块、程序段、或代码的一部分包含一个或多个用于实现规定的逻辑功能的可执行指令。也应当注意，在有些作为替换的实现中，方框中所标注的功能也可以以不同于附图中所标注的顺序发生。例如，两个接连地表示的方框实际上可以基本并行地执行，它们有时也可以按相反的顺序执行，这依所涉及的功能而定。也要注意的是，框图和/或流程图中的每个方框、以及框图和/或流程图中的方框的组合，可以用执行规定的功能或操作的专用的基于硬件的系统来实现，或者可以用专用硬件与计算机指令的组合来实现。

[0126] 描述于本公开实施例中所涉及到的单元可以通过软件的方式实现，也可以通过硬件的方式来实现。其中，单元的名称在某种情况下并不构成对该单元本身的限定。

[0127] 以上描述仅为本公开的较佳实施例以及对所运用技术原理的说明。本领域技术人员应当理解，本公开中所涉及的公开范围，并不限于上述技术特征的特定组合而成的技术方案，同时也应涵盖在不脱离上述公开构思的情况下，由上述技术特征或其等同特征进行任意组合而形成的其它技术方案。例如上述特征与本公开中公开的(但不限于)具有类似功能的技术特征进行互相替换而形成的技术方案。

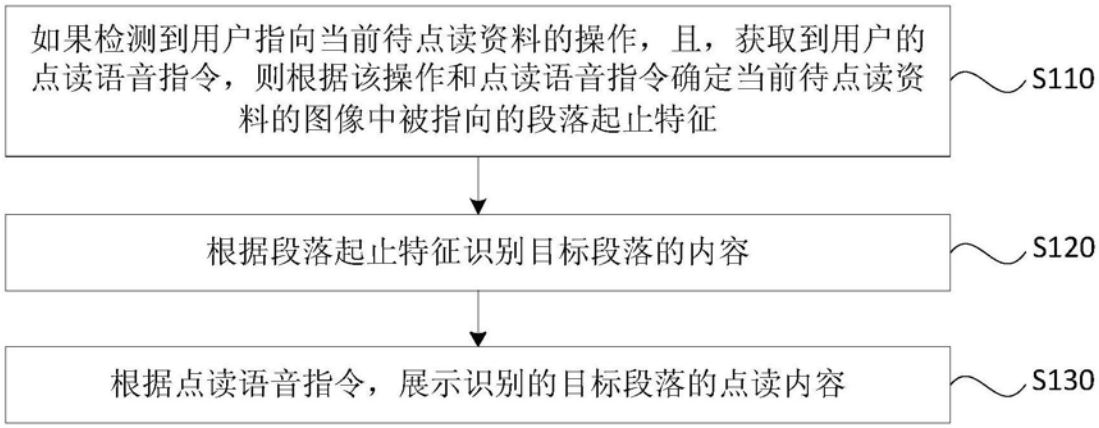

如果检测到用户指向当前待点读资料的操作，且，获取到用户的点读语音指令，则根据该操作和点读语音指令确定当前待点读资料的图像中被指向的段落起止特征
S110
根据段落起止特征识别目标段落的内容
S120
根据点读语音指令，展示识别的目标段落的点读内容
S130

图1A

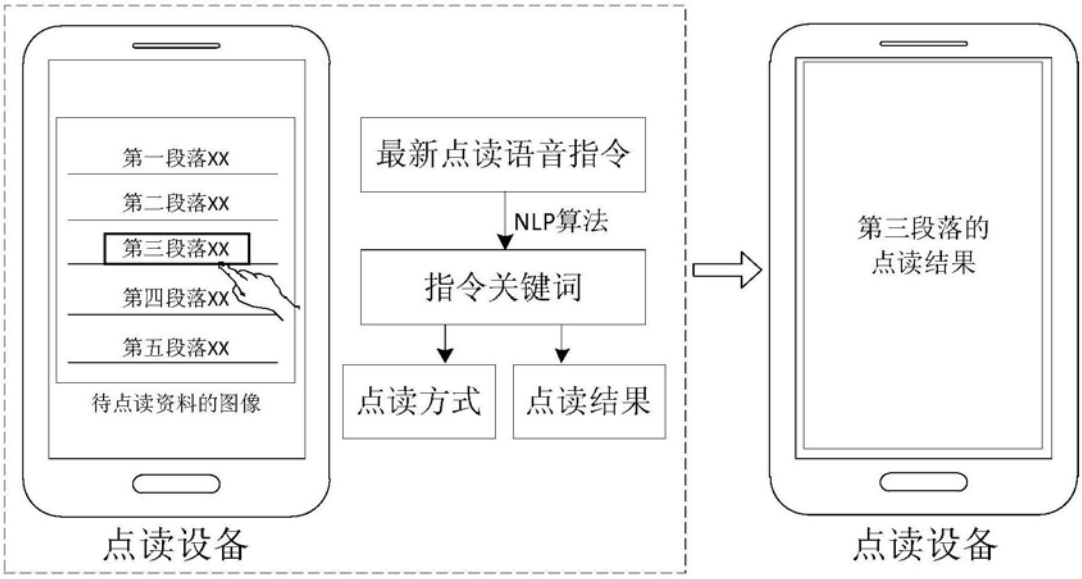

第一段落XX
第二段落XX
第三段落XX
第四段落XX
第五段落XX
待点读资料的图像
点读设备
最新点读语音指令
NLP算法
指令关键词
点读方式
点读结果
第三段落的
点读结果
点读设备

图1B

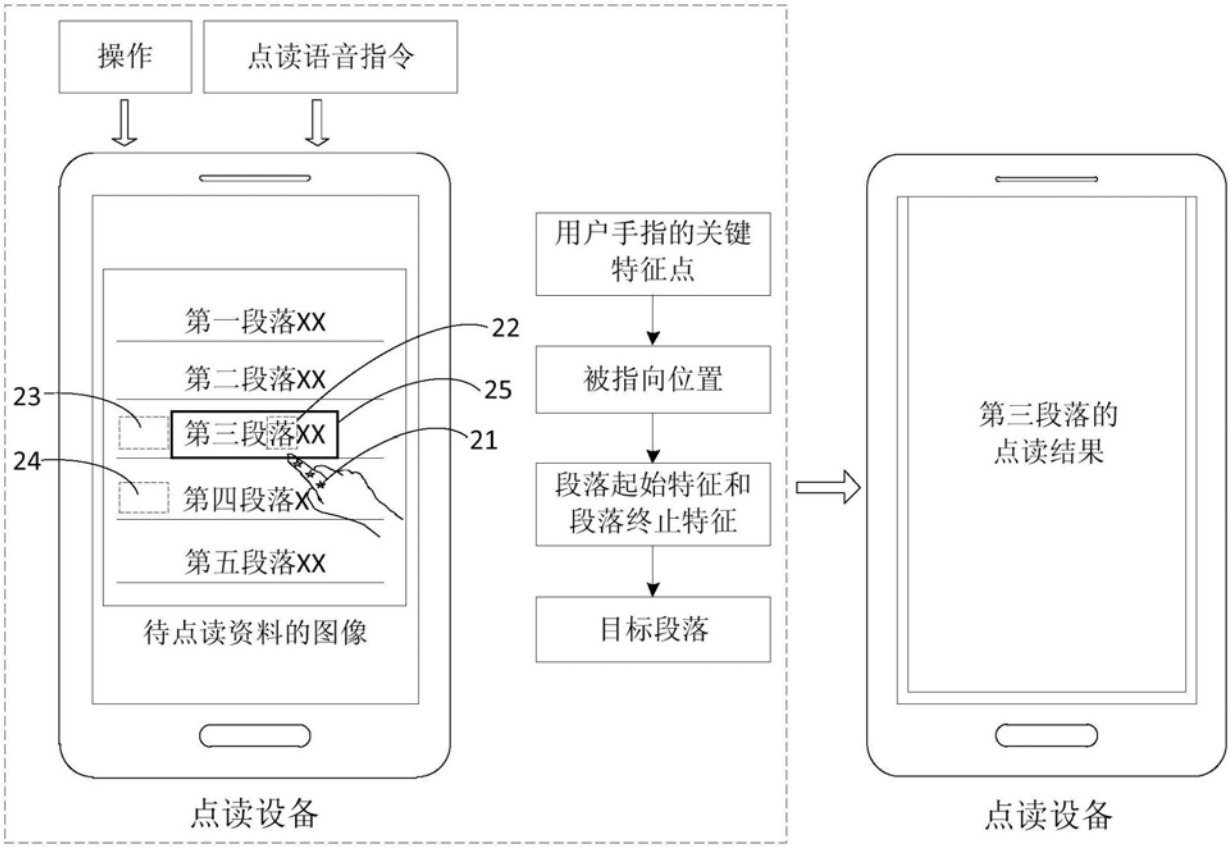

操作
点读语音指令
第一段落XX
第二段落XX
第三段落XX
第四段落XX
第五段落XX
待点读资料的图像
22
25
21
23
24
点读设备
用户手指的关键特征点
被指向位置
段落起始特征和段落终止特征
目标段落
第三段落的点读结果
点读设备

图2

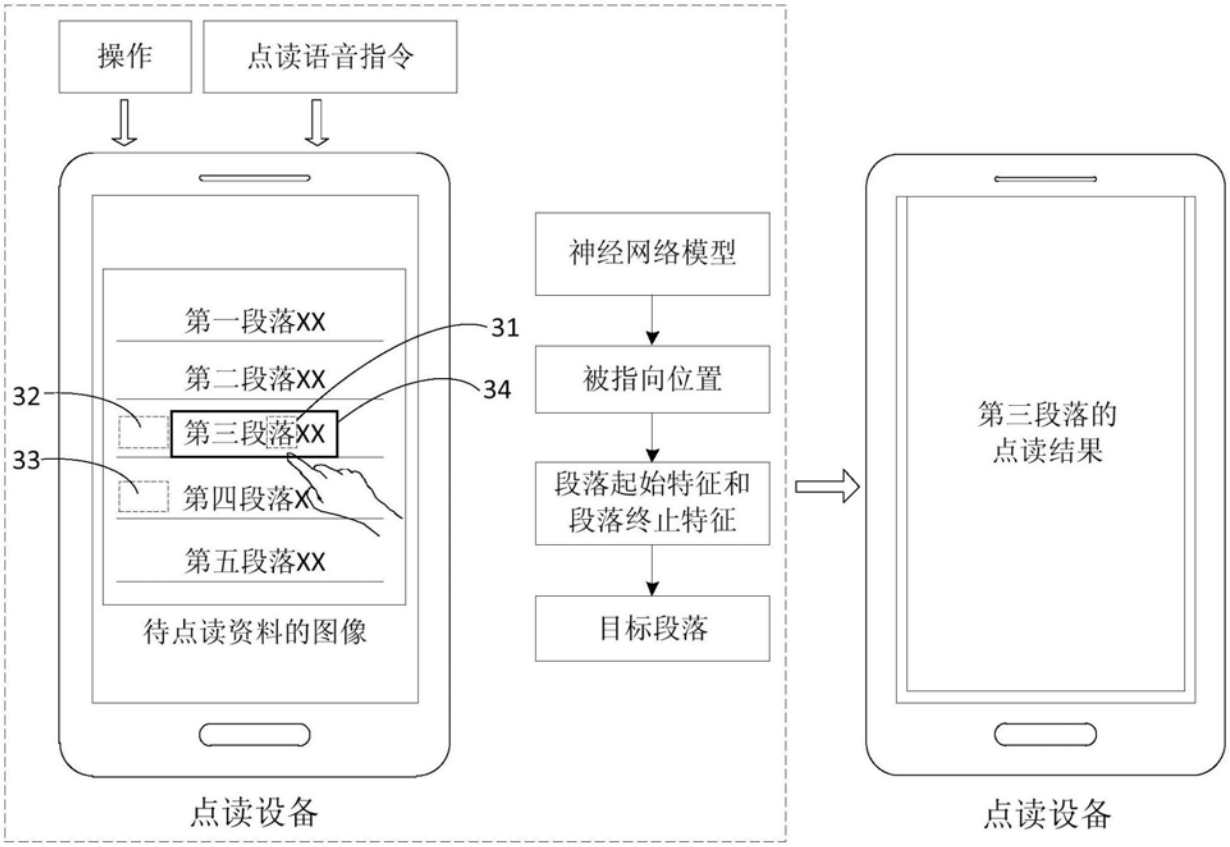

操作
点读语音指令
第一段落XX
第二段落XX
第三段落XX
第四段落XX
第五段落XX
待点读资料的图像
31
32
33
34
神经网络模型
被指向位置
段落起始特征和
段落终止特征
目标段落
第三段落的
点读结果
点读设备
点读设备

图3

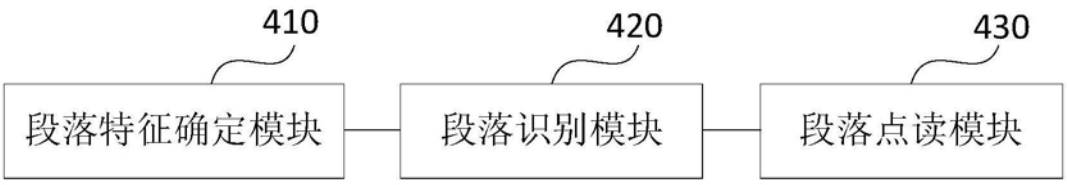

410
420
430
段落特征确定模块
段落识别模块
段落点读模块

图4

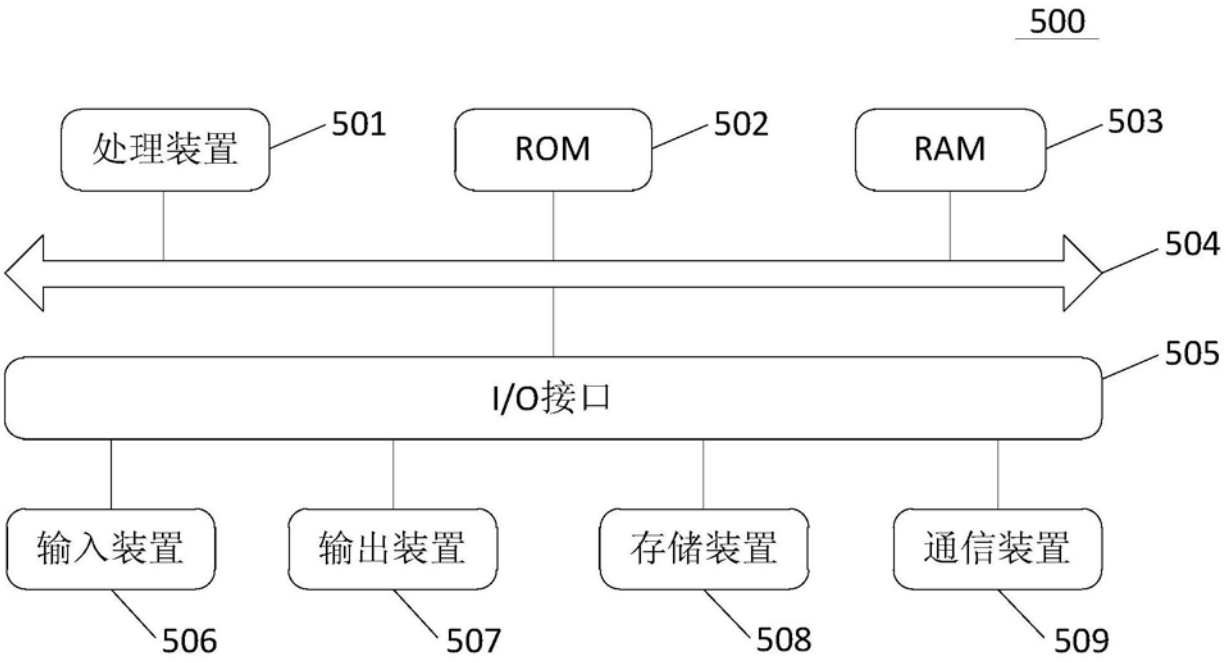

500
处理装置
501
ROM
502
RAM
503
504
I/O接口
505
输入装置
506
输出装置
507
存储装置
508
通信装置
509

图5