



(12) 发明专利

(10) 授权公告号 CN 112364718 B

(45) 授权公告日 2024. 09. 03

(21) 申请号 202011147163.4

G06V 30/18 (2022.01)

(22) 申请日 2020.10.23

G06V 30/19 (2022.01)

(65) 同一申请的已公布的文献号

G06V 30/224 (2022.01)

申请公布号 CN 112364718 A

G06V 30/416 (2022.01)

G09B 5/04 (2006.01)

(43) 申请公布日 2021.02.12

(56) 对比文件

(73) 专利权人 深圳市云希谷科技有限公司

CN 110060524 A, 2019.07.26

地址 518000 广东省深圳市宝安区西乡街

道龙腾社区西乡大道230号满京华艺

峦大厦4座301-F006

审查员 赵乙

(72) 发明人 罗辉

(74) 专利代理机构 深圳市特讯知识产权代理事

务所(普通合伙) 44653

专利代理师 黄彧

(51) Int. Cl.

G06V 30/14 (2022.01)

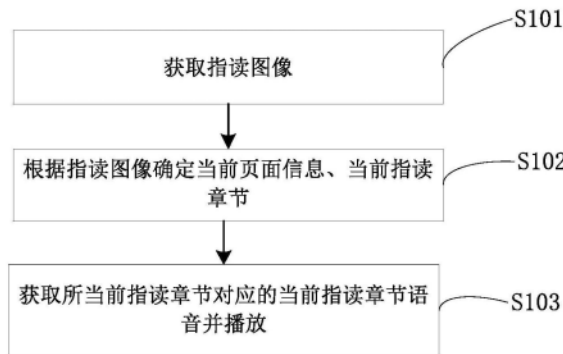
权利要求书2页 说明书13页 附图7页

(54) 发明名称

一种指读方法、装置、机器人、终端及计算机存储介质

(57) 摘要

本发明提供一种指读方法、装置、机器人、终端及计算机存储介质,通过获取指读图像,指读图像包括内容正文、边框花边、章节节点和位置指示工具,根据指读图像确定当前页面信息、当前指读章节,获取当前指读章节对应的当前指读章节语音并播放,本发明还提供一种指读装置、机器人、终端及计算机存储介质,达到了基于现有的部分书籍通过预先处理即可实现指读,不需要重新制作书籍和点读笔,降低了成本,更加便利,提升了用户体验度。



1. 一种指读方法,其特征在于,包括:

获取指读图像,所述指读图像包括内容正文、边框花边、章节节点和位置指示工具;

根据所述指读图像确定当前页面信息、当前指读章节;

获取所述当前指读章节对应的当前指读章节语音并播放;

其中,所述根据所述指读图像确定当前页面信息包括:

对所述指读图像进行掩码处理,分别确定目标边框花边轮廓和目标章节节点轮廓;

根据所述目标边框花边轮廓确定透视后指读图像;

根据所述目标章节节点轮廓确定可疑图片;

将所述透视后指读图像与所述可疑图片进行SIFT特征比对,确定所述当前页面信息;

其中,所述根据所述目标章节节点轮廓确定可疑图片包括,获取各所述目标章节节点轮廓信息,确定章节节点编码;根据所述章节节点编码确定所述可疑图片;所述目标章节节点轮廓信息包括每一页是否存在章节节点,章节节点个数,章节节点的位置信息;

所述根据章节节点编码确定可疑图片包括,根据章节节点编码与预设的基准章节节点编码进行比对,当章节节点编码所对应的章节节点位置信息和章节节点数量信息,为基准章节节点编码所对应的基准章节节点位置信息和基准章节节点数量信息的子集或全集时,将基准章节节点编码对应的页面图像作为可疑图片;

所述根据所述指读图像确定当前指读章节包括,从透视图像掩码集合中确定颜色掩码,并进行图片融合,生成位置指示工具掩码;根据所述位置指示工具掩码确定所述位置指示工具的轮廓,并确定所述位置指示工具的轮廓距离页面顶端最近的点作为指示位置;根据所述指示位置、所述当前页面信息确定当前指读章节。

2. 如权利要求1所述的指读方法,其特征在于,所述对所述指读图像进行掩码处理,确定目标边框花边轮廓包括:

对所述指读图像进行掩码处理,获取多个初始轮廓掩码;

对所述初始轮廓掩码进行逼近处理,确定多个疑似边框花边轮廓;

将多个所述疑似边框花边轮廓进行聚类处理,确定目标边框花边轮廓。

3. 如权利要求2所述的指读方法,其特征在于,所述章节节点轮廓包括圆轮廓,所述对所述指读图像进行掩码处理,确定目标章节节点轮廓包括:

确定所述目标边框花边轮廓的各顶点信息;

根据所述顶点信息对所述指读图像进行透视变换,生成透视后指读图像;

对所述透视后指读图像进行掩码处理,生成透视图像掩码集合,并对所述透视图像掩码集合进行轮廓检测,生成霍夫变换掩码;

对所述霍夫变换掩码进行霍夫圆检测,获取多个疑似章节节点轮廓;

对所述章节节点轮廓进行分类聚类,确定目标章节节点轮廓。

4. 如权利要求3所述的指读方法,其特征在于,所述根据所述目标边框花边轮廓确定透视后指读图像包括:

获取包括所述目标边框花边轮廓的定位范围,根据位于所述定位范围内的所述指读图像生成定位图像;

根据所述顶点信息对所述定位图像进行透视变换,生成透视后指读图像。

5. 如权利要求4所述的指读方法,其特征在于,所述边框花边的轮廓包括矩形,所述确

定所述目标边框花边轮廓的各顶点信息包括：

从所述目标边框花边轮廓的四条轮廓边中确定两条或三条可靠轮廓边；

对所述定位图像进行掩码处理,生成定位图像掩码,并对所述定位图像掩码进行图片融合,确定二值掩码；

对所述二值掩码进行概率霍夫直线检测,生成补全轮廓边；

根据所述可靠轮廓边和所述补全轮廓边确定所述目标边框花边轮廓的各顶点信息。

6. 一种指读装置,其特征在于,用于实现权利要求1所述的指读方法,包括：

图像获取模块,用于获取指读图像,所述指读图像包括内容正文、边框花边、章节节点和位置指示工具；

确定模块,用于根据所述指读图像确定当前页面信息、当前指读章节；

播放模块,用于获取所述当前指读章节对应的当前指读章节语音并播放。

7. 一种指读机器人,其特征在于,包括图像采集设备、发声设备和如权利要求6所述的指读装置,

所述图像采集设备用于采集指读图像,并将所述指读图像发送给所述指读装置；

所述发声设备用于播放所述指读装置反馈的当前指读章节语音。

8. 一种终端,其特征在于,包括处理器、存储器和通信总线；

所述通信总线用于将所述处理器和存储器连接；

所述处理器用于执行所述存储器中存储的计算机程序,以实现如权利要求1-5中一个或多个所述的指读方法。

9. 一种计算机可读存储介质,其特征在于,其上存储有计算机程序,

所述计算机程序用于使所述计算机执行如权利要求1-5中任一项所述的指读方法。

一种指读方法、装置、机器人、终端及计算机存储介质

技术领域

[0001] 本发明涉及通信领域,特别是涉及一种指读方法、装置、机器人、终端及计算机存储介质。

背景技术

[0002] 书籍阅读是人们工作生活中非常重要的一部分,随着技术的发展,人们对于阅读辅助产品的需求日益增长,当前一种常见的阅读辅助技术包括OID点读技术,针对OID点读印刷物,一般配合点读笔使用。在进行点读时,可以通过点读笔,识别隐藏在印刷物中的OID编码,播放与OID编码对应的点读数据。

[0003] 但该OID点读技术需要配套制作其配套的书籍,对于消费者来说成本较高,且必须配合点读笔使用,不够便利,用户体验度不高。

发明内容

[0004] 鉴于以上所述现有技术的缺点,本发明的目的在于提供一种指读方法、装置、机器人、终端及计算机存储介质,用于解决阅读辅助需要配套制作其配套的书籍,对于消费者来说成本较高,且必须配合点读笔使用,不够便利,用户体验度不高的问题。

[0005] 为实现上述目的及其他相关目的,本发明的目的在于提供一种指读方法,包括:

[0006] 获取指读图像,所述指读图像包括内容正文、边框花边、章节节点和位置指示工具;

[0007] 根据所述指读图像确定当前页面信息、当前指读章节;

[0008] 获取所述当前指读章节对应的当前指读章节语音并播放。

[0009] 可选的,所述根据所述指读图像确定当前页面信息包括:

[0010] 对所述指读图像进行掩码处理,分别确定目标边框花边轮廓和目标章节节点轮廓;

[0011] 根据所述目标边框花边轮廓确定透视后指读图像;

[0012] 根据所述目标章节节点轮廓确定可疑图片;

[0013] 将所述透视后指读图像与所述可疑图片进行SIFT特征比对,确定所述当前页面信息。

[0014] 可选的,所述对所述指读图像进行掩码处理,确定目标边框花边轮廓包括:

[0015] 对所述指读图像进行掩码处理,获取多个初始轮廓掩码;

[0016] 对所述初始轮廓掩码进行逼近处理,确定多个疑似边框花边轮廓;

[0017] 将多个所述疑似边框花边轮廓进行聚类处理,确定目标边框花边轮廓。

[0018] 可选的,所述章节节点轮廓包括圆轮廓,所述对所述指读图像进行掩码处理,确定目标章节节点轮廓包括:

[0019] 确定所述目标边框花边轮廓的各顶点信息;

[0020] 根据所述顶点信息对所述指读图像进行透视变换,生成透视后指读图像;

- [0021] 对所述透视后指读图像进行掩码处理,生成透视图像掩码集合,并对所述透视图像掩码集合进行轮廓检测,生成霍夫变换掩码;
- [0022] 对所述霍夫变换掩码进行霍夫圆检测,获取多个疑似章节节点轮廓;
- [0023] 对所述章节节点轮廓进行分类聚类,确定目标章节节点轮廓。
- [0024] 可选的,所述根据所述目标边框花边轮廓确定透视后指读图像包括:
- [0025] 获取包括所述目标边框花边轮廓的定位范围,根据位于所述定位范围之内的所述指读图像生成定位图像;
- [0026] 根据所述顶点信息对所述定位图像进行透视变换,生成透视后指读图像。
- [0027] 可选的,所述边框花边的轮廓包括矩形,所述确定所述目标边框花边轮廓的各顶点信息包括:
- [0028] 从所述目标边框花边轮廓的四条轮廓边中确定两条或三条可靠轮廓边;
- [0029] 对所述定位图像进行掩码处理,生成定位图像掩码,并对所述定位图像掩码进行图片融合,确定二值掩码;
- [0030] 对所述二值掩码进行概率霍夫直线检测,生成补全轮廓边;
- [0031] 根据所述可靠轮廓边和所述补全轮廓边确定所述目标边框花边轮廓的各顶点信息。
- [0032] 可选的,还包括以下至少之一:
- [0033] 所述根据所述目标章节节点轮廓确定可疑图片包括,获取各所述目标章节节点轮廓信息,确定章节节点编码;根据所述章节节点编码确定所述可疑图片;
- [0034] 所述根据所述指读图像确定当前指读章节包括,从所述透视图像掩码集合中确定颜色掩码,并进行图片融合,生成位置指示工具掩码;根据所述位置指示工具掩码确定所述位置指示工具的轮廓,并确定所述位置指示工具的轮廓距离页面顶端最近的点作为指示位置;根据所述指示位置、所述当前页面信息确定当前指读章节。
- [0035] 本发明还提供了一种指读装置,包括:
- [0036] 图像获取模块,用于获取指读图像,所述指读图像包括内容正文、边框花边、章节节点和位置指示工具;
- [0037] 确定模块,用于根据所述指读图像确定当前页面信息、当前指读章节;
- [0038] 播放模块,用于获取所述当前指读章节对应的当前指读章节语音并播放。
- [0039] 本发明还提供了一种指读机器人,包括括图像采集设备、发声设备和如上述任一项实施例所述的指读装置,
- [0040] 所述图像采集设备用于采集指读图像,并将所述指读图像发送给所述指读装置;
- [0041] 所述发声设备用于播放所述指读装置反馈的当前指读章节语音。
- [0042] 本发明还提供了一种终端,包括处理器、存储器和通信总线;
- [0043] 所述通信总线用于将所述处理器和存储器连接;
- [0044] 所述处理器用于执行所述存储器中存储的计算机程序,以实现如上述实施例中的一个或多个所述的指读方法。
- [0045] 本发明还提供了一种计算机可读存储介质,其上存储有计算机程序,
- [0046] 所述计算机程序用于使所述计算机执行如上述实施例中任一项所述的指读方法。
- [0047] 如上所述,本发明提供的一种指读方法、装置、机器人、终端及计算机存储介质,具

有以下有益效果：

[0048] 通过获取指读图像，指读图像包括内容正文、边框花边、章节节点和位置指示工具，根据指读图像确定当前页面信息、当前指读章节，获取当前指读章节对应的当前指读章节语音并播放，基于现有的部分书籍通过预先处理即可实现指读，不需要重新制作书籍和点读笔，降低了成本，更加便利，提升了用户体验度。

附图说明

- [0049] 图1为本发明实施例一提供的指读方法的一种流程示意图；
- [0050] 图2为本发明实施例一提供的一种指读图像示意图；
- [0051] 图3为本发明实施例一提供的另一种指读图像示意图；
- [0052] 图4为本发明实施例一提供的另一种指读图像示意图；
- [0053] 图5为本发明实施例一提供的另一种正文页面示意图；
- [0054] 图6为本发明实施例一提供的另一种指读图像示意图；
- [0055] 图7为本发明实施例一提供的一种图书指读方法流程示意图；
- [0056] 图8为本发明实施例一提供的一种具体的指读方法流程示意图；
- [0057] 图9为本发明实施例二提供的指读装置的一种结构示意图；
- [0058] 图10为本发明实施例二提供的终端的一种结构示意图。

具体实施方式

[0059] 以下通过特定的具体实例说明本发明的实施方式，本领域技术人员可由本说明书所揭露的内容轻易地了解本发明的其他优点与功效。本发明还可以通过另外不同的具体实施方式加以实施或应用，本说明书中的各项细节也可以基于不同观点与应用，在没有背离本发明的精神下进行各种修饰或改变。需说明的是，在不冲突的情况下，以下实施例及实施例中的特征可以相互组合。

[0060] 需要说明的是，以下实施例中所提供的图示仅以示意方式说明本发明的基本构想，遂图式中仅显示与本发明中有关的组件而非按照实际实施时的组件数目、形状及尺寸绘制，其实际实施时各组件的型态、数量及比例可为一种随意的改变，且其组件布局型态也可能更为复杂。

[0061] 实施例一

[0062] 请参阅图1，本发明实施例提供了一种指读方法，包括：

[0063] S101：获取指读图像。

[0064] 在一些实施例中，指读图像包括内容正文、边框花边、章节节点和位置指示工具。

[0065] 其中，位置指示工具包括不限于手指、笔、尺子、等可以用来指示内容正文的某一位置的工具。

[0066] 在一些实施例中，章节节点包括区别于文字、标点符号的具有一定形状或花纹的图案。

[0067] 在一些实施例中，本发明实施例提供的指读方法应用于页面包括如图2所示的正文四周带有边框花纹4的书籍或文件等的指读，且该书籍或文件的正文部分穿插有如图2所示的明显标识的章节节点5，需要说明的是，图2边框花纹以及章节节点仅是一种示例，本实

施例提供的指读方法还可以适用于其他形状的边框花纹、章节节点的书籍或文件。

[0068] 在一些实施例中,本发明实施例应用于页面正文具有边框,正文具有章节节点的图书、期刊、杂志等的指读,以图书为例,通常对于不同出版机构所出版同一内容的图书的排版略有不同,若这些图书的正文页面均有明显的边框花纹,参见图2、图3、图4和图5所示的正文页面示例,正文内容中也包括有章节节点。因此,可以通过预先处理并存储现有的该图书的印刷本信息,基于已出版的图书所自带的边框花纹和章节节点实现指读。

[0069] 在一些实施例中,参见图2,指读图像1包括边框花边4、章节节点5、内容正文3、以及位置指示工具2。

[0070] 在一些实施例中,位置指示工具如图2所示可以是手指,也可以是笔等其他工具,在此不做限定。

[0071] 在一些实施例中,指读图像可以由其他设备采集后发送给当前执行指读方法的设备,也可以是由当前执行指读方法的设备自身直接所采集的,在此不做限定。

[0072] 在一些实施例中,指读图像包括页面的正面图像。需要说明的是,指读图像也可以是由与页面具有一定角度的图像采集设备所采集的,此时,由于视角的缘故,指读图像具有一定的“扭曲”,如图3所示,此时,在不影响对于边框花纹4、章节节点5的识别的前提下,此类指读图像也可以适用于本实施例所提供的指读方法。

[0073] 在一些实施例中,获取指读图像之前,该指读方法还包括:

[0074] 获取触发动作,抓拍指读图像。

[0075] 可选的,触发动作包括但不限于以下至少之一:执行工具停留同一位置超过预设时间、语音指令、双击页面、点头、摇头等。

[0076] 在一些实施例中,该指读方法还包括:

[0077] 获取当前指读图书或文件的身份识别标识,后续根据该身份识别标识以及当前指读章节确定当前指读章节语音。

[0078] 在一些实施例中,获取指读图像之前,该指读方法还包括:

[0079] 获取指读对象的每一页正文页面的基准图像;

[0080] 获取各标准图像中的基准章节节点信息;

[0081] 根据基准章节节点信息确定基准章节节点编码。

[0082] 可选的,预先存储指读对象的每一页正文与基准章节节点编码,以供后续确定可疑页面。

[0083] 可选的,指读对象即为应用该指读方法的内容正文包括有边框花纹、章节节点的书籍或文件等对象。

[0084] 可选的,以指读对象为X图书为例,对整本X图书进基准图片信息获取,即对所有的内页图片进行正面俯拍;获取每一页的章节节点信息,章节节点信息包括但不限于每一页是否存在章节节点,章节节点个数,章节节点的位置信息,并根据上述章节节点信息确定章节节点编码,以便于后续的内容识别,并通过数据库存储;将该X图书的每页每章节的语音信息按页数和章节数通过数据库存储。由于针对于同一X图书各版本可能存在排版的差异,但其各章节的内容均一致,因此语音信息可以用同一套语音信息,当需要适用于新的排版的X图书时,仅需要对新的X图书的各页面的基准章节节点编码进行重新获取,并重新关联页数、章节数即可。

[0085] 可选的,可以预先存储若干本常见的指读对象的基准章节节点编码、基准图像以及章节语音存入基础数据库中,并对每一本设定对应的身份识别信息,当用户更换了新的指读对象时,基于身份识别信息选择对应服务的基础数据库的数据。

[0086] 在一些实施例中,可以设定每隔预设时间自动获取指读图像,其中预设时间可以根据本领域技术人员的需要进行设定,也可以由用户自行设定,在此不做限定。

[0087] 在一些实施例中,指读图像为图像采集设备采集到的预设采集区域中的图像,该指读图像可能还会包括如图3、图4、图5所示的外围干扰图像,如桌面的图像等。

[0088] S102:根据指读图像确定当前页面信息、当前指读章节。

[0089] 可选的,当前页面信息包括页码,当前指读章节包括章节数。

[0090] 在一些实施例中,根据指读图像确定当前页面信息包括:

[0091] 对指读图像进行掩码处理,分别确定目标边框花边轮廓和目标章节节点轮廓;

[0092] 根据目标边框花边轮廓确定透视后指读图像;

[0093] 根据目标章节节点轮廓确定可疑图片;

[0094] 将透视后指读图像与可疑图片进行SIFT特征比对,确定当前页面信息。

[0095] 在一些实施例中,对指读图像进行掩码处理,确定目标边框花边轮廓包括:

[0096] 对指读图像进行掩码处理,获取多个初始轮廓掩码;

[0097] 对初始轮廓掩码进行逼近处理,确定多个疑似边框花边轮廓;

[0098] 将多个疑似边框花边轮廓进行聚类处理,确定目标边框花边轮廓。

[0099] 可选的,对指读图像进行掩码处理包括但不限于对指读图像指读图像进行Canny、Sobel、Laplace中任意一种处理。

[0100] 可选的,对指读图像的初始轮廓掩码进行逼近处理之前,还包括,对各初始轮廓掩码分别进行轮廓检测,并根据第一预设条件进行筛选,将筛选后的初始轮廓集合进行加和,再对筛选后的初始轮廓集合进行逼近处理。其中,第一预设条件包括第一预设面积等,该第一预设面积与页面的边框花边面积所包括的面积大小接近。这样可以筛选掉噪点或文字内容轮廓以免对后续处理造成影响。

[0101] 可选的,对初始轮廓掩码进行逼近处理之后,确定多个疑似边框花边轮廓之前,还包括,根据第二预设条件对逼近处理之后得到的粗边框花边轮廓进行筛选,通过筛选的粗边框花边轮廓为疑似边框花边轮廓。其中第二预设条件包括第二预设面积、顶点数量、长宽比等。这样可以筛选出较为精确的页面边框花边轮廓。参见图6中,示例展示了一种疑似边框花边轮廓6。

[0102] 可选的,将多个疑似边框花边轮廓进行聚类处理,确定目标边框花边轮廓包括,对各疑似边框花边轮廓进行分类聚类,以将疑似边框花边轮廓中单一页的网页轮廓分类,再对分类聚类后的每一类轮廓进行最优解的聚类处理,以获取目标边框花边轮廓,该目标边框花边轮廓包括较为精确的网页轮廓信息。

[0103] 可选的,边框花边的轮廓包括矩形,确定目标边框花边轮廓的各顶点信息包括:

[0104] 从目标边框花边轮廓的四条轮廓边中确定两条或三条可靠轮廓边;

[0105] 对定位图像进行掩码处理,生成定位图像掩码,并对定位图像掩码进行图片融合,确定二值掩码;

[0106] 对二值掩码进行概率霍夫直线检测,生成补全轮廓边;

- [0107] 根据可靠轮廓边和补全轮廓边确定目标边框花边轮廓的各顶点信息。
- [0108] 可选的,继续参见图6,若疑似边框花边轮廓6就是目标边框花边轮廓,定位范围可以通过对目标边框花边轮廓进行外接矩形7提取,并对外接矩形7在纵横方向进行一定比例的扩张,并根据扩张后的外扩外接矩形8所在的定位范围提取指读图像中对应位置的图像以作为定位图像,该定位图像用于后续确定当前页面信息。
- [0109] 可选的,可以通过确定目标边框花边轮廓在定位图像中的4个顶点的位置信息是否存在临界外扩外接矩形的边界的顶点来确定定位图像所包括的页面图像是否完整,具体的,继续参见图6,4个顶点9与外扩外接矩形均存在一定距离,此时,定位图像中的页面图像时完整的。参见图4,定位图像中左下的顶点与外扩外接矩形8存在重合,此时可以说明定位图像中存在页面图像缺失。
- [0110] 可选的,当判断出定位图像中存在页面图像缺失,从目标边框花边轮廓的四条轮廓边中确定两条或三条可靠轮廓边。其中,可以通过定位图像中目标边框花边轮廓各边的斜率、偏移情况、四个顶点的方向信息、各顶点与外扩外接矩形的位置关系等确定定位图像中可能缺失的顶点信息,以及四条边中较为可信的两条或三条可靠轮廓边。
- [0111] 可选的,对定位图像进行掩码处理的方式包括但不限于以下任意之一:Canny处理、Sobel 处理、自定义二值处理等。
- [0112] 可选的,对定位图像掩码进行图片融合可以采用包括但不限于Canny7:0.5, Canny3:0.5等方式。
- [0113] 可选的,二值掩码用于直线检测。
- [0114] 可选的,对二值掩码进行概率霍夫直线检测,生成补全轮廓边包括:将二值掩码进行概率霍夫直线检测,并将检测到的直线设置条件进行第三预设条件筛选,并根据之前的可靠轮廓边数目确定一条或两条的补全轮廓边。其中,第三预设条件包括但不限于直线所在的范围是否在正确的范围内,直线与可靠轮廓边的交点的位置信息等。
- [0115] 在一些实施例中,章节节点轮廓包括圆轮廓,对指读图像进行掩码处理,确定目标章节节点轮廓包括:
- [0116] 确定目标边框花边轮廓的各顶点信息;
- [0117] 根据顶点信息对指读图像进行透视变换,生成透视后指读图像;
- [0118] 对透视后指读图像进行掩码处理,生成透视图像掩码集合,并对透视图像掩码集合进行轮廓检测,生成霍夫变换掩码;
- [0119] 对霍夫变换掩码进行霍夫圆检测,获取多个疑似章节节点轮廓;
- [0120] 对章节节点轮廓进行分类聚类,确定目标章节节点轮廓。
- [0121] 可选的,根据目标边框花边轮廓确定透视后指读图像包括:
- [0122] 获取包括目标边框花边轮廓的定位范围,根据位于定位范围之内的指读图像生成定位图像;
- [0123] 根据顶点信息对定位图像进行透视变换,生成透视后指读图像。
- [0124] 可选的,对透视后指读图像进行掩码处理包括但不限于以下任意之一:Canny处理、Sobel 处理、自定义二值处理等。
- [0125] 可选的,对透视图像掩码集合进行轮廓检测之后,生成霍夫变换掩码之前还包括:对轮廓检测的结果进行第四预设条件筛选,根据第四筛选的结果生成霍夫变换掩码。其中,

第四预设条件包括但不限于面积大小,外接矩形长宽比例,位置是否在识别书本内容范围内等。

[0126] 可选的,对霍夫变换掩码进行霍夫圆检测之后,获取多个疑似章节节点轮廓之前,还包括:对进行霍夫圆检测的结果进行第五预设条件筛选,第五筛选之后确定疑似章节节点轮廓。其中第五预设条件包括但不限于圆心位置,圆半径等。

[0127] 可选的,若章节节点为其他图形,如五边形等,霍夫变换掩码进行与章节节点形状相适应的形状检测,以获取疑似章节节点轮廓。

[0128] 可选的,根据目标章节节点轮廓确定可疑图片包括:

[0129] 获取各目标章节节点轮廓信息,确定章节节点编码;

[0130] 根据章节节点编码确定可疑图片。

[0131] 可选的,目标章节节点轮廓信息包括但不限于每一页是否存在章节节点,章节节点个数,章节节点的位置信息等。

[0132] 可选的,根据章节节点编码确定可疑图片包括,根据章节节点编码与预设的基准章节节点编码进行比对,当章节节点编码所对应的章节节点位置信息和章节节点数量信息,为基准章节节点编码所对应的基准章节节点位置信息和基准章节节点数量信息的子集或全集时,将基准章节节点编码对应的页面图像作为可疑图片。需要说明的是,这样可以不用考虑当指读图像中的页面部分缺失导致的部分章节节点信息缺失所带来的问题。

[0133] 可选的,根据指读图像确定当前指读章节包括:

[0134] 从透视图像掩码集合中确定颜色掩码,并进行图片融合,生成位置指示工具掩码;

[0135] 根据位置指示工具掩码确定位置指示工具的轮廓,并确定位置指示工具的轮廓距离页面顶端最近的点作为指示位置;

[0136] 根据指示位置、当前页面信息确定当前指读章节。

[0137] 当指示工具包括手指时,进行肤色掩码获取,从透视图像掩码集合中提取HSV和YCbCr 颜色掩码进行图片融合,生成位置指示工具掩码;对生成位置指示工具掩码中进行轮廓检测,并通过第六预设条件进行筛选,确定人手轮廓,并提取轮廓纵轴上最大值点作为指尖位置。其中,第六预设条件包括但不限于面积大小,轮廓所处的区域等。其中,面积大小与人手大小相近。

[0138] S103:获取所当前指读章节对应的当前指读章节语音并播放。

[0139] 可选的,当指示工具包括手指时,根据指尖位置和页面信息确定当前指读章节信息,并根据页数和章节数在音频数据库中进行提取音频信息;将音频信息通过设备进行播放,完成整个指读操作。

[0140] 在一些实施例中,获取所当前指读章节对应的当前指读章节语音并播放包括:

[0141] 获取所当前指读章节对应的当前指读章节语音;

[0142] 获取播放信号,播放当前指读章节语音。

[0143] 在一些实施例中,可以设置每隔预设时间获取一次指读图像,并根据该指读图像确定当前指读章节,但只有获取到播放信号后再行播放当前指读章节语音。其中播放信号包括但不限于语音控制信号、机械控制信号等。

[0144] 在一些实施例中,当前指读章节语音存储在云端,播放当前指读章节语音可以通过联网的播放装置进行播放,如接入网络的音箱。

[0145] 在一些实施例中,也可以通过蓝牙连接扬声器,通过扬声器播放当前指读章节语音。

[0146] 在一些实施例中,根据顶点信息对指读图像进行透视变换,生成透视后指读图像后,还包括:获取历史透视后指读图像,该历史透视后指读图像为根据上一次获取的指读图像生成的透视后指读图像;若历史透视后指读图像与透视后指读图像相同,则获取上一次执行指读方法所确定的当前页面信息作为本次获取的指读图像对应的当前页面信息。可以直接根据本次获取的指读图像确定当前指读章节,进而根据当前页面信息、当前指读章节确定当前指读章节语音。

[0147] 在一些实施例中,指读方法可以在本地执行,也可以在云端执行,在此不做限定。

[0148] 可选的,本实施例的指读方法,可以将预先采集的指读对象的各页面图像、以及各页面图像对应的基准章节节点编码、各章节语音均存储在云端服务器,通过本地图像采集设备进行指读图像采集后,发送到云端服务器,由云端服务器执行指读方法,并将确定的当前指读章节语音通过网络传输的方式通过本地的扬声器进行本地播放。

[0149] 可选的,可以将预先采集的指读对象的各页面图像、以及各页面图像对应的基准章节节点编码、各章节语音均存储在本地,通过本地图像采集设备进行指读图像采集后,由本地执行指读方法,并播放确定的当前指读章节语音。

[0150] 可选的,指读对象的各页面图像、以及各页面图像对应的基准章节节点编码、各章节语音也可以一部分存储在云端服务器,一部分存储在本地,在此不做限定。

[0151] 下面通过一个具体的实施例对本实施例提供的指读方法进行进一步示例性说明,参见图 7,一种图书指读方法包括:

[0152] S701:图书预处理。

[0153] 可选的,预处理方式如下:

[0154] 获取整本图书的基准图像,也即对图书所有的内页进行正面俯拍,得到内页的页面图像;

[0155] 获取每一页的基准章节节点信息,基准章节节点信息包括每一页是否存在章节节点,章节节点个数,章节节点的位置信息,并根据基准章节节点信息生成基准章节节点编码;

[0156] 将上述基准图像、基准章节节点编码存储于数据库;

[0157] 将图书的每页每章节的语音信息按页数和章节数通过数据库存储。

[0158] 其中,基准章节节点编码用于后续在当前页面信息确定。

[0159] S702:获取指读图像,并确定页面信息。

[0160] 可选的,通过设置一个频率(如0.5s),每隔一定的频率进行预设区域范围指读图像的获取。

[0161] 可选的,确定页面信息的方式如下:

[0162] 对获取的预设区域范围的指读图像进行掩码处理,其中处理方式包括但不限于 Canny、Sobel、Laplace等;

[0163] 对处理后的掩码集合中的指读图像的初始轮廓掩码分别进行轮廓检测,并设置第一预设条件(面积大小)进行筛选,以筛选掉噪点或文字内容轮廓以免对后续处理造成影响,再将筛选后的初始轮廓集合进行加和;

- [0164] 对筛选后的初始轮廓集合进行逼近处理,并再次设置第二预设条件(面积大小,顶点数,长宽比)进行筛选,筛选出较为精确的疑似边框花边轮廓;
- [0165] 对再次筛选后的疑似边框花边轮廓集合进行分类聚类操作,将疑似边框花边轮廓中同一页的单页轮廓分类;
- [0166] 对进行分类聚类后的每一类轮廓进行最优解的聚类操作,以获取精确的目标边框花边轮廓;
- [0167] 根据上面聚类后提取到的目标边框花边轮廓,进行外接矩形的提取,并对外接矩形在纵横方向进行一定比例的扩张,并根据扩张后的扩张外接矩形信息提取指读图像中对应位置的图像(定位图像)以作为初步的页面信息定位图片范围,再确定目标边框花边轮廓信息在扩张外接矩形中的相对位置坐标信息,也即,确定目标边框花边轮廓信息在定位图像中的相对位置坐标信息;
- [0168] 将定位图像中的目标边框花边轮廓信息、扩张外接矩形信息、相对位置坐标信息进行zip操作,并作为当前页面的定位信息;
- [0169] 进行缺失性判断,具体的缺失性判断包括根据当前页面的定位信息中目标边框花边轮廓的顶点信息是否临界外扩外接矩形,来判断目标边框花边轮廓的边框在定位图像中是否会存在缺失,若不存在缺失,则根据轮廓顶点信息进行透视变换并将透视后的透视后指读图像添加到透视结果集中;
- [0170] 若存在缺失则将透视得到的透视后指读图像与上一次的透视得到的透视后指读图像进行比对,若未翻页(两者相同),则使用上一次的页数信息,并直接确定当前指读章节。若翻页,则根据定位信息,确定出目标边框花边轮廓的中可用边缘的方向信息(上下左右),可用边缘的基本信息(斜率,偏移等),四个顶点的方向信息(纵向+横向,如上左),并结合缺失性判断以确定定位图像中可能缺失的角点信息,和四边中较为可靠的三条或两条边的信息;
- [0171] 对定位图像进行掩码操作,生成定位图像掩码,其中,掩码操作包括但不限于Canny、Sobel、自定义二值中至少之一;
- [0172] 对处理后的定位图像掩码进行一定比例的图片融合(如Canny7:0.5,Canny3:0.5),再进行阈值筛选获取一个用于直线检测的二值掩码;
- [0173] 将上一步获取的二值掩码进行概率霍夫直线检测,并将检测到的直线设置第三预设条件进行筛选(直线所在的范围是否在正确的范围内,直线与可靠直线的交点的位置信息等),并根据之前的可靠直线数目获取一条或两条计算的补全直线;
- [0174] 根据补全直线和可靠直线重新求出四个顶点信息,并进行透视变换将透视后指读图像添加到透视结果集中;
- [0175] 对获取的透视结果集中内容分别进行掩码操作,该掩码操作包括但不限于Canny、Sobel、自定义二值等中至少之一。
- [0176] 对透视掩码集合进行轮廓检测,并设置第四预设条件(面积大小,外接矩形长宽比例,位置是否在识别书本内容范围内等)来进行轮廓筛选,并根据获取后的轮廓在信息制作一个用于霍夫变换的霍夫变换掩码;
- [0177] 将上一步获取的霍夫变换掩码进行霍夫圆检测,并设置第五预设条件(圆心位置,圆半径)以进行筛选;

- [0178] 对筛选后的疑似章节节点轮廓信息进行分类聚类;
- [0179] 对分类聚类后的轮廓信息进行求最优解的聚类,确定目标章节节点轮廓,根据目标章节节点轮廓的位置信息、确定目标章节节点轮廓的数量信息等进行编码,生成章节节点编码;
- [0180] 根据章节节点编码中章节节点的位置信息,在预处理得到的存储有基准章节节点编码的数据库里进行比对,当章节节点编码的位置内容为基准章节节点编码的编码位置的子集或全集时,将基准章节节点编码对应的页面图像存入可疑图片集中,其中可以不用考虑当内页内容部分缺失导致的部分节点缺失的影响;例如,章节节点编码信息中包括2个章节节点,其位置分别是第一行首端,第二行中部,此时,包括第一行首端,第二行中部有章节节点的基准章节节点编码所对应的页面图像均加入到可疑图片集中;
- [0181] 将当前透视后的透视后指读图像与可疑图片集进行SIFT (Scale-Invariant Feature Transform,尺度不变特征变换)特征比对,并将特征匹配值最高的可疑图片作为当前页面,该可疑图片对应的页码作为当前页面的页码。可选的,当前页面信息包括当前页面的页码。
- [0182] S703:获取所当前指读章节。
- [0183] 当指示工具包括手指时,进行肤色掩码获取,从透视图像掩码集合中提取HSV和YCbCr 颜色掩码进行图片融合,生成位置指示工具掩码;对生成位置指示工具掩码中进行轮廓检测,并通过第六预设条件进行筛选,确定人手轮廓,并提取轮廓纵轴上最大值点作为指尖位置。其中,第六预设条件包括但不限于面积大小,轮廓所处的区域等。其中,面积大小与人手大小相近。
- [0184] S704:获取所当前指读章节对应的当前指读章节语音并播放。
- [0185] 可选的,当指示工具包括手指时,根据指尖位置和页面信息确定当前指读章节信息,并根据页数和章节数在音频数据库中进行提取音频信息;将音频信息通过设备进行播放,完成整个指读操作。
- [0186] 在一些实施例中,图书指读基于机器人实现,其实现过程如图8所示:
- [0187] S801:开机并开启指指读模式;
- [0188] S802:在设备识别区域的预设区域范围内将图书翻到需要进行指读识别的页面;
- [0189] S803:在预设区域范围内进行指读操作;
- [0190] S804:定时检测,获取指读图像并确定当前页面信息;
- [0191] S805:确定当前指示位置;
- [0192] S806:根据当前页面信息、指示位置确定当前指读章节;
- [0193] S807:确定当前指读章节对应的当前指读章节语音,并播放。
- [0194] 可选的,步骤S801-步骤S803可以基于用户操作来实现,步骤S804-步骤S806可以通过对应的机器人设备来识别。
- [0195] 本发明实施例提供了一种指读方法,通过获取指读图像,指读图像包括内容正文、边框花边、章节节点和位置指示工具,根据指读图像确定当前页面信息、当前指读章节,获取当前指读章节对应的当前指读章节语音并播放,基于现有的部分书籍通过预先处理即可实现指读,不需要重新制作书籍和点读笔,降低了成本,更加便利,提升了用户体验度。
- [0196] 可选的,基于视觉识别技术识别指读对象,对于排版存在略微不同的指读对象,只

需重新进行基准图像、基准章节节点信息的采集,即可实现对新的指读图像的指读,成本较低,简单方便。例如,基于视觉识别技术对不同排版的同一图书进行指读,只需重新进行预处理操作,重新采集基准图像、基准章节节点信息并将其与章节语音对应,就可以实现该图书的指读,可以识别到市面上大部分的与该图书同样内容的版本。

[0197] 可选的,从目标边框花边轮廓的四条轮廓边中确定两条或三条可靠轮廓边,对定位图像进行掩码处理,生成定位图像掩码,并对定位图像掩码进行图片融合,确定二值掩码,对二值掩码进行概率霍夫直线检测,生成补全轮廓边,根据可靠轮廓边和补全轮廓边确定目标边框花边轮廓的各顶点信息;可以实现在由于摆放位置等原因导致的边角缺失情况下,也能将指读图像较好的进行透视变换以进行比对最终确定当前页面信息。根据多掩码加和、轮廓筛选和直线检测的获得页面图像,在边角存在一定缺失的情况下也能做到较好的识别。

[0198] 实施例二

[0199] 参见图9,本发明实施例还提供了一种指读装置900,包括:

[0200] 图像获取模块901,用于获取指读图像,指读图像包括内容正文、边框花边、章节节点和位置指示工具;

[0201] 确定模块902,用于根据指读图像确定当前页面信息、当前指读章节;

[0202] 播放模块903,用于获取当前指读章节对应的当前指读章节语音并播放。

[0203] 在本实施例中,该指读装置实质上是设置了多个模块用以执行上述实施例中的指读方法,具体功能和技术效果参照上述实施例一即可,此处不再赘述。

[0204] 本发明实施例还提供了一种指读机器人,该指读机器人包括图像采集设备、发声设备和上述任一实施例所述的指读装置,

[0205] 图像采集设备用于采集指读图像,并将指读图像发送给指读装置;

[0206] 发声设备用于播放指读装置反馈的当前指读章节语音。

[0207] 可选的,图像采集设备包括30万的摄像头,发声模块包括扬声器,指读装置包括云端服务器。

[0208] 该指读机器人的实现无需要重新印刷指读对象(书籍等),将指读对象放置于指读机器人图像采集的预设采集区域,在对指读对象进行预处理后,只需要把指读机器人联网,图像采集设备将采集到的指读图像通过网络传输给指读装置,指读装置确定好当前指读章节语音后,控制发声设备播放当前指读章节语音,就可以实现指到哪里读哪里,设备简单便携,提升用户体验度。

[0209] 在一些实施例中,当前指读章节语音存储在云端,指读装置向发声设备反馈当前指读章节语音,当前指读章节语音可以通过联网的发声设备进行播放,如接入网络的音箱等。

[0210] 可选的,发声设备用于播放指读装置反馈的当前指读章节语音包括,发声设备根据指读装置的控制播放当前指读章节语音。

[0211] 在一些实施例中,指读装置支持蓝牙功能,也可以通过蓝牙连接发声设备如扬声器、蓝牙耳机等,播放当前指读章节语音。在一些实施例中,图像采集设备、发声设备可以集成在可穿戴设备上,该可穿戴设备支持2G、3G、4G、5G、WIFI网络中至少一种网络连接。

[0212] 在一些实施例中,图像采集设备、发声设备可以集成在智能眼镜框上,该智能眼镜

框支持2G、3G、4G、5G、WIFI网络中至少一种网络连接。发声设备包括两个扬声器,两个扬声器分别设置于两个镜腿上。可选的,扬声器包括骨传导扬声器。

[0213] 参见图10,本发明实施例还提供了一种终端1000,包括处理器1001、存储器1002和通信总线1003;

[0214] 通信总线1003用于将处理器1001和存储器连接1002;

[0215] 处理器1001用于执行存储器1002中存储的计算机程序,以实现如上述实施例一中的一个或多个所述的指读方法。

[0216] 本发明实施例还提供了一种计算机可读存储介质,其特征在于,其上存储有计算机程序,

[0217] 计算机程序用于使计算机执行如上述实施例一中的任一项所述的指读方法。

[0218] 本申请实施例还提供了一种非易失性可读存储介质,该存储介质中存储有一个或多个模块(programs),该一个或多个模块被应用在设备时,可以使得该设备执行本申请实施例的实施例一所包含步骤的指令(instructions)。

[0219] 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一一但不限于一一电、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。

[0220] 上述计算机可读介质可以是上述电子设备中所包含的;也可以是单独存在,而未装配入该电子设备中。

[0221] 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)。

[0222] 附图中的流程图和框图,图示了按照本公开各种实施例的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,依所涉及的功能而定。也要注意的,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。

[0223] 上述实施例仅例示性说明本发明的原理及其功效,而非用于限制本发明。任何熟悉此技术的人士皆可在不违背本发明的精神及范畴下,对上述实施例进行修饰或改变。因此,举凡所属技术领域中具有通常知识者在未脱离本发明所揭示的精神与技术思想下所完成的一切等效修饰或改变,仍应由本发明的权利要求所涵盖。

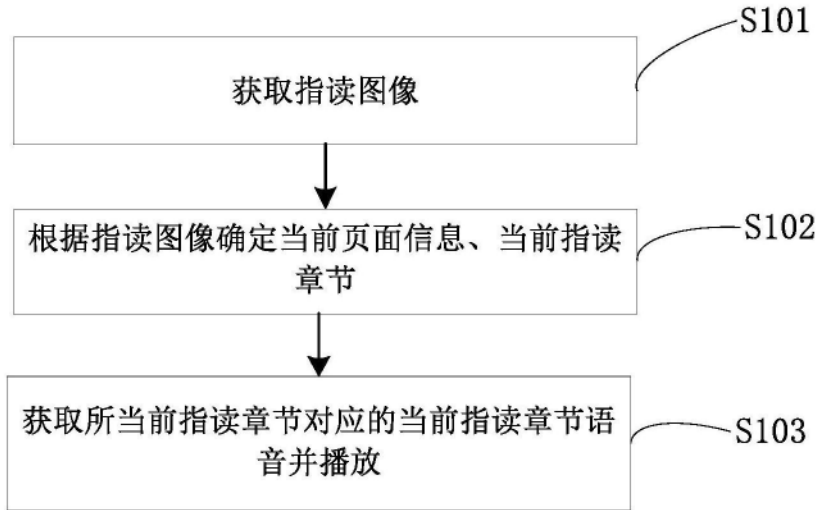


图1

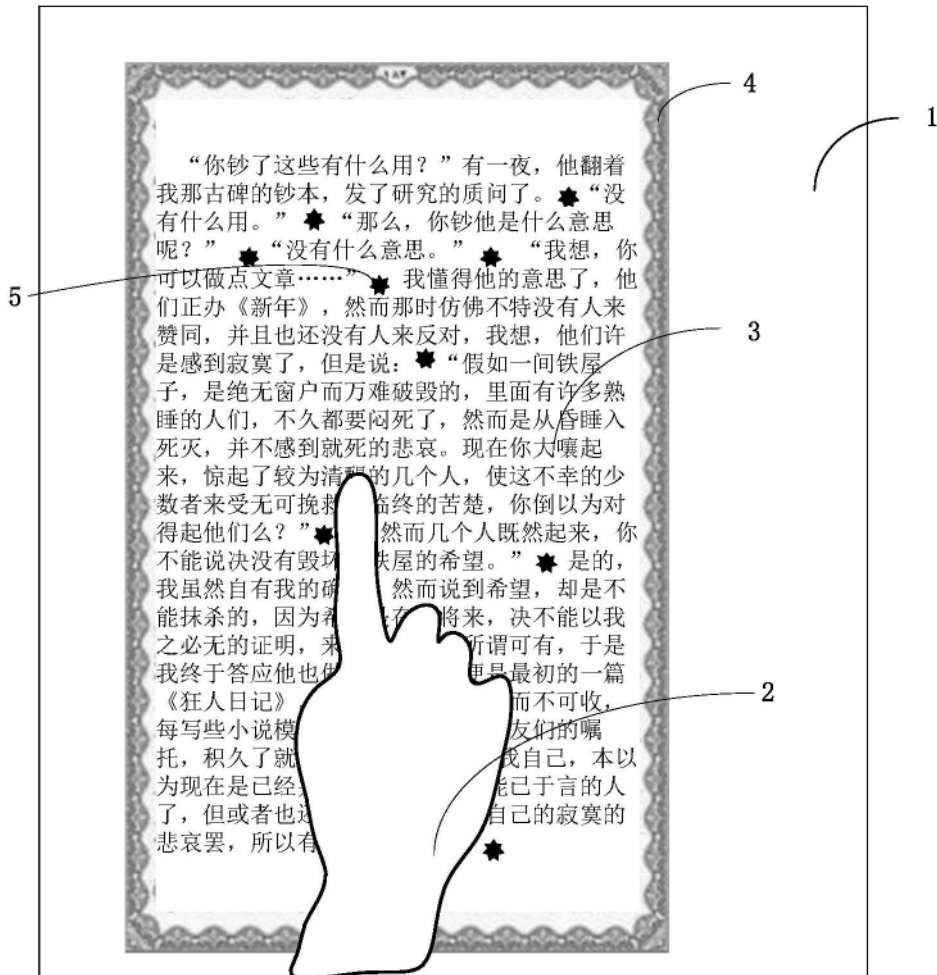


图2

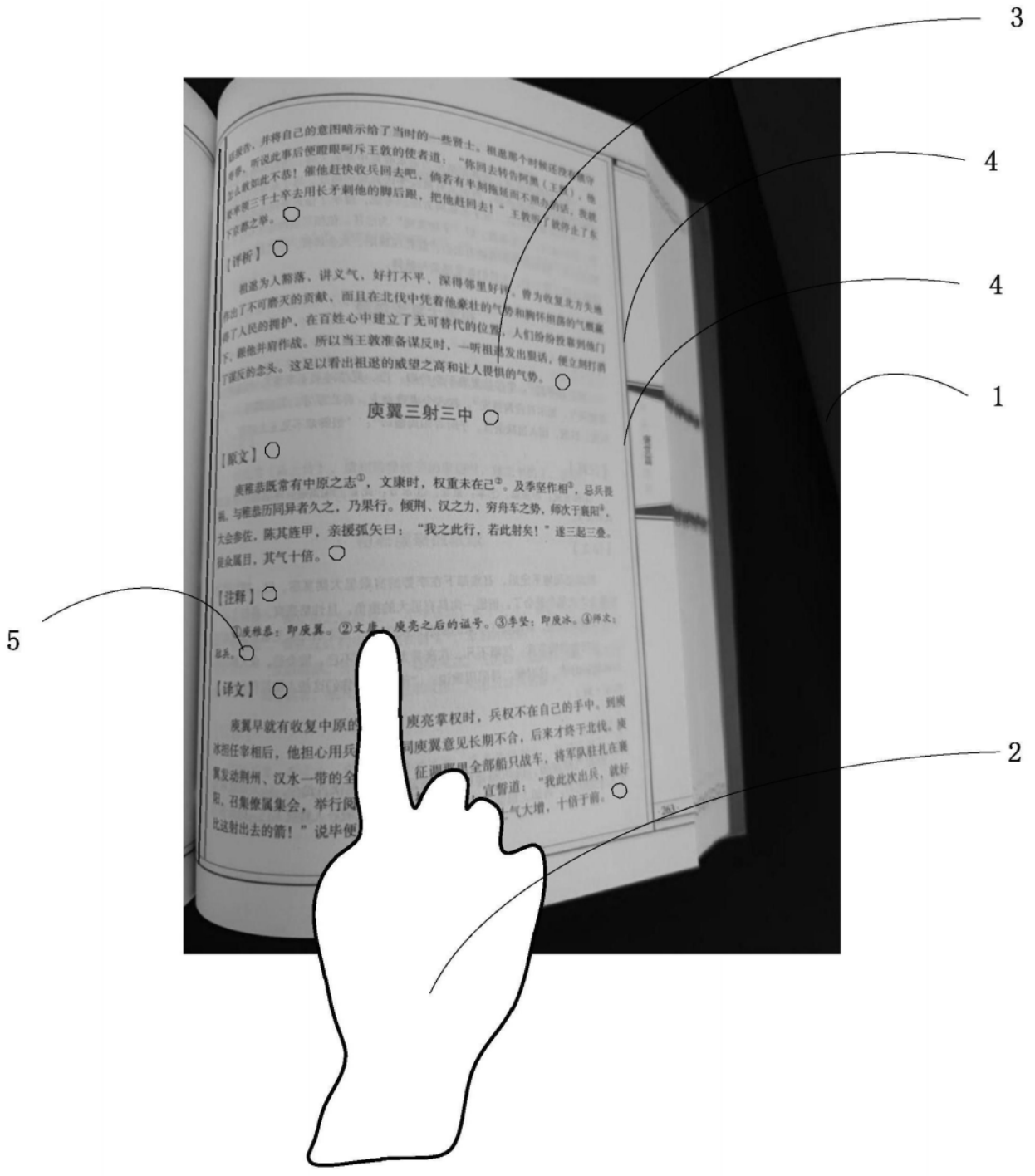


图3

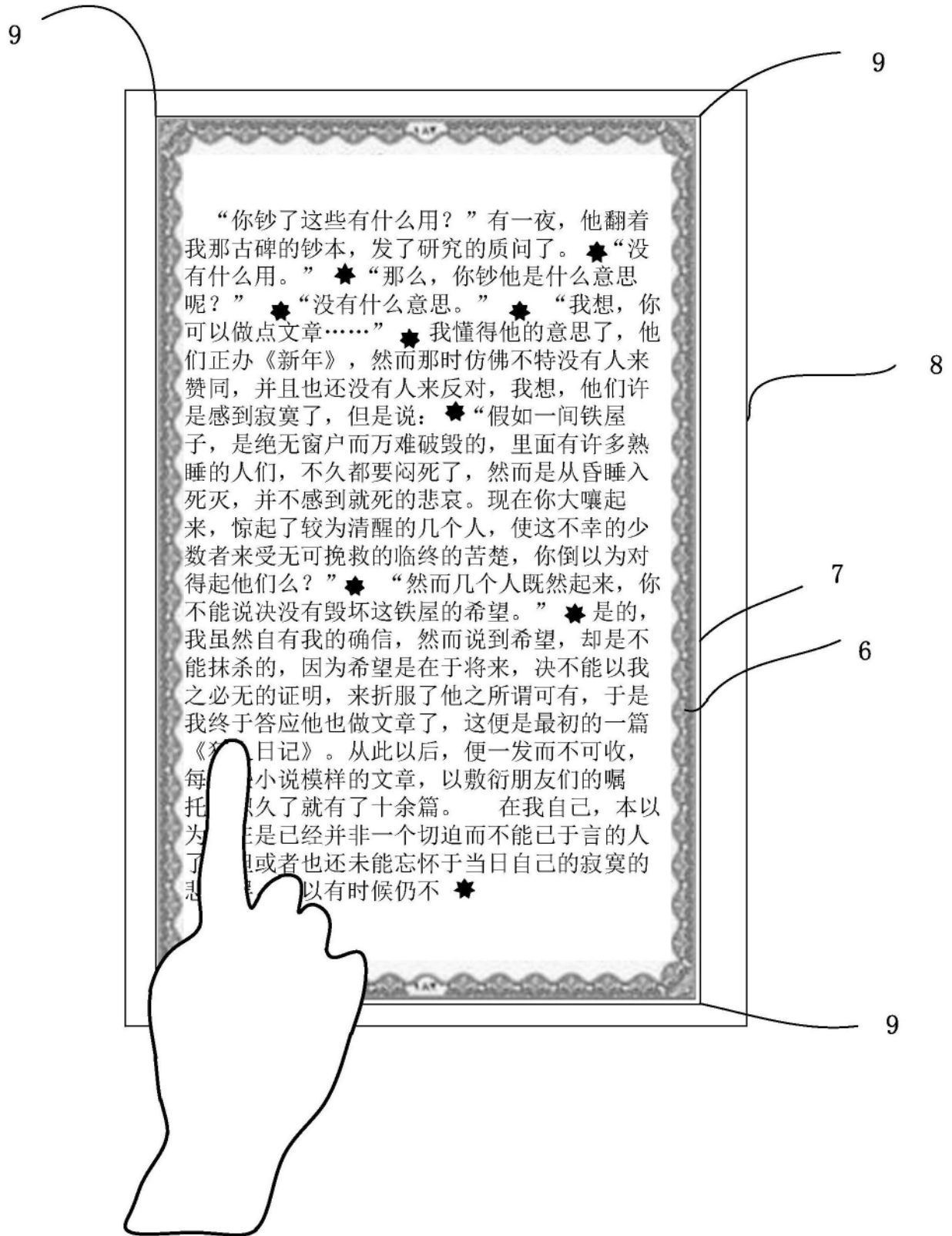


图4

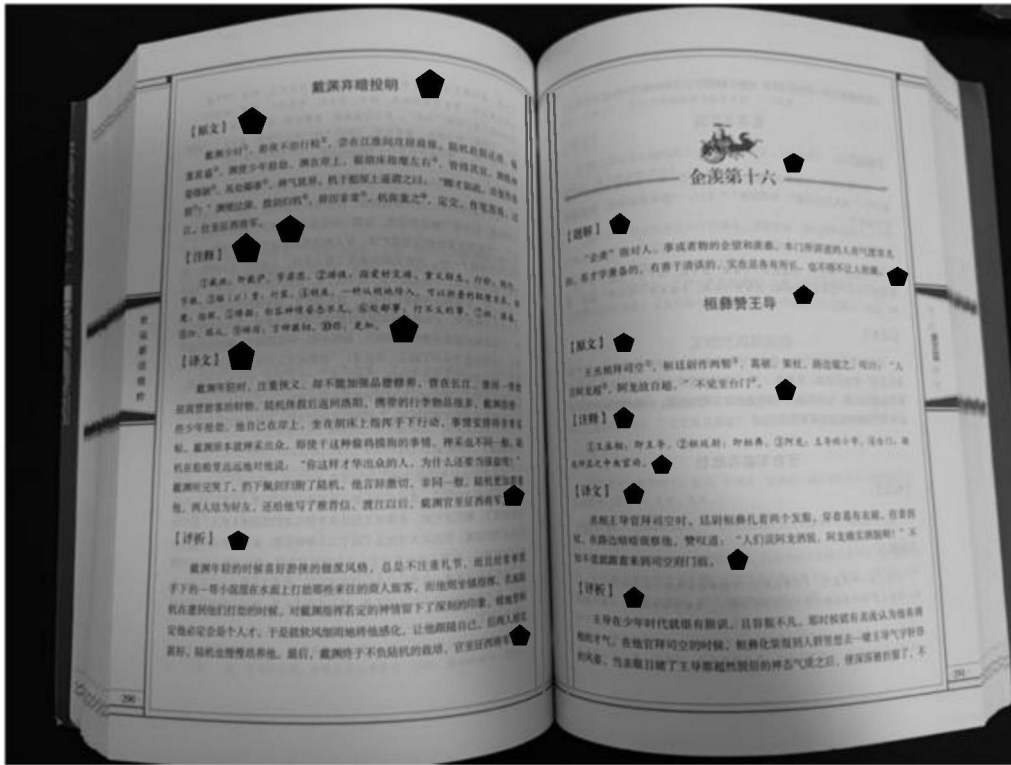


图5

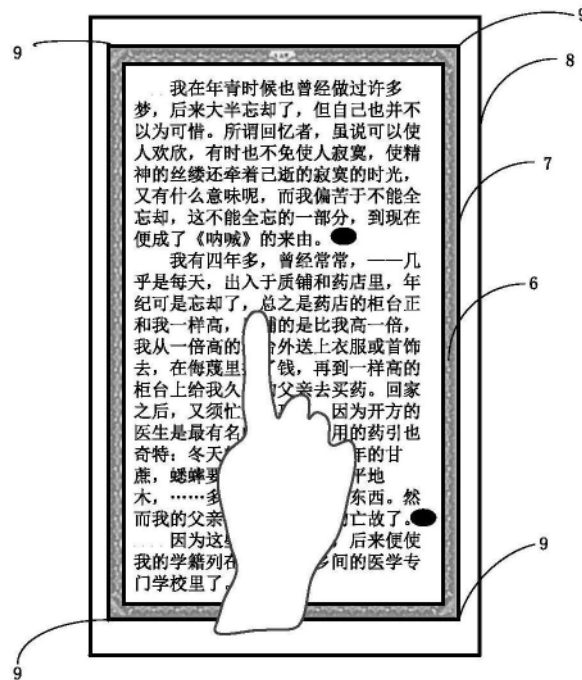


图6

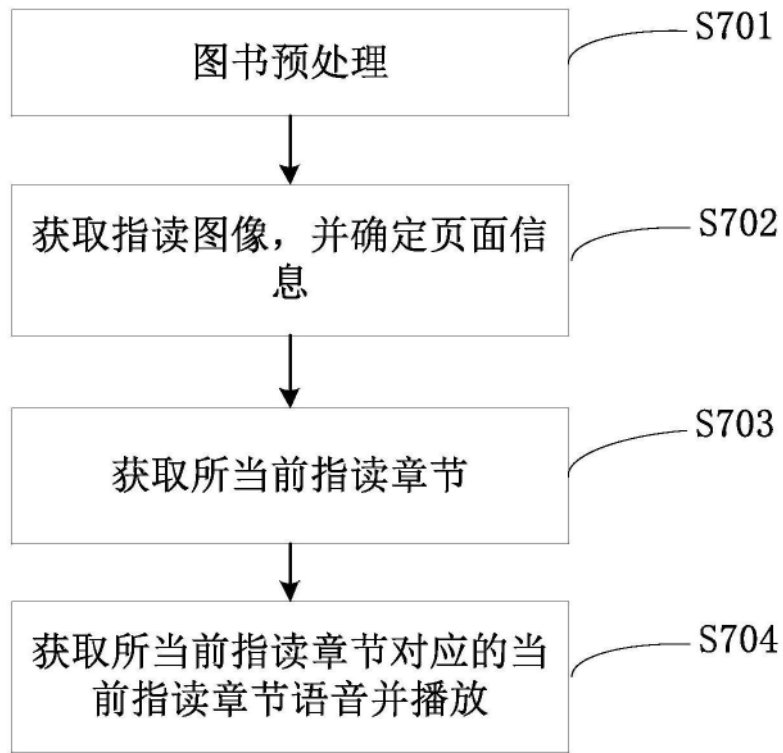


图7

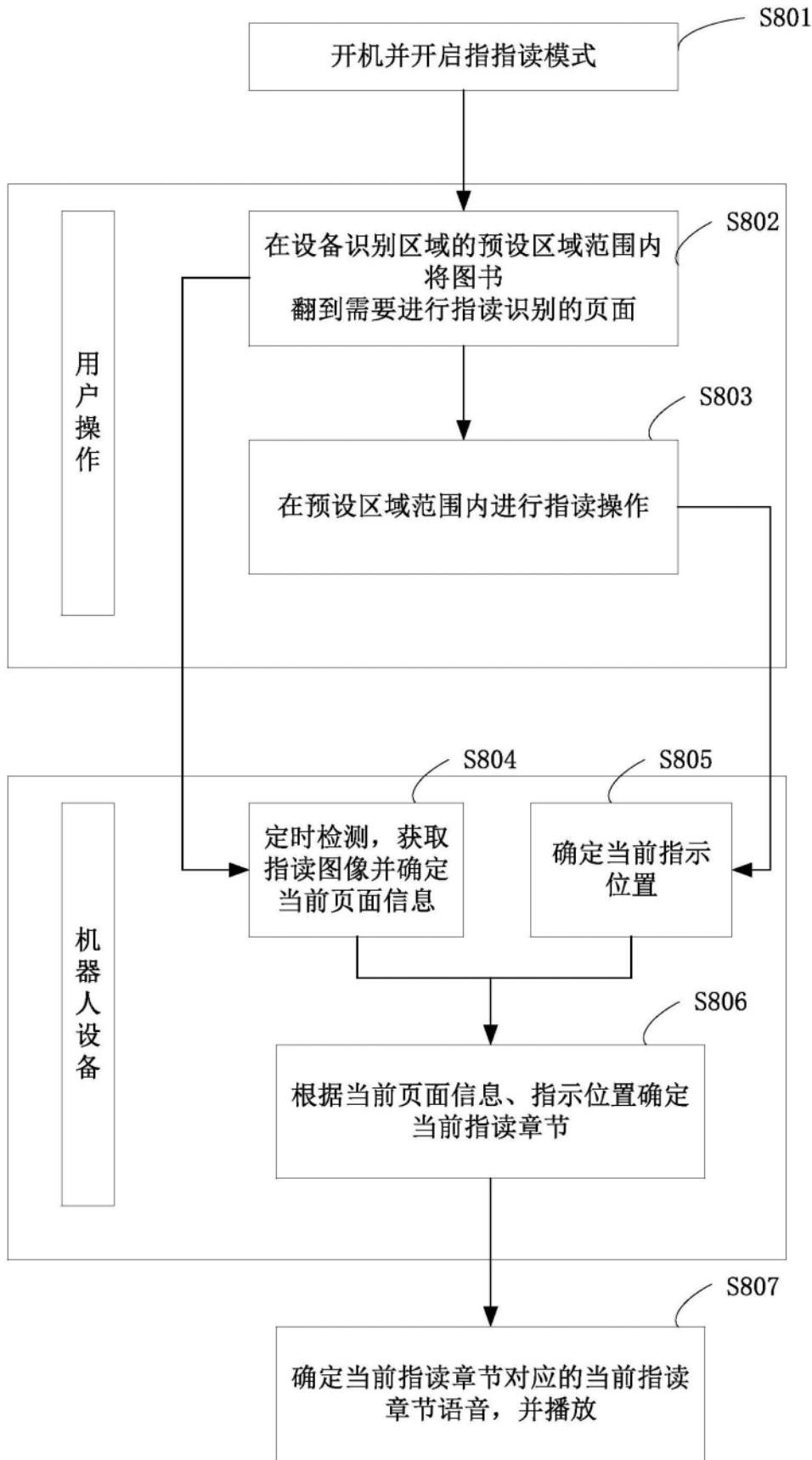


图8

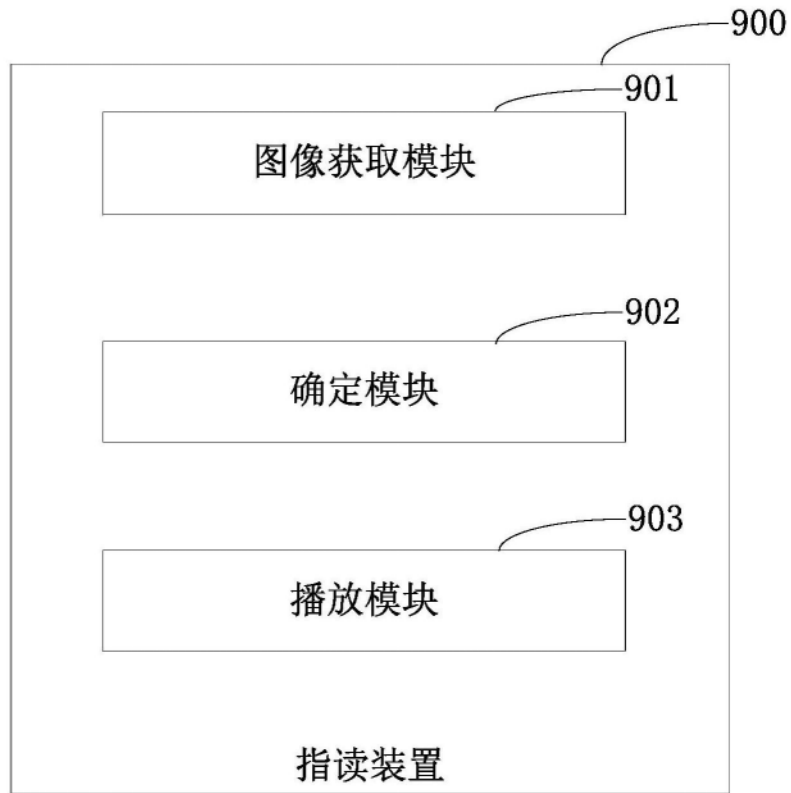


图9

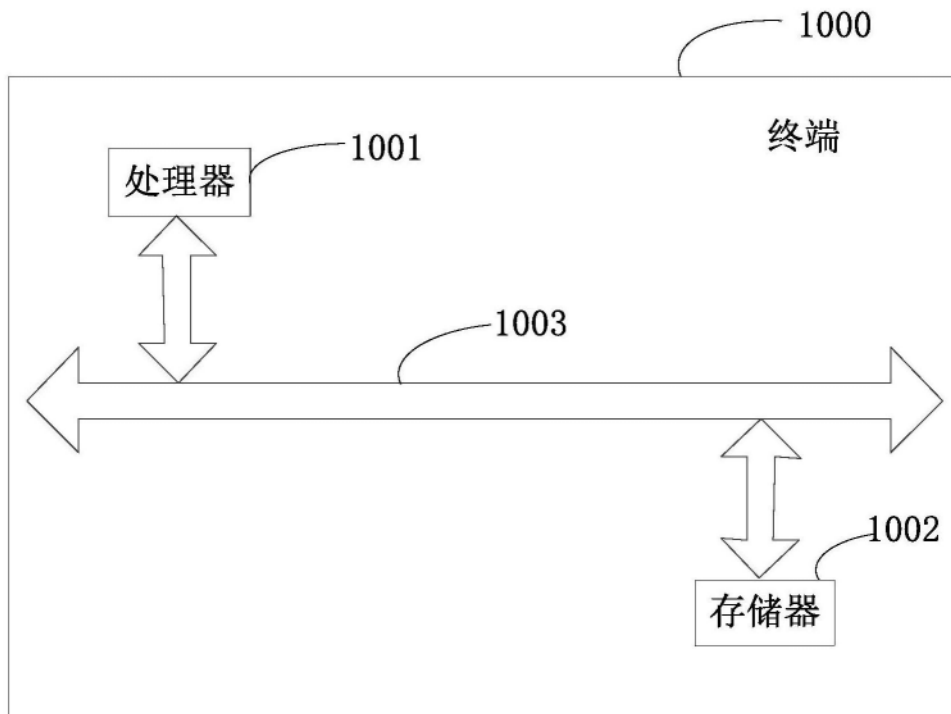


图10