



(12)发明专利申请

(10)申请公布号 CN 110717397 A

(43)申请公布日 2020.01.21

(21)申请号 201910856607.2

(22)申请日 2019.09.11

(71)申请人 中国海洋大学

地址 266100 山东省青岛市崂山区松岭路  
238号

(72)发明人 仲国强 林鑫 岳国华

(74)专利代理机构 北京汇捷知识产权代理事务  
所(普通合伙) 11531

代理人 马金华

(51) Int. Cl.

G06K 9/00(2006.01)

G06K 9/32(2006.01)

G06K 9/38(2006.01)

G06K 9/40(2006.01)

G06F 40/58(2020.01)

权利要求书2页 说明书4页

(54)发明名称

一种基于手机相机的在线翻译系统

(57)摘要

本发明公开了一种基于手机相机的在线翻译系统,自动对焦获取图片;图像预处理;在提取出相应文本信息后,利用谷歌提供的正方体文字识别引擎识别文字;取词识别完成后获取到了需要翻译的文字,将待翻译的文字向翻译模块传递即可返回翻译结果。本发明的有益效果是成功识别及翻译得到准确的结果。

1. 一种基于手机相机的在线翻译系统,其特征按照以下步骤进行:

步骤1:自动对焦获取图片;摄像机对待识别的内容每隔两秒钟自动对焦一次并获取图片,从而通过自动对焦来动态定位并获取图像;

步骤2:图像预处理;对图片的预处理的步骤是:扫描图像-->获得图像-->解析图像-->获取图片参数-->图像去噪-->计算图像阈值-->对图像灰度二值化并保存灰度值-->获得灰度图像;

步骤3:识别功能,在提取出相应文本信息后,利用谷歌提供的正方体文字识别引擎识别文字;

步骤4:翻译功能,取词识别完成后获取到了需要翻译的文字,将待翻译的文字向翻译模块传递即可返回翻译结果。

2. 按照权利要求1所述一种基于手机相机的在线翻译系统,其特征按照以下步骤进行:

所述步骤1中,首先在系统中设定了一个定时器Timer,它会每隔一段时间执行一个TimerTask,并且在TimerTask里执行对焦方法,Camera类提供了自动对焦的方法,它接收AutoFocusCallback回调,这个方法执行后相机就会自动对焦,当它对焦完成后触发回调方法并再次启动Timer,以实现连续间隔的自动对焦,最后设计相机自动对焦方法,判断对焦成功后发出一个消息,间接控制执行拍照获取图片。

3. 按照权利要求1所述一种基于手机相机的在线翻译系统,其特征按照以下步骤进行:

(1) 图像获取:通过调用摄像头,扫描得到固定矩形区域图像,得到图像后,解析图像,获取待处理图像的参数;

(2) 图像去噪,系统所采用的去噪算法是获取设置的固定大小区域中所有像素值的中间值,采取3乘以3的表格区域,pixel[0~8]是一个一维的像素值序列,中心点为pixel[4],对9个像素点的值进行排序,然后取中间值,再存入pixel[4],即中心点,得到去噪后的像素值;

(3) 计算阈值,系统中使用一维最大熵法、大律法和迭代法分别计算图像阈值,综合对比取最优阈值结果;

(4) 灰度二值化处理,系统根据之前计算的最优阈值对图像进行灰度二值化处理;

(5) 提取特征,将经预处理后的图像分割,区域化提取图像特征,取出图像中的文本信息,系统中文字的提取方法使用了Android系统中数据在Activity之间的传递模块,首先将图像置于拍照所在的Activity模块进行处理,同时在该模块完成取词的功能,然后将图片经处理得到的结果返回给翻译模块的Activity,再对文字进行识别与翻译。

4. 按照权利要求3所述一种基于手机相机的在线翻译系统,其特征按照以下步骤进行:

①利用浮点算法对扫描得到的区域图像进行灰度化处理,手机摄像头扫描得到的图像通常是24位深的RGB图像,灰度化处理得到8位深的灰度层次图像;

②图像通过普通二值化处理后得到黑白二色的图像,系统在利用二值化处理时保存了图像灰度信息,根据不同图像的不同像素值,对灰度信息进行二值化,将图片归为只有两个灰度值的图像,最终获得的图片是灰度图片;

③在进行灰度二值化处理之后,先判断深色值像素数是否多于浅色值像素数,如果深

色值像素数比浅色值像素数多,则进行反色处理。

5. 按照权利要求1所述一种基于手机相机的在线翻译系统,其特征在于:所述步骤3包括

(1) 正方体文字识别引擎分为两部分:图片布局分析和字符分割和识别;

① 图片布局分析主要分析连通区域,是字符识别的准备工作,通过一种混合的基于制表位检测的页面布局分析方法,将图像的表格、文本、图片等内容进行区分;

② 字符分割和识别过程是先找到块区域,定位文本行和单词,识别文本,先进行第一次字符识别,通过字符区域类型判定,根据判定结果对比字符库识别字符,然后,根据识别出来的文本字符,进行粘连字符的分割,同时把错误分割的字符合并,完成字符的精细切分;

(2) 正方体文字识别引擎训练:

利用jTessBoxEditor工具来训练语言字库,在该系统中需要完成的任务是对中文和英文的识别与翻译,所以在训练过程中,对tessdata中以下几个数据包进行训练,从而得到系统中应用的语言识别库:

chi\_sim.traineddata,

chi\_tra.traineddata,eng.traineddata,chi\_tra\_vert.traineddata,chi\_sim\_vert.traineddata,在初次使用此系统时,需要先加载识别库,将识别语言库加载到手机存储根目录中才能正确运行识别功能,通过调用训练得到的语言识别库。

6. 按照权利要求1所述一种基于手机相机的在线翻译系统,其特征在于:所述步骤4中翻译功能流程为开始-->获取识别得到的文字信息-->将文本信息编码转换-->MD5加密,生成签名-->传输信息-->获取翻译结果-->在程序界面上显示翻译结果-->结束。

## 一种基于手机相机的在线翻译系统

### 技术领域

[0001] 本发明属于在线翻译技术领域,涉及一种基于手机相机的在线翻译系统。

### 背景技术

[0002] 随着经济的发展,越来越多的人选择出国旅游,但是语言障碍始终困扰着人们。因此,一个能将旅游场景中的文字识别并翻译的软件,将能够满足国内外游客的需求。旅游翻译助手具有OCR实时识别及翻译功能,能帮助游客解决旅游过程中语言障碍的问题。旅游翻译助手实现了基于Android平台的对不同场景中的英文和中文的识别翻译功能。该应用程序基于Java语言开发,安装在安卓手机上,可在多种自然场景下进行调试。测试场景有书本文档识别与翻译、菜单识别与翻译、路牌标识识别与翻译等等。常用翻译软件需手工输入所要翻译的文本才可翻译,但该旅游翻译助手为方便用户使用,设计了相机取词识别后直接翻译的功能,省去了手工输入翻译文本的麻烦。

### 发明内容

[0003] 本发明的目的在于提供一种基于手机相机的在线翻译系统,本发明的有益效果是成功识别及翻译得到准确的结果。

[0004] 本发明所采用的技术方案是按照以下步骤进行:

[0005] 步骤1:自动对焦获取图片;摄像机对待识别的内容每隔两秒钟自动对焦一次并获取图片,从而通过自动对焦来动态定位并获取图像;

[0006] 步骤2:图像预处理;对图像的预处理的步骤是:扫描图像-->获得图像-->解析图像-->获取图片参数-->图像去噪-->计算图像阈值-->对图像灰度二值化并保存灰度值-->获得灰度图像;

[0007] 步骤3:识别功能,在提取出相应文本信息后,利用谷歌提供的正方体文字识别引擎识别文字;

[0008] 步骤4:翻译功能,取词识别完成后获取到了需要翻译的文字,将待翻译的文字向翻译模块传递即可返回翻译结果。

[0009] 进一步,步骤1中,首先在系统中设定了一个定时器Timer,它会每隔一段时间执行一个TimerTask,并且在TimerTask里执行对焦方法,Camera类提供了自动对焦的方法,它接收AutoFocusCallback回调,这个方法执行后相机就会自动对焦,当它对焦完成后触发回调方法并再次启动Timer,以实现连续间隔的自动对焦,最后设计相机自动对焦方法,判断对焦成功后发出一个消息,间接控制执行拍照获取图片。

[0010] 进一步,步骤2包括

[0011] (1) 图像获取:通过调用摄像头,扫描得到固定矩形区域图像,得到图像后,解析图像,获取待处理图像的参数;

[0012] (2) 图像去噪,系统所采用的去噪算法是获取设置的固定大小区域中所有像素值的中间值,采取3乘以3的表格区域,pixel[0~8]是一个一维的像素值序列,中心点为pixel

[4],对9个像素点的值进行排序,然后取中间值,再存入pixel[4],即中心点,得到去噪后的像素值;

[0013] (3) 计算阈值,系统中使用一维最大熵法、大律法和迭代法分别计算图像阈值,综合对比取最优阈值结果;

[0014] (4) 灰度二值化处理,系统根据之前计算的最优阈值对图像进行灰度二值化处理;

[0015] (5) 提取特征,将经预处理后的图像分割,区域化提取图像特征,取出图像中的文本信息,系统中文字的提取方法使用了Android系统中数据在Activity之间的传递模块,首先将图像置于拍照所在的Activity模块进行处理,同时在该模块完成取词的功能,然后将图片经处理得到的结果返回给翻译模块的Activity,再对文字进行识别与翻译。

[0016] 进一步,步骤3包括

[0017] (1) 正方体文字识别引擎分为两部分:图片布局分析和字符分割和识别;

[0018] ①图片布局分析主要分析连通区域,是字符识别的准备工作,通过一种混合的基于制表位检测的页面布局分析方法,将图像的表格、文本、图片等内容进行区分;

[0019] ②字符分割和识别过程是先找到块区域,定位文本行和单词,识别文本,先进行第一次字符识别,通过字符区域类型判定,根据判定结果对比字符库识别字符,然后,根据识别出来的文本字符,进行粘连字符的分割,同时把错误分割的字符合并,完成字符的精细切分;

[0020] (2) 正方体文字识别引擎训练:

[0021] 利用jTessBoxEditor工具来训练语言字库,在该系统中需要完成的任务是对中文和英文的识别与翻译,所以在训练过程中,对tessdata中以下几个数据包进行训练,从而得到系统中应用的语言识别库:

[0022] chi\_sim.traineddata,

[0023] chi\_tra.traineddata,eng.traineddata,chi\_tra\_vert.traineddata,chi\_sim\_vert.traineddata,在初次使用此系统时,需要先加载识别库,将识别语言库加载到手机存储根目录中才能正确运行识别功能,通过调用训练得到的语言识别库。

[0024] 进一步,步骤(4)包括

[0025] ①利用浮点算法对扫描得到的区域图像进行灰度化处理,手机摄像头扫描得到的图像通常是24位深的RGB图像,灰度化处理后得到8位深的灰度层次图像;

[0026] ②图像通过普通二值化处理后得到黑白二色的图像,系统在利用二值化处理时保存了图像灰度信息,根据不同图像的不同像素值,对灰度信息进行二值化,将图片归为只有两个灰度值的图像,最终获得的图片是灰度图片;

[0027] ③在进行灰度二值化处理之后,先判断深色值像素数是否多于浅色值像素数,如果深色值像素数比浅色值像素数多,则进行反色处理。

[0028] 进一步,步骤4中翻译功能流程为开始-->获取识别得到的文字信息-->将文本信息编码转换-->MD5加密,生成签名-->传输信息-->获取翻译结果-->在程序界面上显示翻译结果-->结束。

## 具体实施方式

[0029] 下面结合具体实施方式对本发明进行详细说明。

[0030] 实施例:

[0031] 步骤1:自动对焦获取图片;

[0032] 自动对焦:摄像机对待识别的内容每隔两秒钟自动对焦一次并获取图片,从而通过自动对焦来动态定位并获取图像。首先在系统中设定了一个定时器Timer,它会每隔一段时间执行一个TimerTask,并且在TimerTask里执行对焦方法。Camera类提供了自动对焦的方法,它接收AotoFocusCallback回调。这个方法执行后相机就会自动对焦,当它对焦完成后(成功或失败)触发回调方法并再次启动Timer,以实现连续间隔的自动对焦。最后设计相机自动对焦(raiseEvent\_OnAutoFocusSuccess)方法,判断对焦成功后发出一个消息,间接控制执行拍照获取图片。

[0033] 步骤2:图像预处理;

[0034] 对图片的预处理的步骤是:扫描图像-->获得图像-->解析图像-->获取图片参数-->图像去噪-->计算图像阈值-->对图像灰度二值化并保存灰度值-->获得灰度图像。

[0035] (1) 图像获取:通过调用摄像头,扫描得到固定矩形区域图像。得到图像后,解析图像,获取待处理图像的参数,比如颜色色值、位图行宽、像素数量。这是图像处理的第一步。

[0036] (2) 图像去噪。该系统所采用的去噪算法是获取设置的固定大小区域中所有像素值的中间值。采取3乘以3的表格区域。pixel[0~8]是一个一维的像素值序列,中心点为pixel[4]。对9个像素点的值进行排序,然后取中间值,再存入pixel[4],即中心点,得到去噪后的像素值。

[0037] (3) 计算阈值。该系统中使用一维最大熵法、大律法和迭代法分别计算图像阈值,综合对比取最优阈值结果。

[0038] (4) 灰度二值化处理。该系统根据之前计算的最优阈值对图像进行灰度二值化处理。

[0039] ①利用浮点算法对扫描得到的区域图像进行灰度化处理。手机摄像头扫描得到的图像通常是24位深的RGB图像,灰度化处理后得到8位深的灰度层次图像。

[0040] ②图像通过普通二值化处理后得到黑白二色的图像,它只有黑、白两种颜色,也就是说它的每个像素位深度是1,占一个二进制位。但是,该系统在利用二值化处理时保存了图像灰度信息。根据不同图像的不同像素值,对灰度信息进行二值化,将图片归为只有两个灰度值的图像。最终获得的图片是灰度图片,但是通过灰度二值化后,图片识别精度更高。

[0041] ③在进行灰度二值化处理之后,可以先判断深色值像素数是否多于浅色值像素数,如果深色值像素数比浅色值像素数多,则进行反色处理。

[0042] (5) 提取特征。将经预处理后的图像分割,区域化提取图像特征,取出图像中的文本信息。系统中文字的提取方法使用了Android系统中数据在Activity之间的传递模块。首先将图像置于拍照所在的Activity模块进行处理,同时在该模块完成取词的功能,然后将图片经处理得到的结果返回给翻译模块的Activity,再对文字进行识别与翻译。

[0043] 步骤3:识别功能;

[0044] 在提取出相应文本信息后,该系统利用谷歌提供的正方体文字识别引擎(Tesseract-ocr)识别文字。

[0045] (1) 正方体文字识别引擎(Tesseract-ocr)原理:Tesseract引擎可以分为两部分:图片布局分析和字符分割和识别。

[0046] ①图片布局分析主要分析连通区域,是字符识别的准备工作。它通过一种混合的基于制表位检测的页面布局分析方法,将图像的表格、文本、图片等内容进行区分。

[0047] ②字符分割和识别过程是先找到块区域,定位文本行和单词,识别文本。先进行第一次字符识别,通过字符区域类型判定,根据判定结果对比字符库识别字符。然后,根据识别出来的文本字符,进行粘连字符的分割,同时把错误分割的字符合并,完成字符的精细切分。

[0048] (2) 正方体文字识别引擎(Tesseract-ocr)训练:

[0049] 利用jTessBoxEditor工具来训练语言字库。在该系统中我们需要完成的任务是对中文和英文的识别与翻译,所以在训练过程中,对tessdata中以下几个数据包进行训练,从而得到系统中应用的语言识别库:

[0050] chi\_sim.traineddata,

[0051] chi\_tra.traineddata,eng.traineddata,chi\_tra\_vert.traineddata,chi\_sim\_vert.traineddata。在初次使用此系统时,需要先加载识别库,将识别语言库加载到手机存储根目录中才能正确运行识别功能。通过调用训练得到的语言识别库,可以准确高效地识别文字信息。

[0052] 步骤4:翻译功能;

[0053] 取词识别完成后获取到了需要翻译的文字,我们将待翻译的文字向翻译模块传递即可返回翻译结果,翻译功能流程为开始-->获取识别得到的文字信息-->将文本信息编码转换-->MD5加密,生成签名-->传输信息-->获取翻译结果-->在程序界面上显示翻译结果-->结束。

[0054] 在该系统中,为保证调用接口的安全性,接口采用了IP限制并且通过MD5编码加密传输数据,与服务器交互。首先对待翻译信息利用MD5加密,生成签名进行验证,通过HTTP的Post方式将待翻译信息传递给服务器。翻译结果通过服务器以Json格式返回后,对返回的XML信息解析,将最终的翻译结果以常见的表示形式呈现在翻译结果栏中,这样就实现了翻译功能。信息-摘要算法5(Message Digest Algorithm 5,MD5)为计算机安全领域广泛使用的一种散列函数,用以提供消息的完整性保护。它用于确保信息传输完整一致,是计算机广泛使用的杂凑算法之一。

[0055] 本发明系统不仅具有实时取词识别功能,还提供了在线翻译功能,让用户的使用体验更佳,为用户的在日常生活中提供了方便。该系统的主要技术领域是模式识别与自然语言处理,它主要实现了实时的文本信息识别与翻译功能,是人们在旅游时良好的翻译助手。该系统分别通过图片处理技术、特征提取技术、文字识别技术和翻译技术等技术手段完成了每个子模块的设计,实现了相机取词识别后直接翻译的功能,省去了人们手工输入翻译文本的麻烦。该系统具有广泛的应用前景和重要的实践意义。

[0056] 以上所述仅是对本发明的较佳实施方式而已,并非对本发明作任何形式上的限制,凡是依据本发明的技术实质对以上实施方式所做的任何简单修改,等同变化与修饰,均属于本发明技术方案的范围。