

[19] 中华人民共和国国家知识产权局



# [12] 发明专利申请公布说明书

[21] 申请号 200810167116.9

[51] Int. Cl.

G06K 9/62 (2006.01)

G06K 9/20 (2006.01)

[43] 公开日 2009 年 3 月 18 日

[11] 公开号 CN 101388078A

[22] 申请日 2008.9.27

[21] 申请号 200810167116.9

[71] 申请人 腾讯科技(深圳)有限公司

地址 518044 广东省深圳市福田区振兴路赛格科技园 2 栋东 403 室

[72] 发明人 张程

[74] 专利代理机构 北京德琦知识产权代理有限公司

代理人 罗正云 宋志强

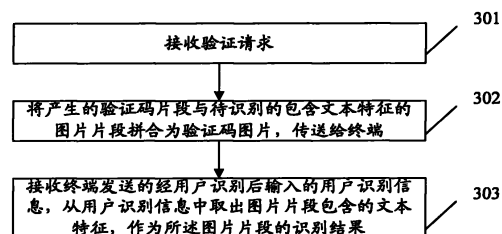
权利要求书 3 页 说明书 10 页 附图 3 页

## [54] 发明名称

基于验证的文本识别的方法及装置

## [57] 摘要

本发明公开了一种基于验证的文本识别的方法及装置,该方法包括:接收验证请求;将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片,传送给终端;接收终端发送的用户对验证码图片识别后输入的用户识别信息,从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果。本发明方案提高了识别结果的正确识别率,解决了计算机识别的局限性。



1、一种基于验证的文本识别的方法，其特征在于，该方法包括：

接收验证请求；

将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片，传送给终端；

接收终端发送的用户对验证码图片识别后输入的用户识别信息，从用户识别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果。

2、如权利要求1所述的方法，其特征在于，所述从用户识别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果之后，该方法包括：

判断得到所述图片片段的识别结果的次数是否大于或等于设定值，如果是，则从各次得到的所述图片片段的识别结果中选取存在概率最大的识别结果，作为所述图片片段的最终识别结果。

3、如权利要求1所述的方法，其特征在于，所述从用户识别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果之后，该方法包括：

判断得到所述图片片段的同一识别结果的次数是否大于或等于设定值，如果是，则将该同一识别结果作为所述图片片段的最终识别结果。

4、如权利要求1所述的方法，其特征在于，所述将验证码片段与图片片段拼合为验证码图片包括：

将验证码片段与图片片段按照固定的位置顺序或按照随机的位置顺序拼合为验证码图片。

5、如权利要求1所述的方法，其特征在于，所述接收终端发送的用户对验证码图片识别后输入的用户识别信息之后，该方法包括：

从用户识别信息中取出验证码信息，判断该取出的验证码信息与所产生的验证码片段中的原始的验证码是否相同，如果相同，则执行所述从用户识别信息中取出图片片段包含的文本特征的步骤。

6、如权利要求1所述的方法，其特征在于，所述将产生的验证码片段与

待识别的包含文本特征的图片片段拼合为验证码图片包括:

提取所述图片片段的背景信息,根据该背景信息在所述验证码片段中加入信息噪声,将所述图片片段与加入信息噪声后的验证码片段拼合为验证码图片。

7、如权利要求 1 至 6 中任一项所述的方法,其特征在于,获取所述待识别的包含文本特征的图片片段的方法包括:从待识别的图像中抽取包含文本特征的图片区域,对图片区域进行分段及编号,得到各个包含文本特征的图片片段,存储于数据库中,从数据库中获取待识别的包含文本特征的图片片段。

8、如权利要求 7 所述的方法,其特征在于,所述从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果之后,该方法包括:

将一幅图像的各个图片片段的识别结果按照各个图片片段的编号进行组合,得到该图像的完整文本特征。

9、一种基于验证的文本识别的装置,其特征在于,该装置包括拼合模块和提取模块;

所述拼合模块,用于接收验证请求,将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片,传送给终端;

所述提取模块,用于接收终端发送的用户对验证码图片识别后输入的用户识别信息,从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果。

10、如权利要求 9 所述的装置,其特征在于,所述提取模块中包括验证码验证子模块和识别结果提取子模块;

所述验证码验证子模块,用于从用户识别信息中取出验证码信息,判断该取出的验证码信息与所产生的验证码片段中的原始的验证码是否相同,如果相同,则向识别结果提取子模块发送启动指定;

所述识别结果提取子模,用于在接收到启动指令后,从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果。

11、如权利要求 9 或 10 所述的装置,其特征在于,该装置进一步包括数据库和图片片段划分模块;

所述图片片段划分模块，用于从待识别的图像中抽取包含文本特征的图片区域，对图片区域进行分段及编号，得到各个待识别的包含文本特征的图片片段，存储于数据库中；

所述拼合模块从数据库中获取待识别的包含文本特征的图片片段。

12、如权利要求 9 所述的装置，其特征在于，所述提取模块进一步用于从用户识别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果之后，判断得到所述图片片段的识别结果的次数是否大于或等于设定值，如果是，则从各次得到的所述图片片段的识别结果中选取存在概率最大的识别结果，作为所述图片片段的最终识别结果。

## 基于验证的文本识别的方法及装置

### 技术领域

本发明涉及图像识别技术，尤其涉及基于验证的文本识别的方法及装置。

### 背景技术

图像识别技术包括对图像中的文本特征进行识别，这里将对图像中文本特征的识别称为文本识别，文本特征包括文字、数字以及符号等。目前，常通过计算机进行文本识别，其中又多采用光学字符识别（OCR，Optical Character Recognition）方法，OCR方法属于图型识别（PR，Pattern Recognition）技术。

参见图1，为现有技术中OCR方法流程图，该方法包括以下步骤：

步骤101，影像输入。

通过扫描仪对含有文本特征的印刷制品等文本图像进行扫描，将扫描后的图像输入计算机。

步骤102，计算机对图像进行影像前处理。

影像前处理包括将图像的尺寸调整为规范尺寸，以及对图像进行降噪处理等。

步骤103，计算机对图像进行文本特征抽取处理。

计算机从待识别的图像中抽取出文本特征。

步骤104，对比识别。

计算机将抽取出的文本特征与标准字库中的文字、数字以及符号等进行比较，将匹配度最高的文字、数字以及符号作为抽取出的文本特征的识别结果，同时，计算机还将提供每个识别结果的正确识别率。

步骤 105, 字词后处理。

计算机将正确识别率不高的词, 与标准词库中的词进行比较, 选择出标准词库中与该正确识别率不高的词最能匹配的词作为该词的识别结果。在比较过程中, 还可同时结合语法规则库判断与该正确识别率不高的词最能匹配的词, 语法规则库中包括了字词组合的规则。例如, 步骤 4 识别出的“中华人民共和国”一词中, “中华人民共和国”的正确识别率为 100%, “国”字的正确识别率只有 60%, 本步骤根据标准词库和语法规则库获知该正确识别率为 60% 的字为“国”字的可能性最大, 因此, 确定该词中正确识别率只有 60% 的字为“国”。

步骤 106, 人工校正。

本步骤对步骤 105 得到的识别结果进行人为调整。根据人为读取的习惯, 调整步骤 105 中得到的识别结果中不准确的字词。

步骤 107, 将最后的识别结果输出。

现有的 OCR 方法存在以下缺点: 步骤 104 和 105 中涉及计算机对文本特征的识别, 由于计算机自身的特点, 对某些文本特征无法正确识别, 导致最后的识别结果的正确识别率较低。

## 发明内容

本发明提供一种基于验证的文本识别的方法, 该方法能够提高识别结果的正确识别率。

本发明提供一种基于验证的文本识别的装置, 该装置能够提高识别结果的正确识别率。

一种基于验证的文本识别的方法, 该方法包括:

接收验证请求;

将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片, 传送给终端;

接收终端发送的用户对验证码图片识别后输入的用户识别信息, 从用户识

别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果。

一种基于验证的文本识别的装置，该装置包括拼合模块和提取模块；

所述拼合模块，用于接收验证请求，将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片，传送给终端；

所述提取模块，用于接收终端发送的用户对验证码图片识别后输入的用户识别信息，从用户识别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果。

从上述方案可以看出，本发明在验证码验证过程中，将待识别的图片片段包含在验证码图片中，通过人眼进行识别，这样，在得到用户输入的验证码的同时，也得到了用户识别出的图片片段中的文本特征，实现对待识别图像中的文本特征的提取，从而，提高了文本特征的正确识别率。

#### 附图说明

图 1 为现有技术中 OCR 方法流程图；

图 2 为现有技术中的验证码图片实例；

图 3 为本发明基于验证的文本识别的方法流程图；

图 4 为本发明基于验证的文本识别的方法流程图实例；

图 5 为本发明基于验证的文本识别的装置结构示意图。

#### 具体实施方式

为使本发明的目的、技术方案和优点更加清楚明白，下面结合实施例和附图，对本发明进一步详细说明。

现有的 OCR 技术中采用计算机对图像进行识别，由于计算机自身的特点，其正确识别率不高；而人眼却能够识别出 OCR 技术不能够正确识别的文本特征，尤其可体现在验证码验证过程中。本发明将待识别的图像包含在验证码图片中，由人眼识别出待识别的图像中的文本特征，实现对待识别图像中的文本特征的提取，这样，提高了文本特征的正确识别率。

下面首先对现有的验证码验证方法进行说明。验证码是攻击-响应 (Challenge-response) 的一种类型, 用来确保产生的响应不是计算机生成的, 通常的办法是让计算机通过验证码向用户询问一个计算机能够生成并能校验的简单问题, 由于其他计算机不能识别验证码, 则输入的正确答案被推定是人类产生的。通过验证码验证正确响应的产生者是人类的方法, 发问者是计算机, 回答者是人类, 因此也被称为反向图灵测试 (Reverse Turing Test) 法。验证码验证过程包括: 网络侧接收终端发送的验证请求; 网络侧产生验证码片段, 验证码片段中包含验证码, 由该验证码片段形成验证码图片, 该验证码图片中包含的验证码可能是对原始验证码加入信息噪声之后的变形, 将验证码图片传送给终端, 终端将验证码图片展示给用户; 用户通过肉眼识别出验证码图片中的验证码, 将用户识别出的验证码输入终端, 终端将用户识别出的验证码发送给网络侧; 网络侧将识别出的验证码与验证码片段中包含的原始验证码进行比较, 如果两者相同, 则验证通过, 如果两者不相同, 则验证失败。

计算机生成的验证码需要具备以下要求: 当前的计算机软件不能够精确识别; 大多数用户能够通过人眼识别; 没有规律可寻, 不能由逻辑分析得出结果, 因为如果有规律可寻, 计算机便有可能分析出结果, 发起进攻。为了加强验证码的安全性, 常在验证码图片中加入噪点、线段、颜色、图形等信息噪声使验证码变形, 图 2 所示便为加入信息噪声后的字母 smwm。在验证码图片中加入信息噪声增加了计算机进行图像识别的难度, 从而进一步防止了计算机的进攻。

本发明通过验证码验证过程实现对图像中文本特征的识别, 对含有文本特征的印刷制品等通过扫描等手段获得待识别的数字图像, 从待识别的图像中抽取包含文本特征的图片区域, 对图片区域进行分段及编号, 得到各个待识别的包含文本特征的图片片段, 在验证码验证过程中, 将待识别的图片片段包含在验证码图片中, 通过人眼对其进行识别, 由于人眼的正确识别率比计算机高, 从而, 实现了高质量的文本特征提取。下面通过图 3 的流程对本



发明基于验证的文本识别的方法进行说明，该流程包括以下步骤：

步骤 301，接收验证请求。

本发明中，由文本识别的装置实现验证码验证，文本识别的装置接收验证请求后，触发验证码验证过程。

步骤 302，将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片，传送给终端。

验证码片段中包含了需要用户识别的验证码，该验证码可能是加入信息噪声之后的变形，需要说明的是，验证码片段与图片片段相互独立，图片片段中的文本特征没有经过变形，是待识别图像中的原始形式。待识别的包含文本特征的图片片段可以预先存储在数据库中，当需要时从数据库中提取；待识别的包含文本特征的图片片段也可以在验证码验证过程中根据需要实时产生，这种情况下，无需将待识别的包含文本特征的图片片段预先存储于数据库中。每个图片片段对应一个唯一的编号，图片片段的设置方法为：对含有文本特征的印刷制品等通过扫描等手段获得待识别的数字图像，从待识别的图像中抽取包含文本特征的图片区域，对图片区域进行分段及编号，得到各个包含文本特征的图片片段，存储于数据库中。需要识别的文本特征包括呈现于图像中的文字、数字及字符等。

步骤 303，接收终端发送的经用户对验证码图片识别后输入的用户识别信息，从用户识别信息中取出图片片段包含的文本特征，作为所述图片片段的识别结果。

终端将接收的验证码图片展示给用户，用户通过人眼识别后将识别结果输入终端，该识别结果便是所述的用户识别信息，包括验证码信息和对图片片段进行识别后的文本特征，终端再将用户识别信息传送给文本识别的装置。

参见图 4，为本发明基于验证的文本识别的方法流程图实例，本实施例以注册过程中的验证码验证为例，该方法预先设置数据库，从待识别的图像中抽取包含文本特征的图片区域，对图片区域进行分段及编号，得到图片片

段，存储于数据库中。该方法包括以下步骤：

步骤 401，接收验证请求。

进行注册的过程中，需要防止计算机的攻击，排除进行注册的主体是计算机的情况，因此，可在用户完成注册表填写之后，提交注册表之前，进行验证码验证操作，验证请求的发起可通过点击用户界面的“提交”键实现。

步骤 402，从数据库中抽取图片片段，将验证码片段与抽取的图片片段拼合为验证码图片。

验证码片段中包含的验证码可能是被加入噪声之后的变形，这里将原始的验证码表示为 A，将从数据库中抽取的图片片段所包含的真正的文本特征表示为 B。

将验证码片段与抽取的图片片段拼合为一张验证码图片时，可以按照固定的位置顺序进行拼合，例如，按照验证码片段在左，图片片段在右的位置顺序进行拼合，或者按照验证码片段在右，图片片段在左的位置顺序进行拼合；为了防止用户得知该固定的位置顺序而随意输入对应图片片段的识别结果而造成识别质量下降，也可以按照随机的位置顺序进行拼合。采用随机的位置顺序进行拼合，可进一步提高识别质量。

在将验证码片段与抽取的图片片段拼合为验证码图片时，进一步地，还可以提取图片片段的背景信息，根据该背景信息在验证码片段中加入信息噪声，然后，将图片片段与加入信息噪声后的验证码片段拼合为验证码图片。例如，图片片段的背景为红色，则将验证码片段的背景也设置为红色，然后再进行拼合；再如，图片片段的背景为条纹，则将验证码片段的背景也设置为条纹，然后再进行拼合。这样，可减小图片片段与验证码片段之间的背景差别，防止用户根据图片片段与验证码片段之间的差别得知图片片段而随意输入对应图片片段的识别结果而造成识别结果下降，从而，可以进一步提高识别质量。

步骤 403，将验证码图片传送给终端，由终端展示给用户。

步骤 404，接收终端传送的用户识别信息。

将拼合的验证码图片展示给用户后,用户通过肉眼识别出验证码片段和图片片段中的信息,将识别出的信息输入终端,这里将用户输入终端的信息称为用户识别信息。用户识别信息中包含用户识别出的验证码,表示为  $a$ ,还包括除  $a$  外的文本特征,表示为非  $a$ ,非  $a$  便是用户对图片片段进行识别后得到的文本特征。

步骤 405,从用户识别信息中取出  $a$ ,判断  $a$  与  $A$  是否相同,如果相同,则验证成功,执行步骤 406,否则,验证失败,执行步骤 401。

本步骤是为了进一步提高识别率,为可选。

步骤 406,存储用户识别信息中的非  $a$ 。

至此,完成了对图片片段的一次识别,得到的非  $a$  便是相应图片片段的识别结果。为了进一步提高识别率,可以对同一图片片段进行多次识别,从多次识别后得到的识别结果中选取存在概率最大的非  $a$ ,作为最后的识别结果,下面通过步骤 407~408 进行说明。

步骤 407,判断存储的与步骤 402 中取出的图片片段的编号对应的非  $a$  的数目是否大于或等于预先设置的设定值,如果是,则执行步骤 408,否则执行步骤 401。

存储的与步骤 402 中取出的图片片段的编号对应的非  $a$  的数目,也就是得到与步骤 402 中取出的图片片段的编号对应的非  $a$  的次数。

假设存储的与某一图片片段的编号对应的非  $a$  信息包括:

用户  $U_1$  输入的非  $a$ , 为  $B_1$ ,

用户  $U_2$  输入的非  $a$ , 为  $B_2$ ,

用户  $U_3$  输入的非  $a$ , 为  $B_1$ ,

用户  $U_4$  输入的非  $a$ , 为  $B_1$ ,

用户  $U_5$  输入的非  $a$ , 为  $B_1$ ,

用户  $U_6$  输入的非  $a$ , 为  $B_3$ 。

与某一图片片段的编号对应的非  $a$  信息,也就是多个用户识别同一编号的图片片段后输入的非  $a$  信息。

步骤 408, 从存储的与步骤 402 中取出的图片片段的编号对应的非 a 信息中, 选取存在概率最大的非 a, 作为该编号的图片片段的识别结果。

以步骤 407 中存储的与某一图片片段的编号对应的非 a 信息为例, 六位用户识别出的非 a 信息中, 用户 U1、用户 U3、用户 U4 和用户 U5 这四位用户识别出的非 a 都为 B1, 也就是 B1 的存在概率为  $4/6$ , 只有用户 U2 识别出的非 a 为 B2, 也就是 B2 的存在概率为  $1/6$ , 只有用户 U6 识别出的非 a 为 B3, 也就是 B3 的存在概率为  $1/6$ , 可见, 非 a 信息中, 识别概率最大的是 B1。

步骤 407~408 是为了进一步提高识别率, 为可选。

从对同一图片片段进行多次识别后得到的识别结果中确定出最后的识别结果, 可以采用步骤 407~408 描述的方式, 即从多次识别后得到的识别结果中选取存在概率最大的非 a, 作为最后的识别结果; 还可以采用如下方式: 判断得到同一图片片段的同一识别结果的次数是否大于或等于设定值, 如果是, 则将该同一识别结果作为该图片片段的最终识别结果, 判断时机的确定可以采用多种方法, 例如, 每接收一个识别结果就判断一次, 也可以随机地对识别结果进行判断。这里以步骤 407 中所举的某一图片片段的编号对应的非 a 信息的实际例子进行说明, 假设采用随机对识别结果进行判断, 且设定值为 4, 如果在接收到 4 个识别结果时进行判断, 此时用户 U1、用户 U2、用户 U3 和用户 U4 分别输入了识别结果, 判断出得到 B1 的次数为 3, 得到 B2 的次数为 1, 可见, 得到 B1 和 B2 的次数都未达到设定值 4, 此时还不能确定出最终的识别结果; 如果在接收到 6 个识别结果时进行判断, 判断出得到 B1 的次数为 4, 得到 B2 的次数为 1, 得到 B3 的次数为 1, 可见, 得到 B1 的次数已经等于 4, 则将 B1 作为最终的识别结果。

按照步骤 401 至步骤 408, 逐一对数据库中存储的其他编号的图片片段进行识别。

最后, 按照图片片段的编号进行组合, 得到图像的完整文本特征。

参见图 5, 为本发明基于验证的文本识别的装置结构示意图, 该装置包

括拼合模块和提取模块;

所述拼合模块,用于接收验证请求,将产生的验证码片段与待识别的包含文本特征的图片片段拼合为验证码图片,传送给终端;

所述提取模块,用于接收终端发送的用户对验证码图片识别后输入的用户识别信息,从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果。

可选地,所述提取模块中包括验证码验证子模块和识别结果提取子模块;

所述验证码验证子模块,用于从用户识别信息中取出验证码信息,判断该取出的验证码信息与所产生的验证码片段中的原始的验证码是否相同,如果相同,则向识别结果提取子模块发送启动指定;

所述识别结果提取子模块,用于在接收到启动指令后,从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果。

可选地,该装置进一步包括数据库和图片片段划分模块;

所述图片片段划分模块,用于从待识别的图像中抽取包含文本特征的图片区域,对图片区域进行分段及编号,得到各个待识别的包含文本特征的图片片段,存储于数据库中;

所述拼合模块从数据库中获取待识别的包含文本特征的图片片段。

可选地,所述提取模块进一步用于从用户识别信息中取出图片片段包含的文本特征,作为所述图片片段的识别结果之后,判断得到所述图片片段的识别结果的次数是否大于或等于设定值,如果是,则从各次得到的所述图片片段的识别结果中选取存在概率最大的识别结果,作为所述图片片段的最终识别结果。

本发明在进行验证码验证时,验证码图片不仅包括验证码片段还包括图片片段,这样,在得到用户输入的验证码的同时,也得到了用户识别出的图片片段中的文本特征,从而,实现对图片片段中文本特征的提取。由于通过人眼能够识别出 OCR 技术不能正确识别的文本特征,从而,提高了正确识别率。

进一步地,进行文本识别时,可统计多个用户输入的文本特征,也就是得到多个用户对图片片段的解读,从多个用户输入的文本特征中选取最多用户输

入的相同的文本特征，作为这个图片片段的识别信息，从而完成高质量的文本提取工作。

以上所述的具体实施例，对本发明的目的、技术方案和有益效果进行了进一步详细说明，所应理解的是，以上所述仅为本发明的具体实施例而已，并不用于限定本发明的保护范围，凡在本发明的精神和原则之内，所做的任何修改、等同替换、改进等，均应包含在本发明的保护范围之内。

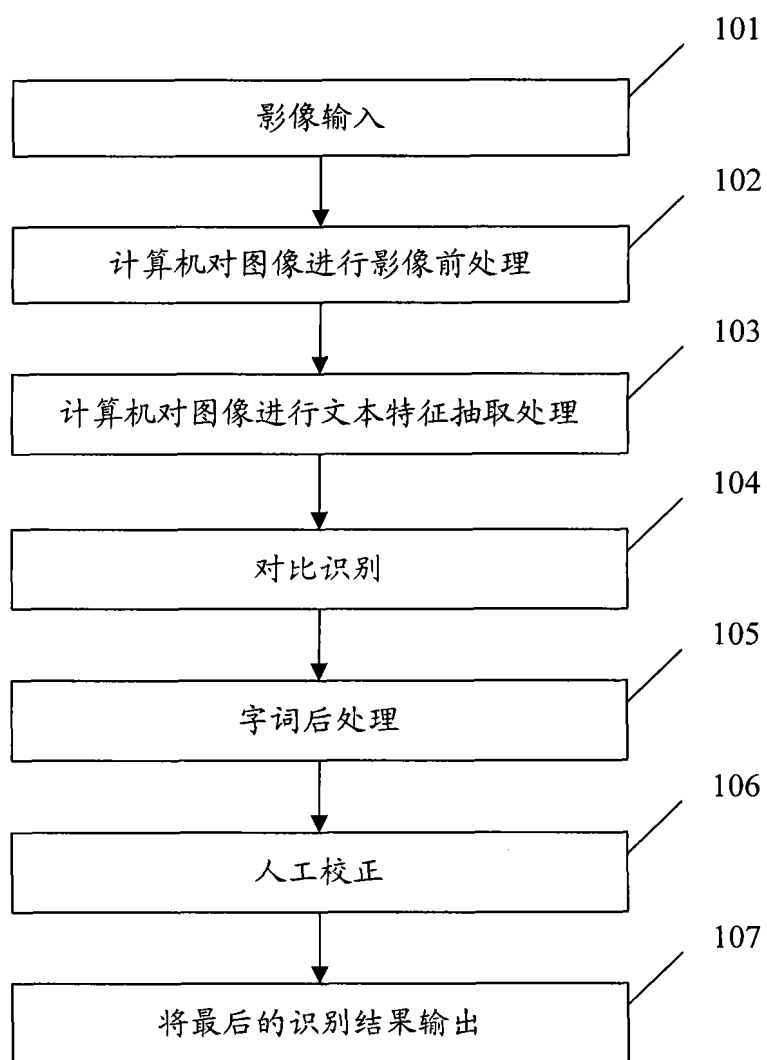


图 1

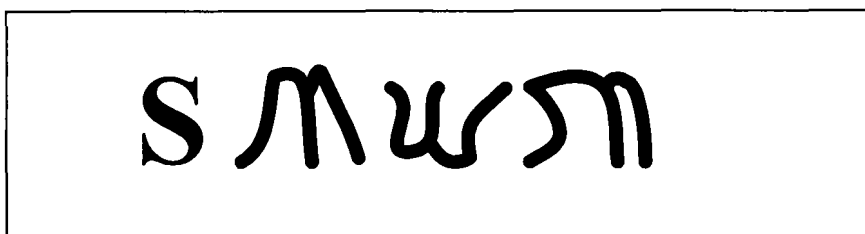


图 2

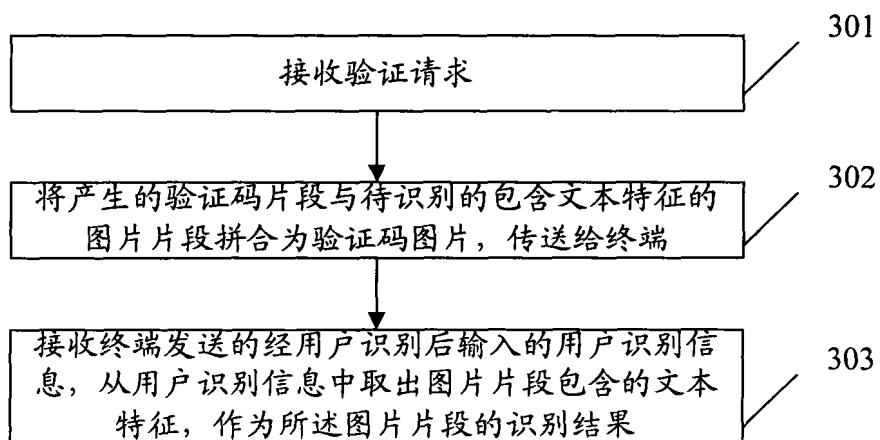


图 3

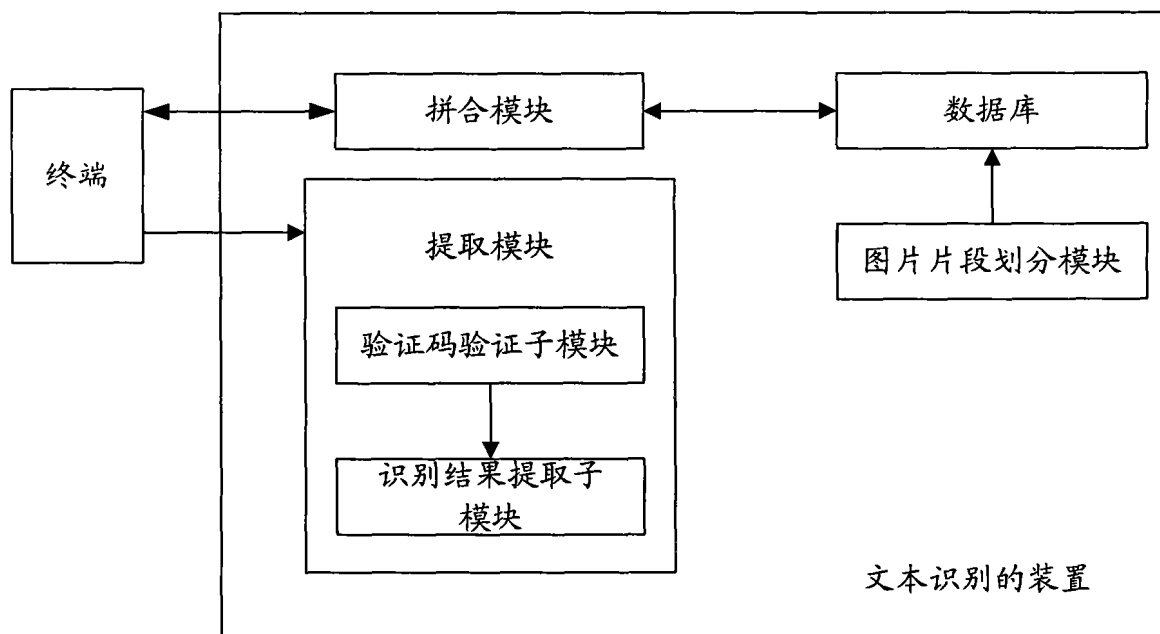


图 5



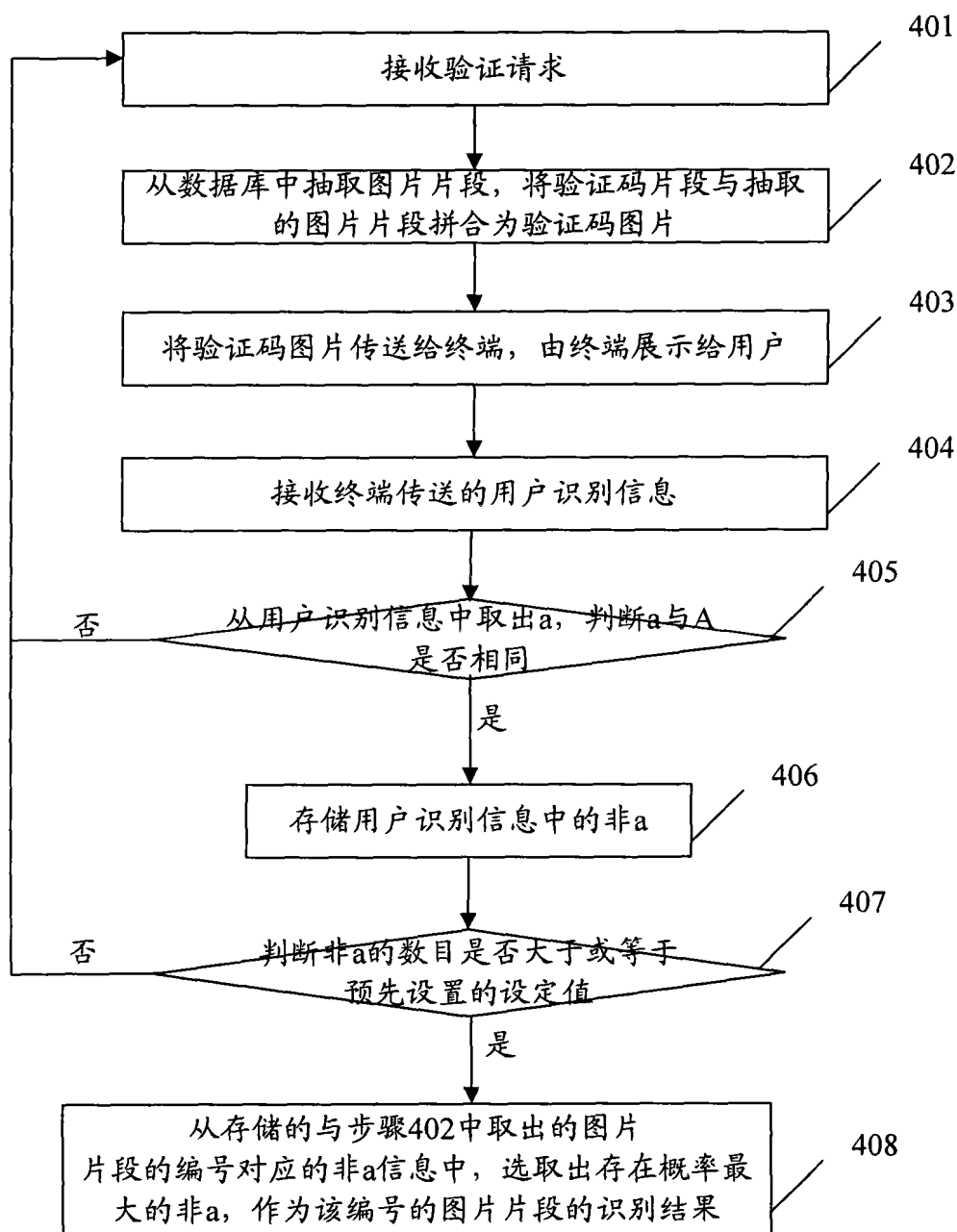


图 4