

[19] 中华人民共和国国家知识产权局

[51] Int. Cl<sup>7</sup>

G06K 9/18

G06K 9/32 G06K 9/36

G06T 1/00



# [12] 发明专利申请公开说明书

[21] 申请号 03121356.1

[43] 公开日 2004年9月29日

[11] 公开号 CN 1532768A

[22] 申请日 2003.3.26 [21] 申请号 03121356.1

[71] 申请人 摩托罗拉公司

地址 美国伊利诺斯

[72] 发明人 镇立新 李俊 黄建成

[74] 专利代理机构 北京市中咨律师事务所

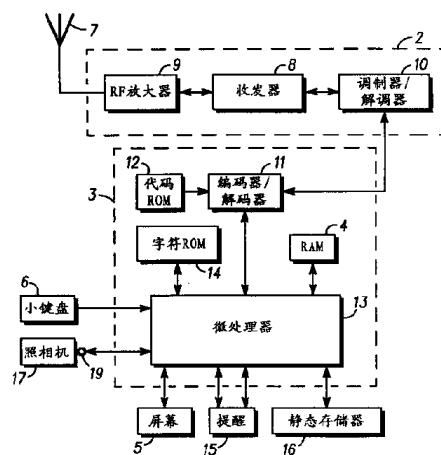
代理人 杨晓光 李峰

权利要求书 2 页 说明书 6 页 附图 2 页

[54] 发明名称 将数字图像的字段类型分类的方法和设备

[57] 摘要

将照相机(17)捕捉的图象中文本字段类型分类的方法(20)和电子设备(1)。在使用时,方法(20)和设备(1)从照相机(17)获得图象,并且识别图象的文本区域。接着对文本区域中的文本进行字符识别,以提供被分类成字段类型的输出字符数据。接着在存储器(16)中指示字段类型的位置上存储输出字符数据。



ISSN 1008-4274

1. 一种通过电子设备完成的方法，该方法用于将照相机捕捉的图象中的文本字段类型分类，该方法包含：

获得图象；

识别图象的至少一个文本区域；

对至少一个所述文本区域中的文本进行字符识别，以提供输出字符数据；

将所述至少一个文本区域分类为一个类型；和

在指示所述字段类型的位置存储所述输出字符数据。

2. 根据权利要求1的方法，该方法还包含旋转图象以基本消除文本偏斜的步骤，该步骤在所述分类之前进行。

3. 根据权利要求1的方法，其中所述的获得图象步骤包括对所述图象执行分辨率增强的步骤。

4. 根据权利要求3的方法，其中如果图象的分辨率低于阈值，进行分辨率增强。

5. 根据权利要求3的方法，其中如果图象被表示成灰度图象，则所述的获得步骤包括对所述图象执行二进制化的步骤。

6. 根据权利要求1的方法，其特征还在于图象是名片图象。

7. 根据权利要求1的方法，其中所述的将所述至少一个文本区域分类为一个字段类型是根据一组规则进行的。

8. 根据权利要求1的方法，其中所述存储步骤在电子设备的至少一个地址簿字段中存储所述输出字符数据。

9. 根据权利要求8的方法，其中所述地址簿字段是电话号码字段。

10. 根据权利要求8的方法，其中所述地址簿字段是人名字段。

11. 根据权利要求8的方法，其中所述地址簿字段是街区地址字段。

12. 根据权利要求8的方法，其中所述地址簿字段是公司名称字段。

13. 一种用于将图象中文本字段类型分类的电子设备，该设备包括：  
处理器；

- 与所述处理器耦合的存储器；和
- 照相机输入端口，其中在使用时，照相机输入端口允许获得图象，而后处理器对图象的至少一个文本区域进行识别，此后处理器完成：
- 对至少一个所述文本区域进行字符识别，以提供输出字符数据；
- 将所述至少一个文本区域分类为一个字段类型；和
- 在存储器中指示所述字段类型的位置上存储所述输出字符数据。
14. 根据权利要求 13 的电子设备，其中所述的照相机输入端口被耦合到照相机。
15. 根据权利要求 13 的电子设备，其中所述处理器提供图象旋转以基本消除文本偏斜，所述旋转在处理器将文本区域分类为一个字段类型之前进行。

## 将数字图像的字段类型分类的方法和设备

### 技术领域

本发明涉及将包含文本的数字图像的字段类型分类的方法和设备，特别涉及，但不限于此，将名片的数字图像的字段类型分类的方法和设备。

### 背景技术

众所周知，人们可以通过桌面扫描设备从名片获得数据。这样的扫描设备需要以对准方式将名片插入槽口，使得可以方便地识别诸如个人名称字段和电话字段的字段。一旦字段被识别，扫描仪可以存储字段的相关信息（人名、地址、电话号码、公司名称等等），相关信息可以被下载到计算机、个人数字助理、蜂窝电话或任何其他适当的电子设备中的用户个人地址簿。然而，出差或外出开会时不便携带这种扫描设备，因此用户通常不得不等到其回到办公室或家里之后，才能扫描其收到的名片。

在美国专利 6,178,270 中描述了一种处理，例如书面文件的图象的方法和设备。根据该专利，使用照相机捕捉文档图象，然后用户选择图象中的文本区域以进行处理。接着确定偏斜角度以用于选定文本区域的处理。虽然这种方法和设备被用于处理图象，然而它们不适于捕捉与字段类型的分类相关的数据，其中根据照相机或其它类似设备捕捉的名片数字图像识别这些字段类型。

在本说明书和权利要求书中，术语"包括"或类似术语被用来表示非排它性的包含，使得包括一组单元的方法或装置不仅包含那些单元，而且可以包含未列出的其它单元。

### 发明内容

根据本发明的一个方面，提供一种通过电子设备完成的方法，该方法用于将照相机捕捉的图象中的文本字段类型分类，该方法包含：

获得图象；

识别图象的至少一个文本区域；

对至少一个所述文本区域中的文本进行字符识别，以提供输出字符数据；

将所述至少一个文本区域分类为一个字段类型；和

在指示所述字段类型的位置存储所述输出字符数据。

该方法最好包含旋转图象以基本消除文本偏斜的步骤，所述旋转在所述分类之前进行。

获得步骤可以适当包含增强所述图象分辨率的步骤。如果图象的分辨率低于阈值，最好进行分辨率增强。

如果图象是灰度图象，获得步骤最好包含对所述图象进行二进制化的步骤。

适当地，该方法的特征还在于图象是名片图象。

最好根据一组规则将所述至少一个文本区域分类为一个字段类型。

存储步骤最好在电子设备的至少一个地址簿字段中存储所述输出字符数据。

所述地址簿字段最好是电话号码字段。

适当地，所述地址簿字段是人名字段。

所述地址簿字段最好是街区地址字段。

适当地，所述地址簿字段是公司名称字段。

根据本发明的另一个方面，提供一种电子设备，该电子设备用于将图象中的文本的字段类型分类，该设备包括：

处理器；

与处理器连接的存储器；和

照相机输入端口，其中在使用时，照相机输入端口允许获得图象，接着处理器对图象的至少一个文本区域进行识别，此后处理器完成：

对至少一个所述文本区域进行字符识别，以提供输出字符数据；

将所述至少一个文本区域分类为一个字段类型；和

在存储器中指示所述字段类型的位置上存储所述输出字符数据。

照相机输入端口最好被连接到照相机。

适当地，处理器旋转图象以基本消除文本偏斜，旋转在处理器将文本区域分类为一个字段类型之前进行。

#### 附图说明

为了方便理解和实施本发明，现在参照附图描述优选实施例，其中：

图 1 的方框图示出了基于本发明的电子设备的实施例；

图 2 的流程图示出了将图 1 电子设备的照相机捕捉的图象中的文本的字段类型分类的方法；而

图 3 的流程图示出了图 2 中使用的获得图象过程的方法。

#### 具体实施方式

在所有附图中，类似的附图标记被用来表示类似的单元。参照图 1，图 1 示出了包括射频通信单元 2 的电子设备 1，其中射频通信单元 2 与处理器 3 耦合以便通信。具有屏幕 5 的形式的输入接口和小键盘 6 也与处理器 3 耦合以便通信。此外，照相机输入端口 19 也以与处理器 3 耦合以便通信，其中照相机输入端口 19 与相关的照相机 17 耦合。本领域技术人员可以理解，照相机 17 可以是设备 1 的不可分割的部分，也可以是分离的附件。

处理器 3 包括具有相关只读存储器 (ROM) 12 的编码器/解码器 11，所述只读存储器 12 存储用于对电子设备 1 发送或接收的语音或其它信号进行编码和解码的数据。处理器 3 还包括与编码器/解码器 11 和相关字符只读存储器 (ROM) 14 耦合的微处理器 13。微处理器 13 还耦合到随机存取存储器 (RAM) 4、小键盘 6、屏幕 5、照相机 17 和静态可编程存储器 16。

微处理器 13 的辅助输出被耦合到提醒模块 15，提醒模块 15 通常包含扬声器、振动马达和相关的驱动器。字符只读存储器 14 存储小键盘 6 上输入的、用于对可以被通信单元 2 接收的文本消息进行解码或编码的代码。在本实施例中，字符只读存储器 14 还存储用于微处理器 13 的操作代码 (OC)，和用于实施将照相机 17 捕捉的图象中文本字段类型分类的方法的代码。

码。

射频通信单元 2 由具有公共天线 7 的接收器和发送器合并而成。通信单元 2 具有通过射频放大器 9 连接到天线 7 的收发器 8。收发器 8 也耦合到合并的调制器/解调器 10，调制器/解调器 10 将通信单元 2 连接到处理器 3。对本领域技术人员来说，电子设备 1 显然可以是任何电子设备，包含蜂窝电话、传统电话、膝上型计算机或 PDA。

参照图 2，其中示出了将照相机 17 捕捉的图象中文字字段类型分类的方法 20。方法 20 包含起始步骤 21，其中操作小键盘 6 上的命令功能的用户调用该步骤。在获得图象步骤 22 获得图象，其中用户通常会将照相机 17 指向名片，直到名片基本落在显示正被照相机 17 检测的区域的屏幕 5 的边界内。然而，不需要在屏幕 17 上显示整个名片，而是只需要显示名片中所需的文本部分。一旦用户对屏幕 5 上显示的内容感到满意，便捕捉并在 RAM 4 中存储图象。

处理器 3 接着执行旋转步骤 23(如果需要)，其中旋转图象以便基本消除与文本相关的偏斜。通过角度检测算法计算基准轴(水平轴)和文本区域中的文本的延伸轴之间的角度，来确定偏斜。美国专利 6,178,270 中描述了确定偏斜角度的算法，该专利被引入本申请作为参考。

在处理器 3 执行旋转步骤 23(如果需要)之后，处理器 3 控制方法 20 完成识别 24，识别图象的至少一个文本区域，该文本区域包括文本。识别首先将二进制化图象投影到水平和垂直两个方向上，接着使用已知的布局分析技术分析两个方向的投影分布，从而从图象背景中限定(区分)出每个文本区域。

该方法接着完成测试 25，以便通过检查二进制化图象中黑和白区域之间的可分离比值来检查图象质量。该测试确定黑区域是字符的可能性。可能性是可分离比。较大的比值表明更好的图象质量。如果可分离比小于一个阈值，该阈值通常是通过只包含文本区域的训练样本图象事先估测和计算的，则对于良好质量的图象，可分离比通常应当至少为 0.8。否则，需要进行二进制化改进 26。二进制化改进 26 仅限于识别至少一个文本区域的步骤 24 所检测的纯文本区域。正如本领域技术人员所清楚的那样，使

用仅限于识别的文本区域内的数据重新计算二进制化判决阈值，以避免非文本区域的影响。

在二进制化改进 26 之后，或者如果测试 25 确定图象具有足够好的质量，方法 20 继续到对至少一个文本区域中的文本进行字符识别 27，以提供输出字符数据。包含文本行或字行的每个图象片段被逐一输入到专用光学字符识别器中，并且被转换成成相应的字符。

接着，方法 20 进行分类 28A，使用一组规则将各个文本区域分类成相应的字段类型。这些规则基于进行字段分类的关键字匹配，例如：当关键字 "Address"、"street"、"st."、"Avenue" 出现时，该行或区域可能是地址信息所位于的文本区域。为了识别电话号码字段，识别字 "Tel" 或 "Telephone" 或 "+"，并且电话号码应当直接跟在后面。对于电子邮件地址，识别 "@"、"email"、".com"、".UK" 等等。为了识别人名，识别其头衔 (例如：Dr. Mr. Mrs. Miss. Ms.)，并且其姓名应当直接跟在后面。对于公司、企业或商号，使用名称识别关键字，例如 "Inc"、"Pty"、"Pte"、"Ltd"、"Limited" 和 "Partners"。

接着，该方法执行存储 28B，在指示字段类型的位置中存储 输出字符数据，其中输出字符数据被存储在静态存储器 16 中。输出字符数据通常被存储在地址簿字段中，并且字段和方法可以通过输出字符数据增加地址簿字段，例如：电话号码字段；人名字段；街区地址字段；或公司名称字段。

方法 20 接着在结束步骤 29 终止，并且用户可以再次启动小键盘以便获得另一个名片的输出字符数据。

参照图 3，其中示出了更加详细地描述图象获得处理 22 的方法。该过程通过照相机 17 捕捉图象，并且将图象存储在 RAM 4 中。接着进行测试 35 以确定 RAM 4 中存储的的图象的分辨率是否高于阈值。通常，用每英寸点 (DPI) 来表示图象的分辨率，并且将阈值适当设置成 200 DPI 的分辨率。如果测试确定图象高于阈值，则完成测试 37，还可以进行分辨率增强 36，并且增强图象被存储在 RAM 4 中。可以通过图象插值方法进行图象增强。因此，如果放大初始图象，则初始图象上的每个象素会被映射成放大图象

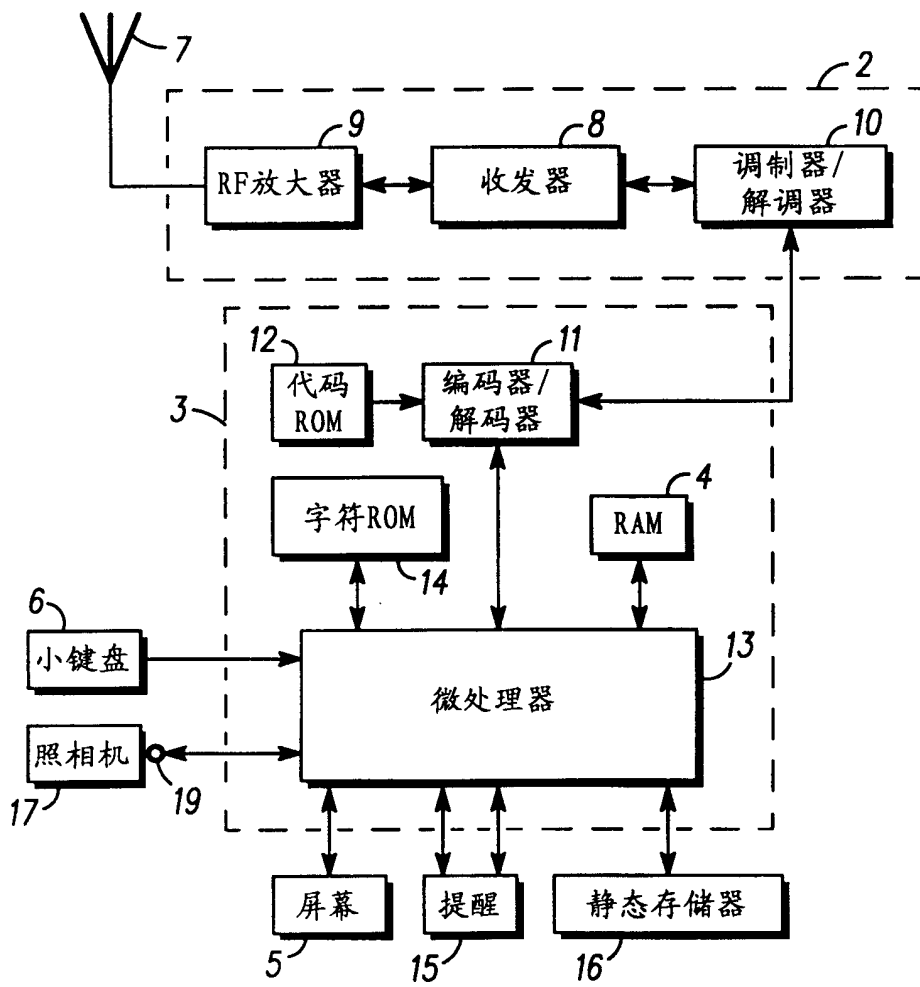


上的若干象素。根据初始图象中象素的相邻象素计算放大图象中映射的那些象素的图象值。在" T.M. Lehmann, C. Gonner, K Spitzer 的 Survey: interpolation methods in medical image processing, Medical Imaging, IEEE Transactions on, Volume: 18 Issue: 11, Nov 1999 Page ( s ): 1049 - 1075 "中描述了图象插值。该文件被引入本申请作为参考。

在测试 37, 检查图象或增强图象以确定其是黑白还是灰度图象。只在图象或增强图象被确定为灰度格式时, 才对图象或增强图象执行二进制化处理 28, 并且接着处理 22 结束。

本发明能提供方便的方法和设备, 这种方法和设备通过对名片照相以获得图象, 从而获得数据, 并且在地址簿字段中存储该数据。本发明处理图象, 图象上的字段被分类, 并且通过从图象中获得的数据更新地址簿字段, 例如: 电话号码字段; 人名字段; 街区地址字段; 或公司名称字段。

前面的详细描述只提供了优选实施例, 不对本发明的范围、适用范围或构造产生任何限制。前面对优选实施例的详细描述只是为了使本领域技术人员能够实现本发明的优选实施例。应当理解, 在不偏离本发明权利要求限定的宗旨和范围的前提下, 可以在单元的功能和布局上进行各种改变。



1

图1

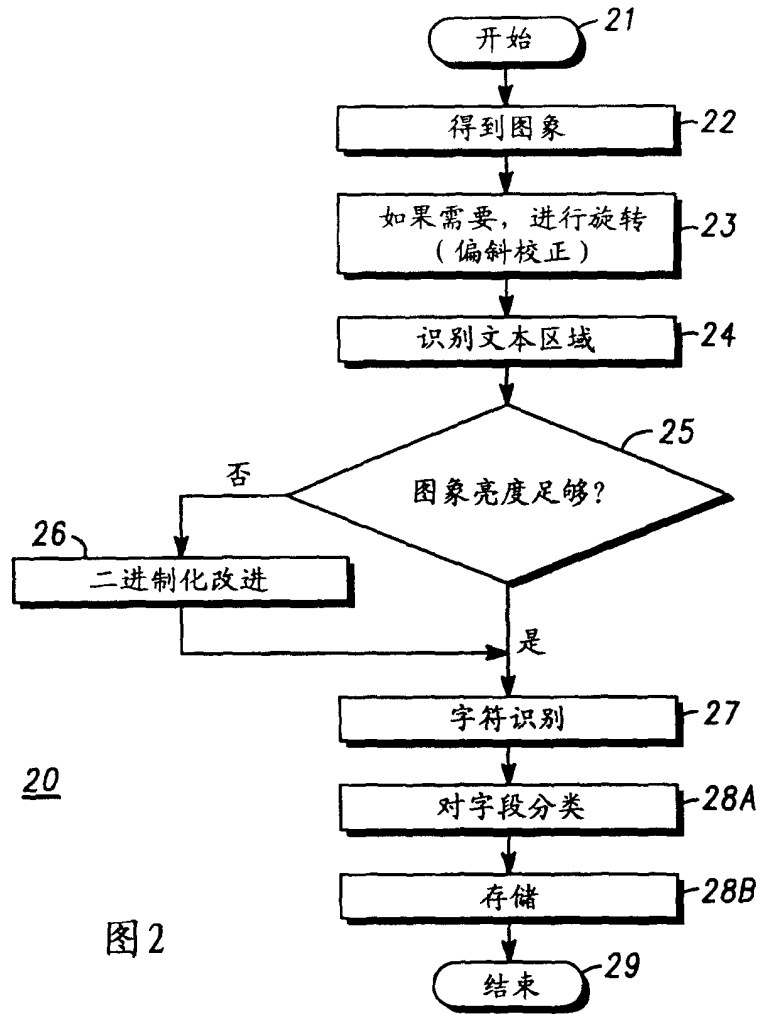


图2

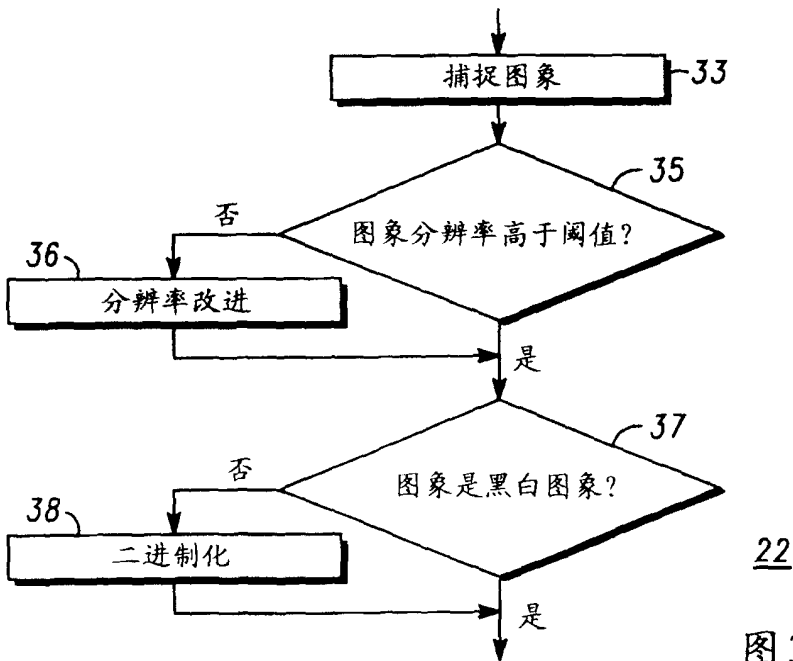


图3