



# [12] 发明专利申请公开说明书

[21]申请号 94119354.3

[51]Int.Cl<sup>6</sup>

H04N 1/41

[43]公开日 1995年11月29日

[22]申请日 94.12.6

[30]优先权

[32]93.12.9 [33]US[31]164,175

[71]申请人 美国电报电话公司

地址 美国纽约

[72]发明人 汉斯·P·格拉夫

丹尼尔·J·梅杰

[74]专利代理机构 中国国际贸易促进委员会专利商  
标事务所

代理人 姜 华

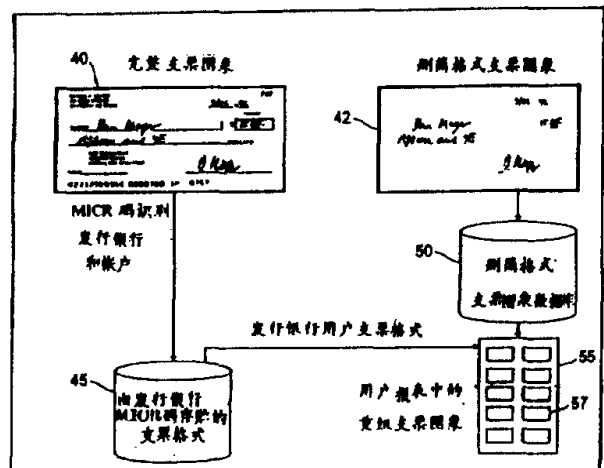
说明书页数:

附图页数:

[54]发明名称 副简格式文件图像压缩

[57]摘要

用于压缩金融证券和其他文件的图象的方法和装置。包括：对多个文件进行扫描以得到每一文件的电子图象；在图象中定位和读出文件识别符，以识别在多个文件中基本维持不变的信息的电子图象的静态部分；识别通常包含每一文件电子图象中多不相同的信息的动态部分，从中分离出仅含动态部分的动态图象；将其存入数据库。本发明为识别和分离文件中的动态信息，例如支票上的手写体正文，提供了有效的方法。



(BJ)第 1456 号

# 权 利 要 求 书

---

1. 用于压缩多个文件的图象的方法,包括下列步骤:

对所述多个文件进行扫描以取得每一所述文件的电子图象;

通过定位和读出所述图象中的文件识别符来识别每一所述多个文件的所述电子图象中的静态部分,所述静态部分包含有在所述多个文件中均基本保持不变的信息;

将所述文件识别符存贮到一数据库中;

识别所述电子图象中的动态部分,所述动态部分包含有在每一所述文件可能各不相同的信息;

将所述图象中的所述动态部分与所述静态部分分隔开,以取得基本上仅包含所述动态部分的动态图象;和

将所述动态图象存入数据库。

2. 权利要求 1 中的方法,进一步还包括下列步骤:

从所述数据库中检索所述文件识别符;

根据所述文件识别符产生所述文件的所述静态部分的图象;

从所述数据库中检索所述动态图象;和

通过将所述动态图象叠加在所述静态部分的所述图象上来重构所述文件的所述电子图象。

3. 权利要求 1 中的方法,其中所述识别所述动态部分的步骤

还包括下列步骤：

从所述文件的所述电子图象中生成至少一种彩色图象；

对所述彩色图象作均衡处理，以改善所述彩色图象的对比度；

和

识别包含有所希望特征的所述彩色图象的动态部分。

4. 权利要求3中的方法，其中所述动态部分包括所述文件上的手写体正文。

5. 权利要求4中的方法，其中所述识别包含有所希望特征的所述彩色图象的部分的步骤还包括下列步骤：

识别包含有笔划和边线的所述图象的部分；

测量所述笔划的粗细；和

识别包含有带有所述经测量的粗细笔划的图象的部分。

6. 权利要求3中的方法，其中所述从所述文件的所述电子图象中生成至少一种彩色图象的步骤包括从所述电子图象中生成红、绿和蓝色图象。

7. 权利要求3中的方法，其中所述作均衡处理以改善所述对比度的步骤包括对所述彩色图象进行直方图均衡处理。

8. 权利要求3中的方法，其中所述识别包含所需特征的所述彩色文件的部分的步骤包括以一神经网络中的多个影响函数作卷积来提取所述特征。

9. 权利要求5中的方法，其中所述测量所述笔划粗细的步骤

包括对从所述彩色图象中产生的特征映象进行交叉相关。

10. 权利要求 1 中的方法,其中所述将所述图象的所述动态部分同所述静态部分相分离的步骤进一步包括下列步骤:

确定所述动态信息域的色谱范围;和

滤除所述图象的处于所述色谱范围以外的成分。

11. 权利要求 10 中的方法,还包括对所述图象进行相关连成份的分析,以分离出所述图象的所述动态部分中的信息的步骤。

12. 权利要求 10 中的方法,还包括应用对应于所述特征的卷积影响函数在一神经网络中求取所述图象的所述动态部分的卷积,以分离出所述图象的所述动态部分中的信息。

13. 用于生成多个文件的压缩图象的方法,包括下列步骤:

识别每一所述文件中的静态部分,所述静态部分包含有在所述多个文件中基本保持不变的信息;

识别每一所述文件中的动态部分;和

将所述动态部分与所述静态部分相分离,以便能将动态部分分别地存贮,而不需要对每一所述多个文件均再存入所述静态部分。

14. 权利要求 13 中的方法,其中所述识别所述动态部分的步骤,还包括下列步骤:

从所述文件的所述电子图象中生成至少一个彩色图象;

对所述彩色图象作均衡处理以改善所述彩色图象的对比度;

和

识别包含所希望特征的所述彩色图象的动态部分。

15. 权利要求 14 中的方法, 其中所述识别所述彩色图象的动态部分的步骤, 还包括下列步骤:

识别包含笔划和边缘的所述图象的部分;

测量所述笔划的粗细; 和

识别包含具有所述测量的粗细的笔划的所述图象的部分。

16. 权利要求 14 中的方法, 其中所述识别所述彩色图象的部分步骤包括以一神经网络中多个影响函数作卷积处理来提取特征。

17. 权利要求 13 中的方法, 其中所述将所述图象的所述动态部分同所述静态部分相分离的步骤还包括下列步骤:

确定所述动态信息域的色谱范围; 和

滤除处于所述色谱范围之外的所述彩色图象的成份。

18. 一种用于生成多个文件的压缩图象的装置, 包括:

用于识别每一所述文件中的静态部分的装置, 所述静态部分包含在所述多个文件中基本保持不变的信息;

用于识别每一所述文件中的动态部分的装置; 和

用于将所述动态部分与所述静态部分分离的装置, 可将所述动态部分分别地存贮, 而不必为每一所述多个文件均再存入所述静态部分。

19. 权利要求 18 的装置, 其中所述文件为支票, 而所述静态部分是预先印制的支票格式。

20. 权利要求 18 的装置, 其中所述静态部分包括所述支票上的 MICR 行。

21. 权利要求 18 中的装置, 其中一第一数据库存贮多个由所述静态部分中的识别符所识别的预先印制的支票格式。

22. 权利要求 18 的装置, 其中一第一数据库存贮多个文件的动态部分。

23. 权利要求 18 的装置, 还包括有为重构所述文件之一的所述图象而将所述动态部分与至少一个所述静态部分重新组合的装置。

24. 权利要求 18 的装置, 其中所述识别所述动态部分的装置包括:

从所述文件的所述电子图象中生成至少一种彩色图象的装置;

对所述彩色图象作均衡处理以改善所述彩色图象的对比度的装置; 和

识别包含所希望特征的所述彩色图象的动态部分的装置。

25. 权利要求 24 的装置, 其中所述动态部分包括所述文件上的手写正文。

26. 权利要求 24 的装置, 其中所述识别所述彩色图象动态部

分的装置包括有：

识别包含笔划和边线的所述图象的部分的装置；

测量所述笔划的粗细的装置，和

识别包含具有所述所测量的粗细的所述图象的部分的装置。

27. 权利要求 24 的装置，其中所述从所述文件的所述电子图象中生成至少一个彩色图象的装置包括有一用于产生红、绿和兰色图象的彩色扫描器。

28. 权利要求 24 的装置，其中所述识别所述彩色图象的部分的装置包括一用于由对多个影响函数作卷积来提取所述特征的神经网络。

29. 权利要求 18 的装置，其中所述用于将所述图象的所述动态部分与所述静态部分分开的装置还包括有：

确定所述动态信息域的色谱范围的装置；和

滤除处于所述色谱范围以外的所述彩色图象的成份的装置。

30. 权利要求 29 的装置，还包括有对所述图象进行相关成份分析的单元，用以分离所述图象的所述动态部分中的信息。

# 说 明 书

---

## 删简格式文件图象压缩

本发明一般涉及文件存贮、检索和传送的改进,更具体地,是涉及对文件的压缩电子图象的改进。

经常总需要存贮多种文件,例如在银行中的支票和其他金融票据,或者在政府机关办公室中的公文、证书、和法庭记录等等。在许多这样大量应用的场合,均是存贮实际的文件。但是,按实际文件的存贮通常引入庞大的存贮设备和人力消耗。因为每一文件均必须如实地被传送到或被安排在一特定的文件库中,所以存贮和检索过程很缓慢,而且常常会使文件不能正确归档或错位。

解决这种存贮实际文件问题的办法之一是存贮这些文件的电子图象而不是实际的文件。目前这一方案通常要对整个文件进行扫描来产生一电子图象,而后再将其存贮一数据库存贮器的一个单元之中。对此电子图象的检索在实际存贮设施中所需时间的一小部分内即可完成。此外,可通过数据传输链路很容易地将文件图象的副本传送至其他的单元位置,而无需真正地将文件由其文档中取出(例如说)加以复印。不过这一方案有一个大问题,就是电子图象在信息位上的大小。一般说来,电子图象均包含有大量的信息位,因此而要求很大的存贮和传输容量。即使采用公知的编码技术来压缩图象信息



位数,例如说用于传真机的 *CCITT Group-3* 和 *Group-4* 图象编码标准,整个图象的大小仍然过大。例如,假定一传真机扫描分辨率为 100~200 点/英寸,则一单元隔行打印正文的电子图象在一维 *Group-3* 格式中将含有约 4000 字节(即 320,000 位);或者在 *Group-4* 二维格式中,将含有约 30,000 字节(即 240,000 位)。在许多大文件量的应用中,为存贮或传送电子图象所需的附加设施容量的费用可能大大超过与实际存贮和检索文件有关联的费用。

当前可用于降低图象大小的技术包括有在生成电子图象之前屏蔽掉文件的某些部分。例如,在对文件进行扫描以生成图象时可以使扫描装置避开这些被屏蔽的部分。换句话说,可以利用一种编码模式,例如 *Group-3* 或 *Group-4* 编码技术,来使得含有大量相同象素的被屏蔽部分被压缩进小量的存贮器。然而这一方案有许多缺点,例如说,被屏蔽的部分一般都必定处于预定的不随文件改变的单元位置中。因此,进行屏蔽并不减少对那些其中所希望信息的存放单元位置是未知的文件中的存贮器的需求量。而且,所有未屏蔽部分中的信息通常均被加以存储,即便它们可能是不必要的背景细节,例如支票上的图案或示例。背景还可能干扰或使得需要的信息模糊不清。

*Chevion* 等人题为“采用大量符号匹配的压缩和还原打印格式的方法”的美国专利 No. 5,182,656(后面简称为“*Chevion*”)中揭示了另一种文件压缩技术。此 *Chevion* 技术包括产生一空的表格和一经填写表格的电子图象。从经填写的表格图象中减去空表格 图象未取得被加到该表中的信息的压缩图象。但是,*Cheion* 方法要求在

相减操作之前在空表格的和经填写表格的图象之间进行精确的重合亦即对准操作,因而增大计算工作量。可参见 *Chevion* 第 5 栏第 39—48 行。另外, *Chevion* 方法亦很不适用于辨认例如在表格中重叠或模糊的手写信息,这种情况在例如银行支票等的文件中是经常存在的。

现有的图象压缩技术一般均不进行被压缩图象不同的部分之间的区分。因此,如果希望进行附加的图象处理,例如自动字符识别,就必须首先对压缩图象加以分段,将适于作字符识别的部分从含有与图形或表格相关信息的部分中区分开。现有的文件压缩技术存在的其他问题,例如在 *M. Kamel* 等人的“由灰度文件图象提取二进制字符/图形图象 (*Extraction of Binary Character/Graphics Images from Grayscale Document Images*)” (*CVGIP: 图形换块和图象处理 (Graphical Models and Image Processing) Vol. 55, No. 3, pp. 203 ~ 217. May 1993*) 一文中有描述,此文被引用为本说明书的参数文件。

由上述可明显看到,需要一个有效的文件图象压缩方法,能够减少对存贮和传输容量方面的要求,并避免现有技术中的信息定位和隔离方面的问题。

本发明提出一种分割和压缩,例如支票和其他金融票据等文件的图象的方法。按照本发明的一个方面,提供了一种方法,其包括有下列步骤:对多个文件进行扫描以取得每一文件的电子图象;对每一图象定位和读取一文件识别符来标明每一文件的电子图象中的那些包含用于多个文件的基本上保持不变的信息的静态部分;将此文件

识别符存入一数据库；识别每一电子图象中的通常包含用于一文件的独特信息的动态部分；将此动态部分与图象中的静态部分相分离，以获得仅仅包含有动态部分的动态图象；以及，将此动态图象存入数据库。本发明为识别和分离一文件中的某些特定类型的动态信息，如支票中的手写体文本等，提供了行之有效的方法。

按照本发明的另一方面，通过检索数据库里的文件识别符，可以重建被存贮的文件图象，并产生对应于该识别符的文件的静态部分的图象。然后从数据库中检索动态图象，再通过将动态图象叠加在静态图象之上重建原来的电子图象。

本发明的一个特征是减少了文件电子图象的大小，使得在一个给定容量的存贮器中可以存贮较大量的文件。例如，对于一普通银行支票的未被屏蔽的电子图象，其图象大小可以被减少到采用现有技术产生的全图象大小的 16% 左右。

本发明的另一个特征是避免了与现有图象压缩技术有关的一些问题。例如，利用本发明的压缩技术，含有手写正文的文件例如支票的签字行等动态部分可以容易被识别，分离、并被存入减少了的空间的存贮器之中。不需要预先确定手写正文的位置，而且不再需要例如屏蔽或精确对准空白格或图象的技术来减少图象尺寸。此外，可以有效地从底线格式背景中分离出手写内容，使得格式不至模糊手写正文，因而不需要重复存贮、检索、或传送不需要的信息。

作为本发明的一个另外特征是，被分割开的动态信息能被用来作后续的例如字符自动识别的处理，而不需要额外的处理步骤去分离出图象的静态部分。

通过对照下面详细说明及附图，本发明的上特征以及其他特点及优点将更容易被明了。

图 1 所示和具有静态及动态部分的银行支票格式的示例性文件；

图 2 说明本发明在银行支票存贮应用中的分割及压缩技术；

图 3 是根据本发明的图象分割及压缩方法中的示例性步骤的流程图；

图 4 是按照本发明的在一文件图象中识别手写正文的示例性工作流程图；

图 5 是按照本发明的在一文件图象中分离手写正文的示例性工作流程图。

本发明提供了用于分割和压缩文件电子图象的方法和装置，以便减小图象尺寸及所需要的存贮器空量或传送容量。虽然下面基本上是对支票的分割及压缩电子图象来说明的，但应该明白，本发明可以更广泛地使用于多种多样的其他相关文件的应用。例如本发明可以用于如 X 射线及核磁共振 (NMR) 图象等医疗文件的压缩电子图象。因此，这里所使用的“文件”这一名词不仅包括例如表格、金融证券及类似这些纸上的文件，而且更广泛地包括可以以电子图象的形式被存贮或处理的任何型式的信息。

图 1 所示为一普通银行支票 10 的示例性文件，根据本发明，该文件可以作为一删简格式图象加以有效地存贮。支票 10 包括静态及动态部分。例如，这里使用的预先印刷好的文件格式的静态部分对许多文件而言是基本上保持不变的部分，而动态部分则是例如被加

入的手写内容是随着文件的不同而有所区别的部分。支票 10 的静态部分包括帐户姓名及地址 11, 收款人栏 12, 支付数量栏 13, 发行银行名称及地址 14, 摘要栏 15, 磁印迹字符识别(MICR)栏 16, 支票号码 17, 日期栏 18, 常规信息 19, 法定总数窗框 20, 及填票者签字栏 21。MICR 行 16 通常包括辨认发行银行、帐户号码及支票号码的编码。静态部分也可以包括预先印刷好的背景结构或图案, 及附加的预先印刷好的常规的或识别信息。这里就将一张实际支票的静态部分称这为支票格式。

支票 10 的动态中分包括若干动态字段, 例如收款人 22, 法定总数 24, 日期 26, 支付数量 28 及填票人签字 30 等。每次使用一张支票时, 填票人通常用手写文体填入这些动态信息域的每一项。另一方面, 这些动态字段也可以包含例如采用打字机, 计算机打印机, 或签字机器输入的正文。应该注意到, 支票号码 17 及包括支票号码在内的 MICR 行 16 的部分通常随着不同的支票而不同。然而, 由于这一信息是预先在印制每张支票上的标识符的部分, 在这里, 将作为支票上的静态部分加以考虑。根据本发明, MICR 码将被用作文件标志符以识别一个特定的预先印制好的包括支票号码在内的支票格式的静态部分。

根据本发明, 支票 10 的电子图象可以被分割静态部分及动态部分, 而动态部分可以作为一个动态的或删简格式的图象而被存贮。将动态部分进行分割, 或进行识别及分离的方式将在下面更详细地讨论。如上所示的, MICR 码将指明在从被存贮的动态图象中再象一原始图象时, 应该使用许多不同支票格式中的哪一种。对于每一支

票均要存贮 MICR 码和动态部分,但静态部分的图象则只存贮一次。本发明在识别和分离通常包括手写正文的动态部分方面提供的优点,使得为每一张支票必须所存贮的图象信息的总量大大减少。正如前面提到的,现有的图象存贮技术,典型地是存贮包括静态和动态部分两者的整个支票的一幅图象。一张被压缩的整张支票电子图象,通常在  $300 \times 300$  点/英寸扫描分辨率时,大约为 320,000 到 400,000 信息位,前面采用黑白或二进制位图象,背面采用 4 位灰度等级图象。因此,例如为对一给定的银行的每一帐户的所有被注销的支票存贮电子图象,就必须要有相当大的存贮容量的存贮器。此外,使用现有技术去检索和/或传送每张支票图象通常还要花费更长的时间。

图 2 是说明根据本明存贮及检索图 1 的示例支票的图象方框图。在这个示例的实施例中,虽然只示出支票 10 正面的图象,但很明白,本发明的技术也可以用来识别及分离该支票背面的动态信息。可以以本技术中公知的技术扫描支票 10,产生一张包括在上面标明的静态及动态部分的完整原始支票图象 40。该电子图象可以例如采用 Group-3 或 Group-4 传真编码加以编码,此二编码技术是以对多个扫描行的图象象素的变化作图象信息编码的。参见 CCITT Fascicle VII. 3, 规程 T. 4 及 T. 6, 本说明书在此引入此二规程作为参考,以便更进一步分别了解 Group-3 及 Group-4 传真编码技术。其他的编码方案,如由 CCITT/ISO Joint Bi-Level Image Experts Group (JBIG) 开发的图案编码算法等,也可以加以采用。然后可以定位和读出电子图象中的 MICR 栏 16, 以便识别发行银行、

帐户号码及支票号码。通常帐户号码,也可以确定填票者的姓名及地址。*MICR* 引通常在每张支票图象的同一位置出现,所以采用本技术领域中公知方式的磁墨读出机能够容易地将其分出及读出。另外,可以采用例如光学字符识别(*OCR*)方法,在产生支票的电子图象之前读出 *MICR* 码。在这两种情况下,*MICR* 码均可以用于识别一张特定的预先印好的支票格式,因而可以作为本实施方案中的一种文件识别符。应该明白,其他一些型的文件识别符也可以被用来识别在其他应用中文件的静态部分。文件识别符及其相应格式的拷贝均可以存入数据库 45。

如果应用 *MICR* 栏中帐户号码不能方便地确定帐户姓名及地址的话,那么可以将原始图象加以分割来决定这些信息。帐户姓名及地址能够通过例如对通常能找到该信息的部分如左上角加以分析,而在图象中加以定位。因为帐户号码及地址通常是以一个有限数目的字根印刷的,因而这个信息可以通过公知的 *OCR* 技术将其读出。另外,可以使用一种特征提取技术去定位姓名及地址信息。因为姓名及地址通常使用单色墨印刷,所以特征提取技术在银行支票的情况下是被简化的。此特征提取技术可以以下面将更详细地加入说明的方式利用例如集成电呼神经网络去完成。*MICR* 行、帐户姓名及地址、以及支票图象上任何其他有关被预先印刷好的部分都可以例如作为被压缩文本有效地加以存贮。

根据本发明,识别出原支票图象 40 的手写体动态部分,然后将其与预先印刷好的静态部分相分离,从而产生一张示例的被分割及压缩的动态图象 42。这种动态图象 42 代表了一种删简格式图象,其

包括加到支票上的全部手写内容,但没有预先印刷好的信息。从支票图象 40 中抽取的 *MICR* 行编码被送进发行银行中的数据库 45。发行银行可以把在该银行为每张支票帐户预先印刷好的支票格式图象保留在数据库 45 中。从而发行银行能通过利用 *MICR* 码访问数据库 45,确定一给定支票的正确格式。发行银行将用于每一支票的动态图象 42 存入支票图象数据库 50,其中每一图象均由 *MICR* 码的一部分例如支票号码及帐户号码予以识别。虽然图 2 所示的支票格式数据库 45 和支票图象数据库 50 是分开的数据库,但可以采用一个单独的数据库来同时存贮静态及动态图象。

在图 2 的示例实施例中,发行银行利用存贮在格式数据库 45 中的静态图象和存贮在图象数据库 50 中的动态图象来产生一定期的用户财务报表 55,它包括某一段时间内所处理的支票的再现图象 57。发行银行通常将被注销的支票或者被注销支票的缩微胶片图象存贮和保留一定的时期。为了降低花费,许多银行不再给填票者返还被注销的支票。但是如果,例如,在填票者与受款人之间发生争议时,通常填票者必须与银行取得联系以得到一张被注销支票的复印件。如果发行银行能提供如图 2 中所示的具有被再现的支票图象 57 的用户财务报表 55 的话,那么填票者就能够避免这种问题,从而能更准确地结算和管理帐目。

可以通过根据表 55 参数据库 45 的检索,印制一份恰当的静态图象,或支票格式,来再现被注销的支票。数据库 45 中的支票格式可以通过被存贮的 *MICR* 行而被识别,因此每张被印制的格式将包括 *MICR* 行以及在该报表期间被处理过支票的其中一张的支票号



码。被印制的格式也可以包括如同上面提到的帐户姓名及地址。另外,姓名及地址的信息段可以从原来图象的其余部分中被分割出来,采用标准的印刷字符识别技术,将其转换为 ASCII 码,利用文件识别符例如发行银行 MICR 码将其存入数据库 45 中,然后再将其印制到报表 55 上的每一张格式中。MICR 码,或者更具体地说支票号码,可被用来辨认哪些支票已经为一给定帐号在此期间处理过,因此动态图象就可以按支票及帐户号来存进数据库 50。而后每一被处理过的支票的手写部分就可按支票号码所识别的印制到报表 55 上的恰当支票格式上。这样,用户财务报表 55 就可能包括有许多注销了的或者经过处理的支票的再现图象 57。

数据库 45 并不一定要包括预先印制格式的图象。相反,银行有可能采用对不同类型帐户通用的格式。这样数据库 45 就只需存放识别帐户号和支票号的 MICR 码。发行银行将带有对每一处理过的支票的合适的 MICR 行和支票号的通用格式印制到报表 55 上。发行银行可应用多种不同的通用格式,而根据帐户号来针对一特定用户指定恰当的通用格式。作为另一种方案,可以不将数据库 45、50 保留在任一特定的发行银行中,而代之以保留于一为数个发行银行共用的支票处理中心。那么 MICR 码就可被存贮在数据库 45 中,并被用于与帐户和支票号码同时识别一特定的发行银行,和该发行银行的通用的或所存贮的格式。此支票处理中心可在访问数据库 45 和 50 后再现支票图象,以生成一可提供给此发行银行或者一银行用户的财务报表。

本发明的删简格式图象压缩还为通过通讯网络传送支票图象提

供了便利。例如象前面说过的那样,原始支票图象 40 可被分成静态和动态部分。识别发行银行、帐户号和支票号的 MICR 码可被送往通讯网络,该网络例如可以是一电话或数据网络。动态部分 42 也被送到此通讯网络。这种将支票图象分为静态和动态部分的图象分割和压缩可以在例如一访问此网络的发行银行中进行。然后,该网络将一特定支票的 MICR 码和动态图象从此发行银行传送到一所希望的目的地,例如,填票者的家用计算机。静态图象,亦即格式,可以被存贮在此目的地的数据库中,以使得能通过将动态图象叠压在静态图象上来再现原始支票图象。

一般说来,按照本发明的经分割和压缩的支票图象可以在较原始支票图象为短的时间内通过一给定的数据信道传送,因而需要的传输容量较小。静态部分仅只传送一次,然后被存放在目的地数据库中,并被用来与同其分别传送的动态图象一起来再现多个支票图象。被再现的支票图象可有助于实时地通过例如家用计算机对通讯网络的访问来通知填票者一特定支票已经在银行被清结或者处理过。这样,填票者就无需等待用户的月财务报表来确定一特定支票的状态,或得到一被注销的支票的副本。另一方面,动态支票图象可随同相应的 MICR 码一起,在一个对所收到的支票上进行扫描,分割和压缩操作的处理中心和一个保持有一支票格式数据库的银行之间进行传输。这样一带有一组图象处理装置的单台支票处理设备就可以被应用于许多银行。

图 3 是说明按本发明的银行支票删简格式图象压缩过程的示例性流程图。在操作方框 300 中,通过采用例如一个 24 位彩色扫描

仪对支票进行扫描得到支票 10 的全电子图象。可以采用基于例如电荷耦合器件 (CCD) 摄象机或其他类型摄象机的成象技术来产生支票适当电子图象。另外,在动态信息域部分通常处于各文件的相同部位的应用场合中,后续图象处理步骤的速度可以因采用屏蔽图象而得到提高。通过将全支票图象的某些静态部分屏蔽或者删掉,可以产生一个经屏蔽的图象。

支票的 MICR 行,以如上所述的方式被定位在电子图象之中,并且被读出以便识别发行银行、帐户号码及支票号码。在操作方框 304 中,以如上术的方式从支票图象的其他静态信息中分离出帐户姓名及地址 11,并利用市场供应的 OCR 程序加以识别。正如上面所提到的,因为帐户姓名及地址通常是以一个有限数目字根印刷的,所以本食施了例中的 OCR 程序是被简化的。另外,因为每一个帐户号码通常只对属于一个单独的帐户姓名及地址,所以发行银行能够使用 MICR 码中的帐户号码来识别例如数据库 45 中的帐户姓名及地址。所以这里要存贮的预先印制的信息包括发行银行标识字符、帐户号码、支票号码、帐户姓名及地址。这些信息可以采用例如 *h<sub>n</sub>ffman* 编码或 *Lemppel-ziv* 压缩技术作为经压缩的文本加以存贮。为了更详细了解请见例如 *T. Welch* 的“高性能数据压缩技术 (*A Technique for High Performance Data Compression*)”, (*Computer* 17 卷 No. 6, pp. 8—9, 1984 年 6 月)。此文在此被引用为本说明的参考文件。如果一家特定的发行银行,仅想要存贮他本行已经发行的支票那么该银行就不需要存贮银行标识符。

操作方框 306 中,对电子图象中通常包括手写文本的支票上的

动态部分加识别。然后被识别的动态部分被与电子图象的剩余部分相分割,如操作方框 308 所示。方框 306 和 308 所示的动态信域段识别和分割的处理将在下面分别结合图 4 及图 5 进一步加以说明。由此产生的动态图象,可以只包括例如图 2 所示的示例动态图象 42 中的手写文本。如操作方框 310 所指示的,然后再将识别发行银行、帐户号码、支票号码的 MICR 码,被分离的帐户姓名及地址、以及从原支票图象中被分离出来的手写体部分存贮到一个或多个数据库之中。如果能方便地从帐户号码中识别姓名及地址,那么就不需要把它们加以存贮。参考图 2 可以看到,本发明从原来图象 40 中排除了不必要的细节内容,这样就使所得到的动态图象 42 所需要的存贮器容量是大大减小。

图 4 示出一组对应于图 3 方框 306 的操作示例,它适用于识别原支票图象 40 中的手写文本。最初在扫描文件时,可产生多种不同彩色版本的原始图象 40。在本实施例方案中,采用红、绿、兰三种不同的颜色图象去识别手写文本所使用的一种墨色的色谱范围,对此下面将更详细地加以说明。可以使用例如一种彩色扫描器首先扫描该文件,以产生原始图象 40 的多种不同彩色版本。彩色扫描器通常产生数幅图象,每一张图象对应于一种颜色,例如红、绿、兰。另一可替换的彩色编三方案是可以利用一幅图象代表亮度而二幅图象表示色度。这后一方案经常被用于对视频图象进行编码。对采用一种具体的彩色信息编码方案并无苛刻要求,只要能够提取文件的彩色信息就行。

而后采用例如直方图均衡处理来增强每一不同彩色样本的对比

度。直方图均衡处理是图象处理领域中所公知的技术，通常包括先产生图象亮度的一直方图，然后对象素值进行换算以使得该直方图成均匀分布的。参看例如 T. Pasvolidis 的“图形和图象处理算法 (Algorithms for Graphics and Image Processing) ” pp50—54 (computerScience Press 出版)，此文在此被引作为参考文件。用于增强图象对比度的另一方法是边缘增强处理。实施边缘增强处理的一种可行措施是，首先采用许多现有的边缘检测装置中任何一个来确定哪些图象象素属于边级部分。然后增加在此边缘一侧的象素值而减少另一侧的象素值，从而增大边缘两侧象素值之差。

根据本发明，为改善此直方图均衡过程，先确定支票上预先印制信息的颜色，然后将该颜色排除在均衡操作以外。通常在支票的某些区域找到的预先印制的文本的彩色，例如在左上角的预先印制的名称和地址，可以由分析这一区域象素值的直方图来确定。例如，可以对被分离的帐户名称和地址的颜色在图 3 的操作方框 304 中加以分析。由于名称和地址通常均以同样的墨色印刷，所以在直方图中它们的颜色将显为一明显的峰顶。从而就能对图象定出一阈值，使得仅有这一颜色成为可见的，从而借此来提取被印刷的文本，而后它们就可以用通常的 OCR 读出。也可以采用另一种技术来确定支票上预印文本的颜色，它包括对在一特定色谱范围内，例如紫外线范围内的预印制文本的反射性进行测量。如果正是以被设计成欲由这种方法来检测的墨水印刷的话，就能很容易地分离出所反射强度的图象。如果支票含有多种颜色的预印制信息，或者如果预印制文本无法方便地进行分析的话，采用全色直方图均衡处理可能取得满意的

结果。

对图象对比度作加强之后,就对每一彩色图象以预定的阈值进行度量,以便生成二进制图象。这里所用的术语“二进制图象”是指双电平黑白图象。每一二进制图象包含支票中符合用于特定颜色的预定阈值的任何特征。对生成每一二进制图象的合适的阈值,可以通过分析图象的直方图来确定。例如,如果直方图显示二个清楚的峰顶,就可以设置阈值来分离此二值。参见例如上面引用作为参考的 *T. Partidis* 一文中 pp65—67。一般常采用更复杂的算法,例如上面引用作为参考的 *M. Kamel* 的论文中所介绍的。

然后通过将所得的二进制图象与一神经网络系统中的一组卷积函数进行比较,以提取特征。一般每一卷积影响函数表示一可与实际图象相比较的或与之作卷积的图象特征的特定排列。这样这些卷积影响函数就可以用通过确定实际图象的哪些部分最与一特定影响函数最吻合,而在一二进制图象中确定具体特征,例如确定在某一取向上的手写体的边线或笔划的位置。借助于选择一组合适的影响函数,就可以提取一组所希望的特征,而其它不需要的特征,例如支票上的背景图案或图形,则被剔除。例如,一组不同的影响函数可被选择来检测出有别于预先印制文件的手写体正文。这种应用一神经网络系统中的影响函数卷积来定位图象特征的过程,在例如 *H. P. Graf* 等的“采用模拟神经网络芯片的图象识别(*Image Recognition with an Analog Neural Net Chip*)” (*Proc. IAPR, Vol. IV, pp. 11—15, IEEE computer Science Press, 1992*)一文中介绍,此文在此作为参考文件。

实际卷积操作可利用一神经网络集成电路 (IC), 如 NET32KIC 来进行。此总片在 H. P. Graf 等人的“利用神经网络硬件作地址单元定位及图象预处理 (Address Block Location and Image preprocessing Using Neural Net Hardware)” (Proceeding of the Advanced Technology Conference, U. S. Postal Service, Vol 3, P. P. A-125~A-135, 1992) 一文中介绍, 此文在此用作为参考。有关神经网络的其他有关细节可参看美国专利 No. 4901271, H. P. Graf 的“计算网络”。美国专利 No. 5093900, H. P. Graf 的“可重组的神经网络”, 以上两专利均已转让给本发明的受让人 (本申请人); 还可参见 H. P. Graaf 等人的“利用可变规换网络作图象分割 (Image Segmentation with Networks of Variable Scales)” (Neural Network Processing, Vol. 4, Morgan Kaufmann Publishers, 1992), 和 H. P. Graf 等人的“可重组的 CMOS 神经网络 (A Reconfigurable COMS Neural Network)” (Digest of the 1990 IEEE International Solid-State Circuits Conference, pp. 144-145, 285, Feb 1990)。上述诸文均在此用参考。神经网络卷积处理的结果是每一二进制图象的一组特征映象, 它们可被用来确定例如在二进制图象这内出现特定取向的手写体行的地点。操作方框 404 中执行的神经网络卷积处理从每一二进制图象中有效地提取特征映象。另外, 一组所希望的特征也可利用一种模仿例如 NET 32K IC 的操作的标准微处理机或数字信号处理机来提取。采用这样一个处理机可以得到同样的结果, 但处理机一般慢得多。例如模仿 NET32K 操作的数字信号处理机较 NET32K IC 约慢 10000 倍。另

外,还可以利用许多市场销售的数字卷积器中的任一种。这些数字卷积器可在大约仅慢 10 至 100 倍的速度上产生与 NET 32K IC 相同的结果。

按照本发明,神经网络特征提取操作可以通过将特征映象生成和分析限定到最可能包含手写体文本的范围来加以简化。在支票的情况下,手写体文本通常将仅仅出现在某些动态区域内,例如那些接近签字栏、受款人栏、支储钱数量栏、法定总数框和日期栏等。因为这些信息段的位置在支票相互间是类同的,所以特征映象就可以只在这些信息段范围内生成和/或得到分析。另外,如上面指示的,可以生成受屏蔽的图象来代替全图象,而特征提取就对整个屏蔽图象进行。在本例子中,整个原始图象被用来产生在最可能包含手写体的区域内分析的特征映象。这些区域可能包括,例如,那些表明为图 1 的支票示例的动态部分。特征映象指明二进制图象的哪些部分含有所希望的特征,由此而提供有效措施以将手写体特征与预先印刷的特征区分开来。在手写体区域未知的应用场合,可对为整个二进制图象生成的特征映象进行分替。

在对含有带所希望的边缘和笔划取向的特征映象的二进制图象部分定位之后,可以测量一特定部分中的笔划粗细,如操作方框 406 中所指示的。一种适用于确定笔划粗细的方法是对沿图象的被识别部分的一条线或截面的特征映象进行自动相关操作。这种操作也可被称之为交叉相关。通常,在进行交叉相关时有两个特征映象相关。所采用的一个特征映象为对应于捩写体部分内的左垂直边缘的边缘映象,而另一个则是对应于右垂直边缘的边缘映象。如果一左边



缘象是一手写笔划的一部分的话，则右边缘一般将代表至该左边缘映象右边的一定的距离或笔划宽度。因此一手写体部分中的左垂直边缘映象与右垂直边缘映象的交叉相关，就将展现为一个对应于手写笔划粗细的量值的峰值。通过手写体部分中的边缘映象的相关处理以及实际测量不同边缘之间的距离，就可得到手写笔划的平均粗细。

对于其他被识别为包含于手写体特征的图象部分可以重复进行笔划粗细测量。然后就可如操作方框 408 中指出的利用测得的手写体笔划粗细即可识别图象手写体部分的一些笔划组。其他的手写特征，例如角，弧线，线交叉，以及线端头等，也可采用特征映象进行识别。虽然利用笔划宽度能较好地将手写体和静态信息分开，但也能够通过直接分析特征映象而不先测量笔划宽度来找到手写体笔划。在此两种情况下，方框 306 的识别过程均被用来以笔划组的形式识别每一包含手写文体的二进制图象部分。然后在图象被识别阅分中的实际手写文本就可与图象的其余部分相分离。

图 5 表示 对应于图 3 方框 308 的适用于按照本发明分离手写文本的一组操作实例。最初对于一个二进制图象测量在上述方框 408 中识别的一组或多组手写体笔划的颜色。一般手写体文本以单色墨水，如兰色或黑色 写入。但是由于通常是从支票图象的数个不同区域提取笔划，因而在笔划颜色上就出现一些变异。这种变异代表了该手写体文本的色谱范围。在操作方框 504 中，此色谱范围被用来通过将所有处于此手写体色谱范围之外的图象成份加以滤除，来分割手写体文本。此滤除操作的执行是通过利用该色谱范围设定

阈值来对每一二进制图象作阈值处理实现的。阈值处理所得的图象主要包括去掉大部分静态信息的手写体文本。如果支票包含有与手写文本相同的颜色、大小、形状、和取向的预先印制的信息,则将可能有必要进行至少一个其它的分离步骤。

操作方框 504 包括一个对经过阈值处理图象进行相关成份分析(本领域的一种公知技术)的分离步骤。相关成份分析通常对由许多象素所组成的图象成份进行识别,其中每一象素均能够通过只移动相同颜色的象素而从该成分内的任何另一象素来达到。对应于手写文本的相关成份可以容易地从其它对应于预印制文本或背景的相关成份中识别和分离出来,后者可能具有与该手写文本类似的颜色和特征。因此相关成份分析就可被用于作经过阈值处理图象中的手写体文本的分离。

还可行额外的特征提取步骤来使得更好地分离手写体文本。例如,上面结合操作方框 404 所讨论的特征提取步骤可以对经过阈值处理的图象重复进行。可选择卷积影响函数来,例如,提取具有在操作方框 406 中所测得的笔划宽度的图象特征。以这种方式,可将图象的任何非手写体部分,例如小污点除去。此辅助特征提取也可以利用,例如,上面提到的 NET32K 神经网络集成电路或任何其他适合作特征提取的硬件或软件来执行。此辅助特征提取步骤可在相关成份分析之后,或者取代后者执行。图 4 方框 400 至 408 和图 5 方框 500 至 508 中的图象处理操作所得结果为一删简格式的支票图象,如图 2 的图象 42,它仅只包含有手写体文本。

虽然上述方法示例仅涉及到单色手写体,但本发明也能容易地

扩展到多色手写体文件。例如,在操作方框 500,可以在测量数组手写笔划的颜色之后确定出存在有数种颜色的手写文本。而后将不同的色谱范围用于每一手写体彩色。应用每一色谱范围进行滤波处理,或对图象作阈值处理,以生成含有一特定颜色的手写体的单独图象。然后在任何相关成份分析或附加的特征提取步骤之前或之后将不同的颜色图象加以组合。而且本发明还可被用来分割较文件背景更深和更淡的手写体。应用卷积影响函数来识别,例如,边缘特征的神经网络卷积过程在此两种情况下都可执行。不过对于在一暗背景上的淡色手写体,右边缘映象与左边缘映象的次序应被反转来,以便决定(例如)笔划宽度。

图 3,4 及 5 的示例流程图描述银行支票的删简格式图象压缩过程。同样的处理步骤也可被用来处理其他类似的文件。其他文件可能包括类似于支票上的 MICR 栏的不同文件类型的文件识别符,它可被分离并读出以识别静态格式。虽然上述的方法特别适宜于用来识别和分离文件中的手写体,但所加入的部分也可能是,例如,印字体,计算机打印体,或盖至文件上的图章印记。不同类型的增加信息可以采用例如修改特征提取中所用的卷积影响函数的措施从原始图象中提取。在上述的实施例中,所选择的函数适应于识别手写文本中的边缘和笔划。

本发明所提供的改进已对数个与图 1 的支票在格式上大本相同的典型的银行支票采用一维 Group-3 和二维 Group-4 电子图象编码进行了模拟。表 1 示出了应用一维 Group-3 所作五个支票图象的删简格式图象压缩的结果。此五个支票中的每一个均具有一

或多种颜色的不同预先印制的信息域,例如,背景图案或图形、包含手写文本的动态信息域以单色墨水写入。表 1 各列表示用于每一支票的测量数据。

表 1

支票#	1	2	3	4	5	平均
全图象(字节)	23,993	40,601	27,151	22,090	27,671	28,301
屏蔽图象(字节)	14,104	24,420	15,442	15,393	13,473	16,566
删简格式图象(字节)	11,298	11,245	7,118	7,663	6,403	8,745
删简格式/全图象(%)	47%	28%	26%	35%	23%	31%
删简格式/屏蔽格式 /屏蔽图象(%)	80%	46%	46%	50%	48%	53%

起初,扫描每一支票以产生分辨率为  $300 \times 300$  点/英寸的二进制或黑的电子图象,然后利用一维 Group-3 传真标准加以编码。另外,在每一空白纸上以手工重现该手写文本以产生每一支票的删简格式图象。经编码的图象采用 *TIFF* (*Tag Image File Format*, 特征图象文件格式) 标准版本 6.0 (可从 *Aldus Corp. Seattle, Washington* 得到) 加以存贮。表 1 中第一行数据包含采用一维 Group-3 标准编码的原始全电子图象中的总字节数。每一字节含 8 位图象信息。第二行的数据指明由屏蔽在支票的某些部分以便仅只包括静态信息时减少的字节数。可以看到,每一支票均可由对支票各不同静态

部分的屏蔽来使得字节数有所减少。但采用按照本发明的删简格式图象能得到大得多的减少。每一删简格式图象中的字节数如第三行的数据所示。例如第五支票的删简格式图象仅含有 6,403 字节。但原始图象有 27,671 字节,而屏蔽图象则含有 13,473 字节。

表 1 中第四行数据表明了用于删简格式图象的占全图象字节的百分比数,而第五行则为用于删简格式图象的占屏蔽图象格式的百分数。可以看出,删简格式图象仅仅需要相应全图象所需字节的 23% 47%,而为对应屏蔽图象所需字节的 46% 至 80%。这种改善随每种支票而不同,主要取决于支票中静态信息的例如背景图案的类型。在此示例中所产生的 5 个删简格式图象,使得平均由全图象的 28,301 字节如屏蔽图象的 16,566 字节减少到 8,745 字节。这就相应于平均约为全图象大小的 31%,或者说在所需的存贮和传输容量上平均减少 69%。

表 2

支票#	1	2	3	4	5	平均值
全图象(字节)	10,050	44,686	21,563	14,142	15,595	21,207
屏蔽图象(字节)	5,953	25,625	12,144	10,934	8,608	12,653
删简格式图象(字节)	3,942	4,813	2,659	2,695	2,577	3,337
删简格式/全图象(%)	39%	11%	12%	19%	17%	16%
删简格式/屏蔽图象(%)	66%	19%	22%	25%	30%	26%

表 2 说明采用二维 *Group-4* 编码的删简格式图象压缩的结果。这些数据是采用的与产生表 1 数据同样的支票、相同的  $300 \times 300$  的扫描分辨率、和同样的图象处理步骤取得的；所不同的是用二维 *Group-4* 标准来代替一维 *Group-3* 传真编码标准。一般说，此删简格式图象字节数的减少大于采用 *Group-3* 编码所得到的，获得的每一删简格式图象从全图象的 21, 207 字节和屏蔽图象的 12, 653 字节平均减少到约 3, 337 字节。因此删简格式图象平均仅需全图象字节的 16%，和屏蔽图象字节的 26%。由表 1 和表 2 的数据可看到，虽然由本发明所得到的改善幅度随特定的支票和所采用的编码标准而异，但在每一种情况下都得到很大的改善。

利用其他在类型的图象编码方案也可以取得同样的改善。例如，对于表 1 和表 2 中第一支票的图象压缩测量，也采用 *JBIG* 图象编码算法来代替 *Group-3* 或 *Group-4* 编码重复进行。所有其他的测量参数不改变。采用 *JBIG* 编码，一号支票的全支票图象含有 7, 593 字节，而屏蔽图象含有 4, 390 字节；对应的删简格式图象仅包含 2, 838 字节，即全 *JBIG* 图象的 37%。虽然上述的示例 *Group-3*、*Group-4* 及 *JBIG* 的结果是用于由手工分割手写信息的删简格式图象的情况，但结果证明采用本发明的分割和压缩方法能取得改善。

应当理解，上述实施例仅仅只是举例。可以在所示的安排中进

行许多改善,包括文件类型文件中静态和动态信息域的数目和型式,以及用于识别和分离各种不同的动态信息域的特定方法。这些以及其他的方案和改型对于熟悉本领域的技术人员来说是很显见的,因此本发明仅仅只为附列的权利要求所限定。

# 说明书附图

1/3

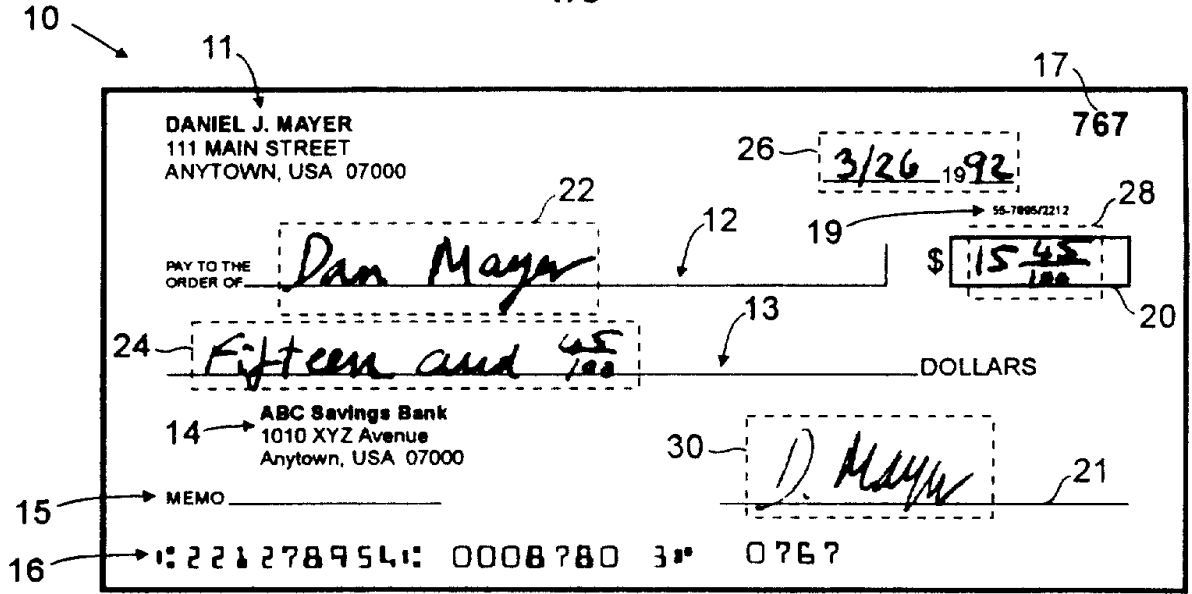


图. 1

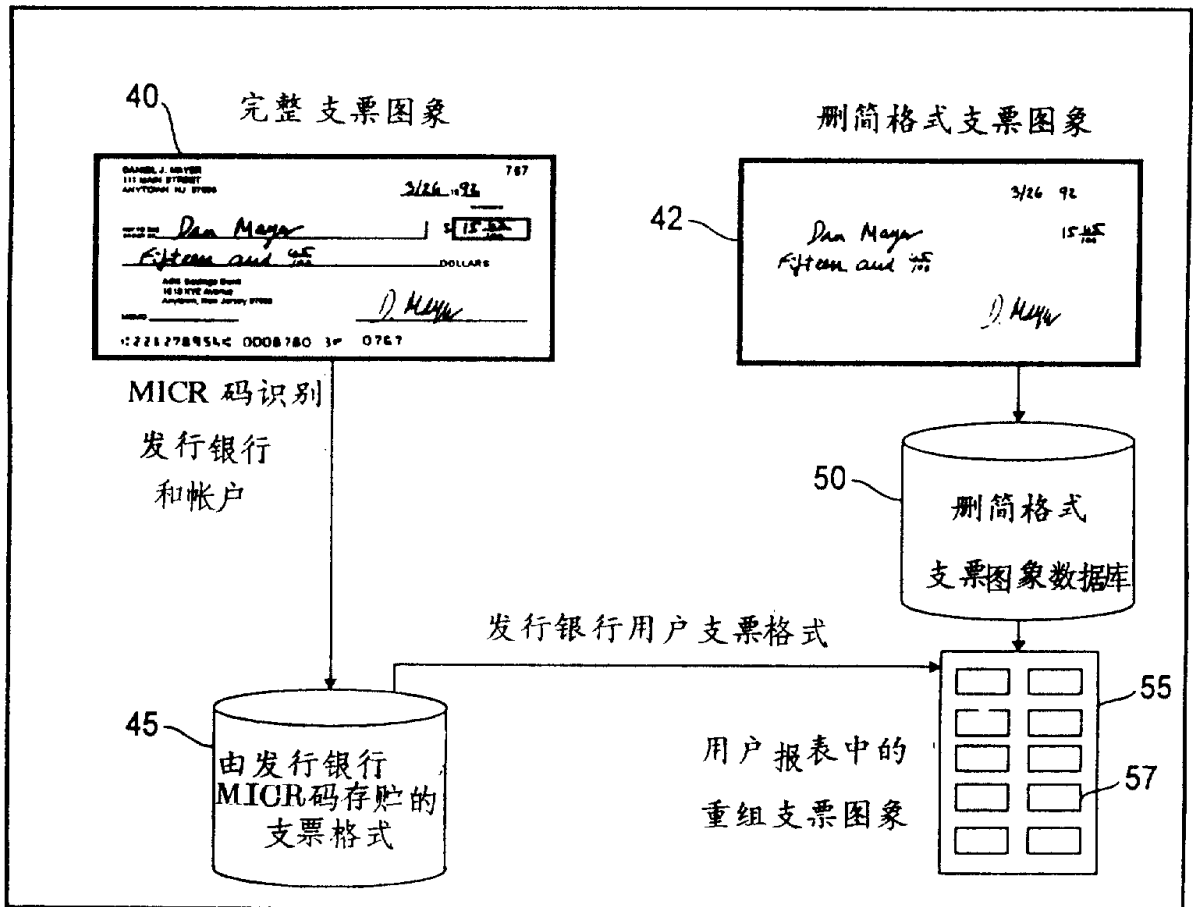


图. 2



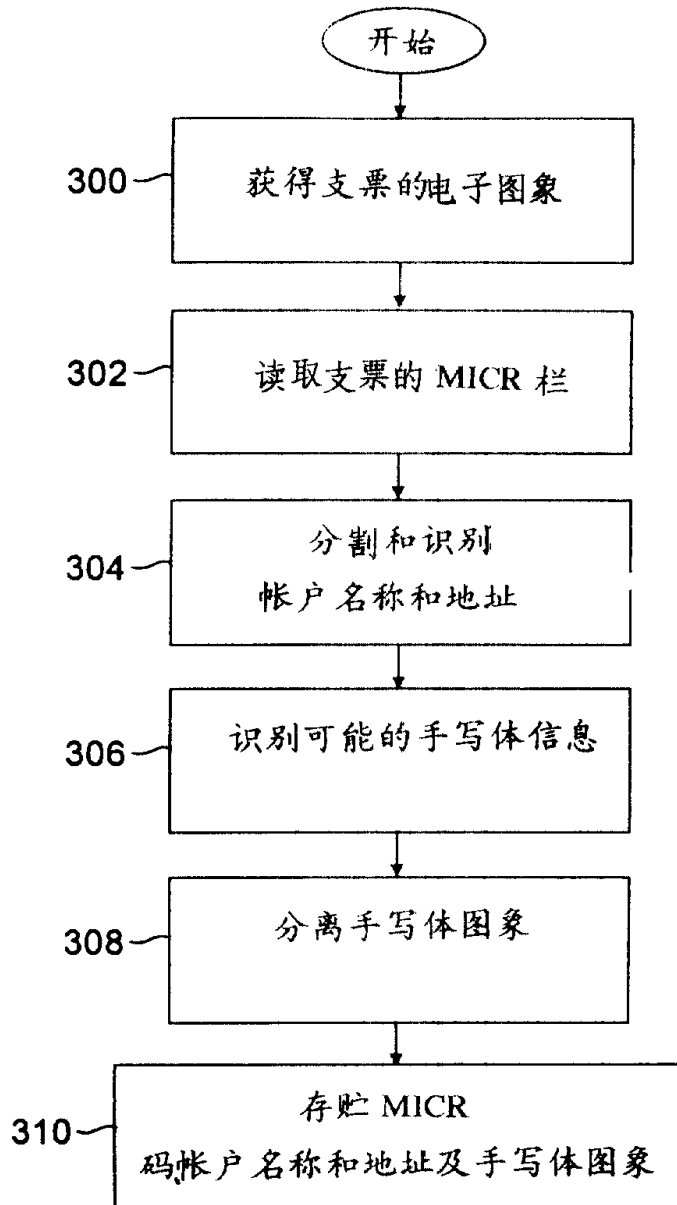


图. 3

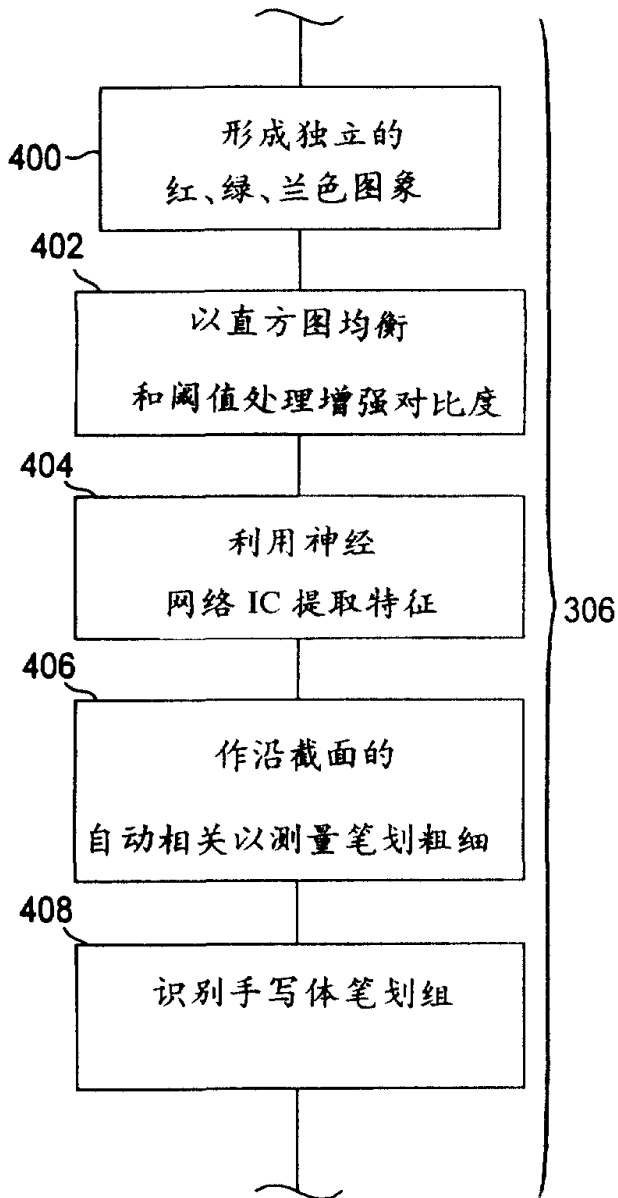


图. 4

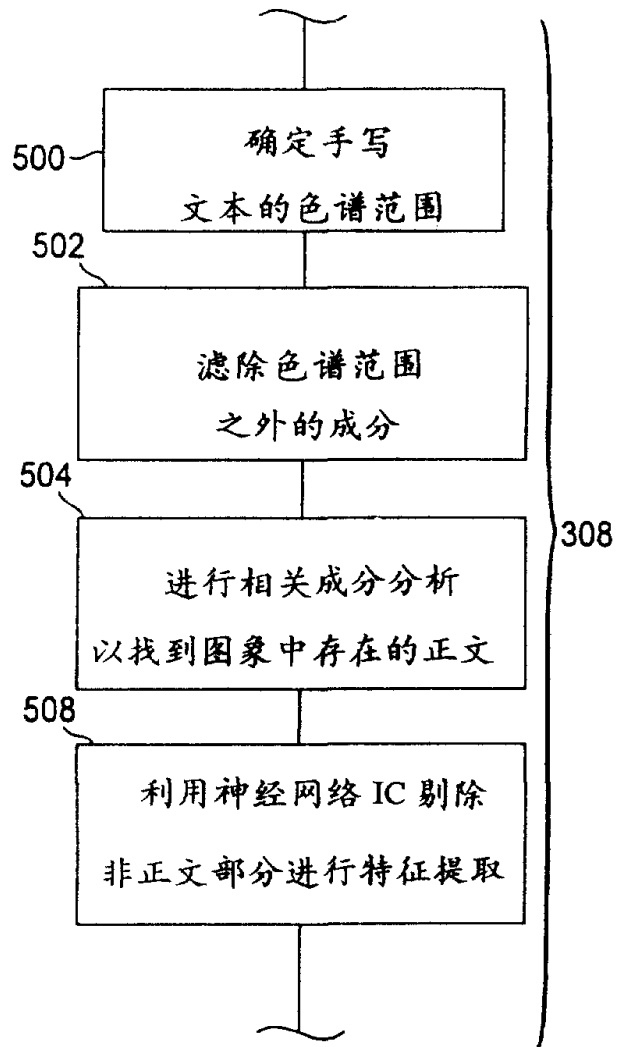


图. 5