

[12] 发明专利申请公开说明书

[21] 申请号 00806687.6

[43]公开日 2002年6月19日

[11]公开号 CN 1354935A

[22]申请日 2000.2.24 [21]申请号 00806687.6

[30]优先权

[32]1999.2.26 [33]US [31]09/259,135

[86]国际申请 PCT/US00/05098 2000.2.24

[87]国际公布 WO00/51286 英 2000.8.31

[85]进入国家阶段日期 2001.10.24

[71]申请人 奥廷提戴特控股有限公司

地址 美国纽约州

[72]发明人 C·D·波罗曼

[74]专利代理机构 中国专利代理(香港)有限公司

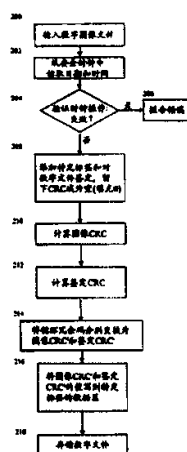
代理人 陈 弄

权利要求书3页 说明书17页 附图页数8页

[54]发明名称 包括安全文件标记的数字文件管理和成像系统及方法

[57]摘要

一种数字文件管理和成像系统,用每个存储的图像(200)记录附加的独立的数据,该图像包括从一个不可被用户设置的安全时钟(202)获取的一个真实日期,由对该图像数据(210)和该真实日期的循环冗余码校验(循环冗余码校验)算法得到的一个数字。在该文件得到后尽可能快的将该附加数据记录在每个数字文件中。如果该文件在记录附加数据后被以某种方式更改,重新对更改过的文件计算的图像CRC将不能同该文件中原始图像CRC相匹配。这样就可以检测到文件已经被更改的事实。同样,如果该真实日期被以某种方式更改,重新计算该日期CRC将同样的揭示这一事实。



权 利 要 求 书

1. 一种数字文件管理和成像系统包括：
输入一个数字文件的设备；
5 一个提供日期和时间信息的安全日期和时间基准；
从上述日期和时间信息生成一个日期/时间值的设备；
从上述数字文件生成一个图像值的设备；
用上述日期和时间信息、上述日期/时间值和上述图像值标记上述数字文
件的设备，以及
10 存储上述标记过的数字文件的设备。
2. 根据权利要求1的系统，其特征在于：上述安全日期和时间基准是一个
本地安全时钟。
3. 根据权利要求1的系统，其特征在于：上述生成该日期和时间值的设备
执行一个循环冗余码校验算法。
15 4. 根据权利要求1的系统，其特征在于：上述生成该图像值的设备执行一
个循环冗余码校验算法。
5. 根据权利要求1的系统，其特征在于：还包括变换该日期/时间值的设备
以及上述设备用变换过的日期/时间值标记该数字文件。
6. 根据权利要求5的系统，其特征在于：变换该日期/时间值的设备执行一
20 个数学变换。
7. 根据权利要求1的系统，其特征在于：还包括变换该图像值的设备以及
上述设备用变换过的图像值标记该数字文件。
8. 根据权利要求7的系统，其特征在于：变换该图像值的设备执行一个数
学变换。
25 9. 根据权利要求1的系统，其特征在于：该数字文件是一个图像文件。
10. 根据权利要求1的系统，其特征在于：该数字文件是一个文本文件。
11. 根据权利要求1的系统，其特征在于：该数字文件是来源于一个数码相
机的文件。
12. 根据权利要求1的系统，其特征在于：该数字文件来源于一个医学成像
30 设备。

13.根据权利要求1的系统，其特征在于：该数字文件是一个计算机应用程序产生的文件。

14.根据权利要求1的系统，其特征在于：还包括验证一个标记过文件的设备。

5 15.一种数字文件管理系统，包括：

一个数字文件；

一个提供日期和时间信息的安全日期和时间基准；

一个来自上述日期和时间信息的日期/时间值；

一个来自上述数字文件的图像值；

10 一个用上述日期和时间信息、上述日期/时间值以及上述图像值标记的标记过的数字文件。

16.根据权利要求15的系统，其特征在于：上述安全日期和时间基准是一个本地安全时钟。

15 17.根据权利要求15的系统，其特征在于：该日期和时间值来自一个循环冗余码校验算法。

18.根据权利要求15的系统，其特征在于：该图像值来自一个循环冗余码校验算法。

19.根据权利要求15的系统，其特征在于：还包括一个变换过的日期/时间值，并且上述标记过的文件是用该变换过的日期/时间值标记的。

20 20.根据权利要求15的系统，其中还包括一个变换过的图像值，并且上述标记过的文件是用该变换过的图像值标记的。

21.根据权利要求15的系统，其特征在于：该数字文件是一个图像文件。

22.根据权利要求15的系统，其特征在于：该数字文件是一个文本文件。

25 23.根据权利要求15的系统，其特征在于：该数字文件是来源于一个数码相机文件。

24.根据权利要求15的系统，其特征在于：该数字文件来源于一个医学成像设备。

25.根据权利要求15的系统，其特征在于：该数字文件是一个计算机应用程序产生的文件。

30 26.一个数字文件管理和成像的方法，包括步骤：

提供一个数字文件;

从一个来自一个本地源的安全日期和时间基准提供日期和时间信息;

从上述日期和时间基准生成一个日期/时间值;

从上述数字文件生成一个图像值;

- 5 用上述日期和时间信息、上述日期/时间值和上述图像值标记上述数字文件; 并且

 存储上述标记过的数字文件。

 27.根据权利要求26的方法, 其特征在于: 上述生成该日期和时间值的步骤执行一个冗余循环码算法。

- 10 28.根据权利要求26的方法, 其特征在于: 上述生成该图像值的步骤执行一个冗余循环码算法。

 29.根据权利要求26的方法, 其特征在于: 还包括变换该日期和时间值以及用变换过的日期和时间值标记该数字文件的步骤。

- 15 30.根据权利要求26的方法, 其特征在于: 还包括变换该图像值以及用变换过的日期和时间值标记该数字文件的步骤。

 31.根据权利要求26的方法, 其特征在于: 上述提供一个数字图像的步骤包括光学扫描一个原始图像成为一个数字图像。

- 20 32.根据权利要求26的方法, 其特征在于: 还包括重新计算该日期/时间值和图像值并且将重新计算的值与标记在该文件中的日期/时间值和图像值分别进行比较。

说明书

包括安全文件标记的数字文件管理和成像系统及方法

5 技术领域

本发明一般涉及数字成像系统，特别是涉及数字文件鉴定。

背景技术

数字成像是作为一个数字光栅图像的一个图像或物体的再现和存储。数字图像在许多行业得到越来越多的使用，部分是由于使其可能的技术越来越多地
10 可以得到，部分是由于它比传统的存储方法提供许多优势：存储空间减小，访问速度加快，集中检索能力（例如，搜索能力），能够方便的制作“多重”和“备份”文档拷贝，并且能够快速转移和发送文档。

对于原始纸质文档，数字成像系统通常扫描该纸质文档并将扫描的一个文档的再现存储为数字光栅图像。通常使用一个光学扫描设备来扫描纸质原件以
15 存储为一个数字图像。扫描的图像是该原件的准确再现（只是受限于该扫描设备的分辨率限制），并可以包括笔迹，签名，照片，图形等。换言之，来自数字相机、医学成像设备或者其他来源的数字图像同样可以被存储在一个数字成像系统中。

已知的成像技术的一个缺点是其固有的数字图像被更改的能力，比如说为了欺骗的目的。例如，虽然一个原始纸质文档可以被修改，这些修改（擦除或
20 添加）通常将会留下可描述的证据，与之相反，这些文档的数字图像可以被完美的修改而不留下这些证据。因此，在图像的鉴定是决定性的并且可能产生问题的情况下（例如法律和医学领域），使用数字图像通常不被优先选择，不可接受或者不容许并且因此被避免使用。

25 虽然可以得到许多不同的数字图像格式，在每种情况下，其数据都有潜在的被更改的可能。即使该数字成像系统没有明确的提供一种编辑功能，该图像可能用一个第三方工具编辑。

一个推荐的解决方法是使用写一次可读多次（WORM）的光学媒体存储数字图像。WORM媒体存储的一个优势是其存储的数据是固有不可以更改的
30 一数据只可以被写到媒体上一次。然而，这种方法也有几个缺点。例如，记录

在WORM媒体上的数据可以从原始记录的WORM盘上拷贝到可重写媒体上，被修改，然后不带任何可追踪的痕迹地被记录在新的WORM盘上，这些操作可以没有。

5 另外，虽然可以非常有信心的宣称在任何一个特定WORM盘上的数据自从被记录到盘上就不曾被更改过，但是不能用任何确定的或明确的方法来判断记录数据的日期和时间或者数据是否任何种类的同原件匹配。

10 一个已知的在文件验证技术的优点在于为“电子签名”的注册提供数字文件（图像，文字处理文档，音频或视频片段等）。已知允许用户就地选择一个文件并就地运行某服务提供商提供的一个程序，以仅仅基于文件内容生成该所选数字文件的“电子签名”。这个签名连同用户提供一个文件名和用户选择的
关键字被上载到该提供商的站点，并且被存储在为这个特定用户建立的账户下该服务提供商维护的注册数据库中。一个特定提供商生成一个显示内部别名、签名的“注册的授权”。

15 在以后验证数字文件的内容和提交的数据需要在线访问该服务提供商的站点，并根据文件名和关键字检索以前的注册记录。该检索到的数据库记录显示出文件签名和注册文件签名时的原始数据。为了完成验证，用户必须（再一次就地）对需要验证的文件运行电子签名程序，并将重新产生的签名和检索到的注册签名进行比较，以判定正被验证的数字文件的签名是否同原始注册文件的签名匹配。

20 用户现在得到的是待处理文件的签名同特定时间注册的文件的签名相匹配的验证。

本发明的概述和目的

本发明通过安全图像标记提供数字文件鉴定解决了已知数字成像系统中在图像鉴定方面的前述及其他问题和缺陷，并取得了技术进步。

25 在多个不同的方面，本发明的目的之一是提供一个通过安全文件标记提供数字文件鉴定的数字文件管理系统和方法。

在本发明的一个实施例中，一个数字文件管理系统包括输入一个数字文件和一个提供日期和时间信息的安全日期和时间基准的设备。生成一个由安全日期和时间信息得到的日期/时间值。一个由数字文件本身得到的图像值。该数字
30 文件用该日期和时间信息、该日期/时间值和该图像值标记。然后存储标记过的

文件。

另一个实施例可以包括通过循环冗余码（CRC）算法产生该时间/日期值和图像值，并且通过数学变换来变换该日期/时间值和图像值，并使用变换后的值来标记该数字文件这些特征。

5 在其他实施例中，该安全日期和时间基准是一个本地安全时钟。

在多个不同的实施例中，该数字文件可以是一个图像文件、一个文本文件或者任何其他文件格式。

本发明的另一个实施例允许通过一个光学扫描仪来扫描一个原始图像成为一个数字图像，或者用直接取自数码相机或医学成像设备的方法输入一个数字
10 图像。标记过的数字文件同样可以被存储在光学存储器。

附图的简要描述

本发明的前述和其他特征以及优点根据后面对示范实施例的详细描述将变得更加明显，如附图所示，其中

图1说明了本发明的DocSTAR实施例的一个系统装置。

15 图2是说明依照本发明的一个实施例对文件作标记的流程图。

图3是说明依照本发明的一个实施例对一个归档标记图像验证CRC的流程图。

图4是说明本发明的一个设置该安全时钟的实施例的流程图。

图5是说明依照本发明的一个实施例对标签图像文件格式图像计算该图像
20 CRC的流程图。

图6是说明依照本发明的一个实施例对标签图像文件格式图像计算该日期CRC的流程图。

图7是说明依照本发明的一个实施例对联合图像专家组格式图像计算该图像CRC的流程图。

25 图8是说明依照本发明的一个实施例对联合图像专家组格式图像计算该日期CRC的流程图。

附图的详细描述

本发明的下述描述用于说明集成在监管文档管理和成像系统DocSTAR™中的Authenticated™图像鉴定系统，上述两者都可以从本发明的受让人BitWise
30 设计有限公司得到。虽然本发明的DocSTAR实施例专用于存储、标记和鉴定纸

质文档原件，将要描述的本发明的方法和系统可以处理任何数字文件。下述参考DocSTAR实施例的讨论只是被用来方便解释和理解本发明的说明，而决不是为了限定。

图1说明本发明的DocSTAR文档管理和成像系统装置的一个示范实施例。

- 5 一个DocSTAR系统主机100被配置成同一个输入设备110、一个存储设备120和一个安全时间和日期基准130通信。

在这个实施例中，系统主机100实现为一台IBM个人计算机或工作站，输入设备110实现为一台光学扫描仪，存储设备120实现为一台光学存储设备，安全时间和日期基准130由集成了一个安全时钟的一个硬件密钥提供。

- 10 光学扫描仪110对原始图像进行扫描。得到的数字图像将依照本发明的方法由系统主机110处理，该方法在这里将作更详细的讨论，然后数字图像存储在光学存储设备120中，以后可以从光学存储设备检索该数字图像。

- 本发明的图像鉴定系统的操作从某一方面来说就是用各个存储的数字文件记录附加独立数据。这些附加数据包括：由不能被用户（Authenticate™）设置的安全时钟获得的一个“真实的日期”（下面有详细描述），由对该图像数据的一个循环冗余码（CRC）算法（后面详述）得到的一个数字，该数字被称作CRC “图像”，以及一个从“真实的日期”得到的CRC,被称作“日期CRC”。
- 15

- 这些附加的数据在系统获得图像后(例如从DocSTAR实施例中的扫描仪110中)尽可能快的优先的被记录在各个数字文件中。如后面更详细讨论的那样，如果该图像在记录附加数据后被以任何形式更改，重新计算被更改图像的图像CRC将不能同该图像中记录的原始图像CRC匹配。这样，图像已经被更改或者已经被损坏的事实就可以被检测到。相类似，如果真实的日期被以任何形式更改，重新计算日期CRC将同样揭示这个事实。
- 20

- 可以在任何时间检查和验证该图像CRC和数据CRC。如果重新计算的值同记录值相匹配，就可以非常有信心的宣称当前记录的图像是在特定日期记录的并且从那以后不曾被以任何形式更改过。没有其他已知系统，包括纸质存储，可以提供类似的对于文档建立日期或鉴定的保证。
- 25

参考图2，这里将描述本发明的操作。

- 首先获得数字文件（从存储设备检索或者从输入设备110接收）。（步骤200）从安全时钟130获得日期和时间信息。（步骤202）评估该安全时钟是
- 30

否正常工作。（步骤204） 如果认为该安全时钟功能正常，那么从该时钟读出并被接受日期和时间数据（在步骤202）。如果判定该安全时钟故障，将返回一个错误指示并且停止图像处理。（步骤206） 在认定该时钟工作正常情况下（在步骤204），添加特殊标签（下文将讨论）和鉴定信息（包括日期和时间）

5 到该数字文件上,并且初始化CRC数据区为0（也就是，数据区用0填充）。（步骤208）

随后分别计算出由该图像内容和该鉴定信息得到的两个计算值。这些计算值可以用基于该数字文件所包括数据的任何允许检测数据修改的方式来计算，例如一个标准求和。在本发明的当前实施例，使用了本质上更复杂的求和计算，循环冗余码（CRC）来得到计算值。然而，任何可以提供从该文档内容数据得到一个数字并适合检测数据修改的计算方法都可以接受。

在当前实施例中，计算值由一种已知的CRC算法（后面详述）生成，该算法通过对该图像内容和该鉴定信息的运算，分别生成一个图像CRC和一个鉴定CRC。（步骤210，212） 该图像CRC和鉴定CRC经过一种为增加安全目的的
15 专有数学变换的“变换”生成一个图像CRC'和一个鉴定CRC'。（步骤214）

该图像文件随后用该图像CRC'和该鉴定CRC'来标记。（步骤216） 标记过的数字文件通过光学存储设备120被存储在光学媒体上。（步骤218）

对该图像以及该时间和日期标记的鉴定可以随后通过检查存储在该数字文件中的计算值来确定，如图3所示描述了在一个归档图像中验证CRC的一个实
20 施例的示范流程图。

对一个数字文件验证CRC的第一步是读该特殊标签和日期域，并检索存储的图像CRC和日期CRC值。（步骤300） 如果该CRC值不能在该数字文件中定位或读出（步骤302），那么，就判定该图像没有被正确的归档，或者该图像已经被更改或被损坏，并给出一个错误。（步骤304） 如果找到该特殊标签，
25 就重新计算该数字文件和该日期字串的CRC。（步骤306） 此时使用与原来所用计算CRC相同的算法来重新生成这些CRC，该重新计算的图像CRC被变换并与从该标签中读出的图像CRC比较。（步骤308） （或者，该存储的图像CRC在同该重新计算值比较前被反变换）如果该重新计算数字文件CRC同存储在特殊标签中的值不匹配，就判定图像已经被更改或被损坏,并指示一个错误。（步
30 骤310）如果该存储的图像CRC同重新计算的图像CRC顺利比较(即它们匹配)，

测试该日期CRC。一个重新计算的日期CRC被变换并同从标签中读出的日期CRC比较。（步骤312）（或者，该存储的日期CRC可以在同重新计算值比较之前被反变换）如果重新计算的日期CRC同存储在该特殊标签中的值不匹配，就判定该日期字符串已经被更改或被损坏，并指示一个错误。（步骤314）如果
5 两个日期CRC匹配，此时图像CRC和日期CRC都已经顺利比较，该数据文件被判定未被更改并且因此鉴定为真的。（步骤316）

从前面的描述可以意识到，使用一个安全的、不可更改的时钟是本发明的基础。它起到一个不可以被用户更改的安全时间和日期源的作用。该安全时钟在电池备份的帮助下即使在计算机关闭的时候还能保持时间和日期。

10 可以使用能提供一个安全时钟的客户设计的硬件或者商业上可以得到的产品。在这两种情况下，都必须存在一个机构来防止欺骗性的或任意的日期/时间调整。

在DocSTAR实施例中，利用了一个商业上可以得到的在物理硬件密钥中集成了一个安全时钟的产品（有时称作“dongle”）。该硬件密钥连接到计算机的并行端口，并且可以通过该制造商提供的一个应用程序接口（API）来访问。
15

在本发明的DocSTAR实施例中选择使用的硬件密钥是可以从Aladdin知识系统有限公司得到的TIMEHASP-4。该硬件密钥的安全性由用户ASIC（专用集成电路）芯片、由该系统提供商（例如本专利受让人和DocSTAR系统的提供商
20 BitWise设计有限公司）专用的唯一一组密码以及在该制造商的编程接口和设备驱动程序中的高级保护算法和反调试技术来保护。这些措施为该安全时钟提供了高度的安全性。

在装配DocSTAR主计算机时，当前日期和时间被现场编程到包括在硬件密钥中的安全时钟中。虽然可以使用任何时间设置，当前实施例中的安全时钟
25 被设置为格林威治标准时间（GMT）以减少为不同本地时区或夏令时而调整该时钟的需要。

可以集成进来一种机构以调整时钟实现复位或更正长时间积累的细微不准确。例如，在图4示例说明的一个实施例中，其安全时钟中的日期和时间可以用驻留在用户系统中的一个专用管理程序来改变，该程序只有当用户给出一个
30 正确的鉴别码后才允许改变该安全日期和时钟。该鉴别码由像本发明的受让人

BitWise设计有限公司的技术支持部门这样的系统供应商提供的。该鉴别码只能
够将该安全时钟的日期和时间从它的当前日期和时间值改变到由该系统供应商
维护的当前格林威治标准时间。这样就防止用户任意更改该安全时钟从而在图
像上标记一个不正确或欺骗性的日期和时间标记。

- 5 在当前实施例中，需要一个鉴别码才能调整该安全时钟。为了获得该鉴别
码，一个系统提供系统的支持技术人员将一个硬件密钥序列号和当前安全时钟
日期输入到在BitWise设计有限公司维护的一个安全客户程序（“鹰呼叫跟踪系
统”）中，（步骤400），该程序将产生一个鉴别码（步骤402）。该鉴别码将
只允许现场技术人员或者终端用户将该安全时钟调整到在BitWise设计有限公司
10 建立和维护的日期和时间上。

当前实施例的鉴别码通过一个数学算法确定，该算法根据当前安全时钟日
期、该硬件密钥序列号和期望对日期和时间的改变给出一个唯一的鉴别码。该
鉴别码是有限有效的，因为它不能在以后的另一天作用将该时钟复位到该鉴别
码产生当天的日期和时间。

- 15 在用户终端输入鉴别码（步骤404）。在用户终端输入期望的时钟设置。
（步骤406） 客户系统使用的管理程序允许一个小的时间窗口（20分钟），在
该窗口范围内的任意时间输入都将同该鉴别码匹配。鉴别码在内部计算为给定
改变时间的前5分钟和后15分钟内。如果给出的鉴别码满足该时间窗口内的某
个鉴别码，就认为鉴别码正确并执行。这将允许一个现场技术人员在传递鉴别
20 码时可以有几分钟的延时。

- 这样由鉴别码对期望的设置作验证以决定该鉴别码是否将鉴定请求的日期
和时间的更改。（步骤408） 如果确定为无效，返回一个错误并且该时钟不更
新。（步骤409）对于有效的请求，对该安全时钟的实际改变只有在更新时钟
命令输入到该用户终端后才发生。（步骤410） 这允许一个现场技术人员精确
25 地将该现场时钟同在BitWise设计有限公司内维护的时钟同步。当该更新命令发
出后，该鉴别码再次对时钟信息验证以保证该鉴别码仍然有效。（步骤412）如
果确定为无效，返回一个错误并且该时钟不更新（步骤412）。该时钟更新（步
骤414）。

- 此外，该安全时钟可以在提供商的设施内(如BitWise设计有限公司)由该服
30 务提供商通过直接连接该硬件密钥到一个指定的在BitWise设计有限公司内的鹰

系统并发出更新安全时钟命令来重新编程。在该硬件密钥序列号被验证后，该安全时钟日期和时间被更新到在BitWise设计有限公司内维护的格林威治标准时间的日期和时间上。

在其他替换实施例中，为了更正长期积累的不准确而调整或者设置该时钟可以通过一种自动过程来实现，在这种自动过程中一个用户可以使得一个时钟从一个远端安全时钟更新但该用户自己不能实际设置该时钟信息。

上面描述的手动或自动设置和更新时钟的方法都能够防止用户用不正确或欺骗行的日期和时间标记图像任意改变安全时钟。

可以预见由于现有可用技术的限制，在每个时钟内的电池都将最终失效，或者该时钟经过一段时间出现故障。这些情况由软件在图像处理前进行测试以保证来自故障时钟（或失效电池）的无效日期不被记录在图像上，从而损害图像标记的可信赖性。在发生时钟失效事件时，图像归档将被禁止直到该时钟被修复或替换。

在本发明的DocSTAR实施例中，上文提到的参考图2的计算值是CRC（RCs）。该CRC是一个32比特的整数值，它代表在一个数据块上执行已知的CRC-32算法的结果。该CRC-32算法是一个通用的不受版权保护的检测数据中甚至细小变化的算法，它有多种应用。例如，CRC被应用在通信领域来保证数据在未知质量的传送线上正确传送。它还被用于像在流行的PKZIP实用程序中那样检测压缩数据损坏。CRC的一个好处是可以检测到不使用它就检测不到的数据的变化。例如，在一个给定的数据块中发生比特错误但其和碰巧同原始数据的一致，如果采用标准的求和算法就检测不到这个错误。该CRC-32算法可以检测到这类变化，因为其结果码不是像标准求和算法中那样简单的数据元素的和。

这里将不给出该CRC-32算法的技术讨论。有许多不受版权限制的CRC-32算法和源代码的来源。下面是在本发明的DocSTAR实施例中实现的一个CRC-32算法的示例C++源代码。如前所述，本发明本身不需要使用CRC，并且任何可以提供由一个图像数据得到一个适合检测数据损坏的数字的计算方法都可以接受。示范C++源代码列在下面：

```
long CRCTable[]=
```

```
{
```

	OX00000000L,	Ox77073096L,	OxOEEOE612CL,	Ox990951BAL,
	Ox076DC419L,	Ox706AF48FL,	OxOE963A535L,	Ox9E6495A3L,
	OxOEDB8832L,	Ox79DCB8A4L,	OxOEOD5E91EL,	Ox97D2D988L,
	Ox09B64C2BL,	Ox7EB17CBDL,	OxOE7B82DO7L,	Ox90BFID91L,
5	Ox1DB71064L,	Ox6AB02OF2L,	OxOF3B97148L,	Ox84BE41DEL,
	Ox1ADAD47DL,	Ox6DDDE4EBL,	OxOF4D4B551L,	Ox83D385C7L,
	Ox136C9856L,	Ox646BA8COL,	OxOFD62F97AL,	Ox8A65C9ECL,
	Ox140I5C4FL,	Ox63066CD9L,	OxOFAOF3D63L,	Ox8DO80DF5L,
	Ox3B6E2OC8L,	Ox4C69105EL,	OxOD56O4IE4L,	OxOA2677172L,
10	Ox3CO3E4DIL,	Ox4BO4D447L,	OxOD2OD85FDL,	OxOA50AB56BL,
	Ox35B5A8FAL,	Ox42B2986CL,	OxODBBBC9D6L,	OxOACBCF940L,
	Ox32D86CE3L,	Ox45DF5C75L,	OxODCD60DCFL,	OxOABD13D59L,
	Ox26D930ACL,	Ox5IDE003AL,	OxOC8D75180L,	OxOBFDO6116L,
	Ox21B4F4B5L,	Ox56B3C423L,	OxOCFBA9599L,	OxOB8BDA50FL,
15	Ox2802B89EL,	Ox5FO58808L,	OxOC60CD9B2L,	OxOBIOBE924L,
	Ox2F6F7C87L,	Ox58684CIL,	OxOC16IIDABL,	OxOB6662D3DL,
	Ox76DC4190L,	OxOIDB7106L,	Ox98D22OBCL,	OxOEFD5102AL,
	Ox71B18589L,	OxO6B6B51FL,	Ox9FBFE4A5L,	OxOE8B8D433L,
20	Ox7807C9A2L,	OxOFOOF934L,	Ox9609A88EL,	OxOEIOE9818L,
	Ox7F6AODBBL,	OxO86D3D2DL,	Ox91646C97L,	OxOE6635COIL,
	Ox6B6B51F4L,	Ox1C6C6162L,	Ox85653OD8L,	OxOF262004EL,
	Ox6CO695EDL,	Ox1BOIA57BL,	Ox8208F4CIL,	OxOF50FC457L,
	Ox65BOD9C6L,	Ox12B7E950L,	Ox8BBEB8EAL,	OxOF9CB9887CL,
25	Ox62DDIIDDFL,	Ox15DA2D49L,	Ox8CD37CF3L,	OxOFBD44C65L,
	Ox4DB26158L,	Ox3AB551CEL,	OxOA3BC0074L,	OxOD4BB3OE2L,
	Ox4ADFA541L,	Ox3DD895D7L,	OxOA4DIC46DL,	OxOD3D6F4FBL,
	Ox4369E96AL,	Ox346ED9FCL,	OxOAD678846L,	OxODA6OB8DOL,
	Ox44042D73L,	Ox33031DE5L,	OxOAAOA4C5FL,	OxODDOD7CC9L,
30	Ox5OO5713CL,	Ox270241AAL,	OxOBEOB1010L,	OxOC9OC2086L,

	Ox5768B525L,	Ox206F85B3L,	OxOB966D409L,	OxOCE61E49FL,
	Ox5EDEF90EL,	Ox29D9C998L,	OxOBODO9822L,	OxOC7D7A8B4L,
	Ox59B33DI7L,	Ox2EB4OD81L,	OxOB7BD5C3BL,	OxOCOBA6CADL,
5	OxOEDB88320L,	Ox9ABFB3B6L,	OxO3B6E20CL,	Ox74BID29AL,
	OxOEAD54739L,	Ox9DD277AFL,	OxO4DB2615L,	Ox73DC1683L,
	OxOE363OB12L,	Ox94643B84L,	OxOD6D6A3EL,	Ox7A6ASAA8L,
	OxOE40ECFOBL,	Ox9309FF9DL,	OxOAOOAE27L,	Ox7DO79EBIL,
	OxOFOOF9344L,	Ox8708A3D2L,	OxIEOIF268L,	Ox6906C2FEL,
10	OxOF762575DL,	Ox806567CBL,	OxI96C3671L,	Ox6E6BO6E7L,
	OxOFED41B76L,	Ox89D32BEOL,	OxIODA7A5AL,	Ox67DD4ACCL,
	OxOF9B9DF6FL,	Ox8EBEEFF9L,	OxI7B7BE43L,	Ox6OBO8ED5L,
	OxOD6D6A3E8L,	OxOAIID937EL,	Ox38D8C2C4L,	Ox4FDFF252L,
	OxODIBB67FIL,	OxOA6BC5767L,	Ox3FB506DDL,	Ox48B2364BL,
15	OxOD8OD2BDAL,	OxOAF0A1B4CL,	Ox36034AF6L,	Ox4IO47A60L,
	OxODF60EFC3L,	OxOA867DF55L,	Ox316E8EEFL,	Ox4669BE79L,
	OxOCB61B38CL,	OxOBC66831AL,	Ox256FD2AOL,	Ox5268E236L,
	OxOCCOC7795L,	OxOBBOB4703L,	Ox220216B9L,	Ox5505262FL,
	OxOC5BA3BBEL,	OxOB2BDOB28L,	Ox2BB45A92L,	Ox5CB36AO4L,
20	OxOC2D7FFA7L,	OxOB5DOCF31L,	Ox2CD99E8BL,	Ox5BDEAEIDL,
	Ox9B64C2BOL,	OxOEC63F226L,	Ox756AA39CL,	OxO26D930AL,
	Ox9CO906A9L,	OxOEBOE363FL,	Ox72076785L,	OxO5OO5713L,
	Ox95BF4A82L,	OxOE2B87A14L,	Ox7BB12BAEL,	OxOCB61B38L,
25	Ox92D28E9BL,	OxOE5D5BEODL,	Ox7CDCEFB7L,	OxOBDBDF21L,
	Ox86D3D2D4L,	OxOFID4E242L,	Ox68DDB3F8L,	OxIFDA836EL,
	Ox8IBE16CDL,	OxOF6B9265BL,	Ox6FBO77EIL,	OxI8B74777L,
	Ox88085AE6L,	OxOFFOF6A70L,	Ox66063BCAL,	OxI101OB5CL,
	Ox8F659EFL,	OxOF862AE69L,	Ox616BFFD3L,	OxI66CCF45L,
30	OxOAOOAE278L,	OxOD70DD2EEL,	Ox4EO48354L,	Ox3903B3C2L,

```

OxOA7672661L,   OxOD06016F7L,   Ox4969474DL,   Ox3E6E77DBL,
OxOAED16A4AL,   OxOD9D65ADCL,   Ox4ODFOB66L,   Ox37D83BFOL,
OxOA9BCAE53L,   OxODEBB9EC5L,   Ox47B2CF7FL,   Ox3OB5FFE9L,
OxOBDBDF21CL,   OxOCABAC28AL,   Ox53B39330L,   Ox24B4A3A6L,
5  OxOBAD03605L,   OxOCDD70693L,   Ox54DE5729L,   Ox23D967BFL,
OxOB3667A2EL,   OxOC4614AB8L,   Ox5D68IB02L,   Ox2A6F2B94L,
OxOB40BBE37L,   OxOC3OC8EAIL,   Ox5AO5DFIBL,   Ox2DO2EF8DL
};

```

```

10  UINT32 CRCFileBlock(UINT16 hfile, UINT32 IOffset, UINT32 ILength, UINT32 ISeed)

```

```

{
    //用给定的开端值（seed）计算文件块的CRC
    //用 OXFFFFFFFFL作为原始开端值
    //成功返回 0, 错误返回 ISeed

```

```

15      int ret;
        char buffer[COPYBUFFERLEN];
        UINT32 IRemainLength;
        UINT16 uBlockSize;
        UINT32 ISourceOff,
20      UINT32 ICRC;
        UINT16 i, index;

        ICRC = ISeed;

25      if(ILength > COPYBUFFERLEN)
            uBlockSize = COPYBUFFERLEN;
        else
            uBlockSize = (UINT16)ILength;

30      IRemainLength = ILength;

```

```

    lSourceOff = lOffset;

    while(lRemainLength){
        ret = ReadFileBlock(buffer, hfile, lSourceOff, uBlockSize);
5         if(ret)
            return lSeed;

        for (i=0; i<uBlockSize; i++)
            index = (UINT16)(lCRC^buffer[i])&(UINT16)0x000000FFL;
10         lCRC = ((lCRC >> 8) & 0X00FFFFFFL) ^ CRCTable[index];
    }

    lCRC = ~lCRC;

    lRemainLength -= uBlockSize;
15    lSourceOff += uBlockSize;
    if(lRemainLength < uBlockSize)
        uBlockSize = (UINT16)lRemainLength;
    }

20    return lCRC;
}

UINT32 CRCBlock(char* buffer, UINT16 nlength, UINT32 lSeed)
{
25    //用给定的开端值计算文件块的CRC
    //用 0xFFFFFFFFL作为原始开端值
    //成功返回 0, 错误返回 lSeed (忽略错误)

    UINT32 lCRC;
30    UINT16 i, index;

```

```

    ICRC = lSeed;

    for (i=0; i<nlength; i++){
5         index = (UINT16)(ICRC ^ buffer[i]) & (UINT16)0x000000FFL;
          ICRC = ((ICRC >> 8) & 0X00FFFFFFL) ^ CRCTable[index];
        }
        ICRC = ~ICRC;
        return ICRC;
10    }

```

示例计算CRC-32的C++源代码

虽然可以单独使用一个CRC值，通过增加对该CRC值的数学变换可以集成一个更高级别的安全性到本发明中以确保图像的鉴定。如指出的那样，一个典型的计算一个CRC-32的算法是不受版权限制的，并因此很容易得到。这个事实加上这里给出的细节将允许任何人重新计算一个改变过的图像的CRC，从而使他们有能力伪造一个“鉴定”并错误的证实图像为真实的和未改变的。在本发明中，实际计算出来的（图像或日期）CRC在标记图像前经过数学变换成一个新的值。对变换的功能的要求是对任何输入值的结果保持一致，并且对于每个唯一的输入值其结果值唯一。该变换可以是，例如，输入值的比特顺序的倒置，输入值与一个恒定的预定的“幻数”的OR操作，或者是这些操作的一个组合。

虽然该特定变换技术的实现不是至关重要,应该认为在本发明实践中用到的实现变换的该技术细节仍然对提供商是机密的,亦即,一个“专有变换技术”，因为对该方法的任何泄露或分发将可能损害系统安全性和有效性。举一个简单类比，不能成功的保护专有变换技术本质上将等同于用密码保护一个文件然后公开这个密码。

在数字文件内的标签中记录信息需要各数字文件格式的知识和管理文件格式的结构的标准。这些标准指示出信息将如何被存储在该文件中，按照何种顺序，使用什么压缩算法等。多数数字文件格式预留空间以容纳除了图像数据外

的用户数据存储。本发明的DocSTAR文件管理和成像系统实施例使用已知的TIFF（标签图像文件）和JPEG（联合图像专家组）文件格式来分别存储（扫描的）黑白或彩色图像。标签图像文件和联合图像专家组图像文件格式标准允许在该图像文件内部以不影响显示图像的方式包括用户数据。很容易理解，本发明同样可以应用于其它文件格式，只要这些文件格式在文件中具备一个存储用户定义数据的机构，或者标记有用户定义数据的文件可以存储在附加文件或分开的数据库中。例如，文字处理文档，表格，数字音频或视频，或者任何其他数字化文件。

已知的标签图像文件格式允许以压缩的方式存储图像数据和该图像相关信息（标签），例如使用的压缩方法，分辨率，大小，颜色数目，标题，日期等。

一份书面的世界范围的标准定义了标签图像文件格式，哪些标签是必选，哪些标签是可选，如何使用特定的标签。标签图像文件标准的维护组织Adobe公司接受那些开发在标签图像文件图像内使用标签的应用程序的公司的客户标签号的请求。Adobe将给每个公司分配唯一的号码以防止客户之间的冲突。例如，本发明的受让人BitWise设计有限公司申请并分配到自己专有的标签号。其他客户同样将分配到他们各自的唯一的专有标签号。使用用户标签允许存储一个用户数据块。标签图像文件规范要求程序忽略它们不理解的并且不在该基本规范中的标签。这允许通用图像阅读器可以观看、显示和打印包含用户标签的图像，因为这种图像仍然符合标签图像文件规范。

就标签图像文件图像文件来说，使用了下列标签图像文件图像标签：

标签# 用途

10Dh 文档名称

10Eh 图像描述

25 132h 日期时间

9244h BitWise DocSTAR用户标签1

用户数据块包括的专有信息有：

图像CRC

鉴定CRC

30 图5所示是说明为一个标签图像文件格式图像文件计算一个图像CRC的一

个示例流程图。对该标签图像文件格式图像文件计算该图像CRC要求使用一个给定的32位开端值计算一个给定数据块的一个CRC-32。该初始开端值被设置为-1。（步骤500）。该子程序根据该文件的图像文件目录（IFD）得到该标签图像文件文件的格式，对每个图像文件目录项和它的相关数据计算一个CRC-32（步骤502）将前面CRC-32的结果作为开端传送到下一个计算（步骤510）直到所有的图像文件目录项都循环过一遍。（步骤506）

除下列标签和数据域外的所有标签和数据域都要处理（步骤508）：

标签#	描述
0x010d	标签图像文件标签_文档名称
0x010e	标签图像文件标签_图像描述
0x0132	标签图像文件标签_日期时间
0x9244	标签图像文件标签_DocSTAR标签1

当该文件的所有图像文件目录项处理完后（步骤506），使用一个专有变换方法（上文描述）将该结果CRC值变换成一个唯一和安全的CRC'值。（步骤512）变换过的图像CRC值CRC'随后被存储在该图像文件中。（步骤514）

图6所示是说明为一个标签图像文件格式图像文件计算一个日期CRC的一个示例流程图。为该标签图像文件格式图像文件计算日期CRC需要一个能够使用给定32位开端值来对一个给定的数据块计算一个CRC-32的子程序。该初始开端值被设置为该图像CRC值（步骤600）。该子程序读取0x0132标签图像文件标签_日期时间标签。（步骤602）如果不能找到并读出该日期时间标签（步骤604），返回一个错误（步骤605）否则，对包括在该日期时间标签内的数据计算一个CRC-32。（步骤606）随后通过专有变换技术的方法将该结果CRC变换成CRC'（步骤608）并存储在该图像文件内。（步骤610）

联合图像专家组开发了同名文件格式并维护着联合图像专家组和JPG文件格式（有时也称作JFIF—联合图像专家组文件图像格式）的标准。该格式是为存储和传送照片图像而开发的。其使用的压缩技术非常适合存储像照片这样颜色变化的细微差别。

已经已知，一个联合图像专家组文件被解释成字符流，通过其中称作“标记”的特殊标识来分离该图像信息的不同元素和图像数据。除了联合图像专家组标准定义的一组供制造商用于特殊或专有特性的标记外，每个标记的确切含

义对本文并不重要。这些（专有）标记被命名为“APPx”，其中x是从0到9的一个数字。

本发明在联合图像专家组文件被存储时向其中添加一个特殊标记和数据块。在当前实施例中将使用“APP8”标记，其原因只是由于该标记很少被其他
5 制造商使用。该标记中有各种专有信息，包括下列信息：

鉴定

图像CRC

鉴定CRC

图7所示是说明为一个联合图像专家组图像文件计算一个图像CRC的一个
10 示例流程图。为该联合图像专家组图像文件计算CRC需要一个能够使用一个给定32位开端值来对一个给定的数据块计算一个CRC-32的子程序。该初始开端值被设置为-1。（步骤700）顺序读入该图像文件数据，同时确定APP8的位置并读入其内容。步骤702）如果不能找到并读入该APP8标记。（步骤704）返回一个错误（步骤705）对该文件中从开始到APP8标记但不包括APP8标记的
15 全部数据计算得到一个CRC-32。（步骤706）该计算的结果被用来作为计算该文件中APP8标记后边数据的一个CRC-32的开端值。（步骤708）该结果CRC通过该专有变换技术方法被变换成CRC'（步骤710）随后将变换过的图像CRC'存储在该图像文件内（步骤712）。

图8所示是说明为一个联合图像专家组图像文件计算一个日期CRC的一个
20 示例流程图。为该联合图像专家组图像文件计算CRC需要一个能够使用一个给定32位开端值来对一个给定的数据块计算一个CRC-32的子程序。该初始开端值被设置为该图像CRC值。步骤800）顺序读入该图像文件数据，同时确定APP8的位置并读入其内容。步骤802）如果不能找到并读出该APP8标记，（步骤804）返回一个错误（步骤805）对该APP8数据区或块内的安全数据字符串计算
25 出一个CRC-32。步骤806）通过专有变换技术方法将该结果CRC变换成CRC'（步骤808）将变换过的日期CRC'存储在该图像文件内。（步骤810）

本发明已经通过相关的特定实施例进行了说明和描述。但是，应该认识到，上面描述的实施例仅仅是对本发明原理的说明，而不是作为专有实施例。为了方便讨论本发明，在本发明的DocSTAR实施例中假设使用扫描成数据图像的
30 纸质文档原件（例如论文，照片等）。但是，这该领域的普通技术人员应该

认识到，本发明将同样应用于不考虑其来源或如何产生的任何数据文件，例如，来源于数码相机、医学成像设备、字处理或表格应用程序或其他来源的数字图像。

这里揭示并列出的实施例的各种变形可以被实现以获得本发明的益处。

- 5 还应该认识到可能由该领域普通技术人员所设计的前述的和许多不同修改、忽略和添加仍然没有脱离本发明的精神和范围。

因此认为本发明不限于给出的实施例，而是应该依照下面的权利要求来定义。

说明书附图

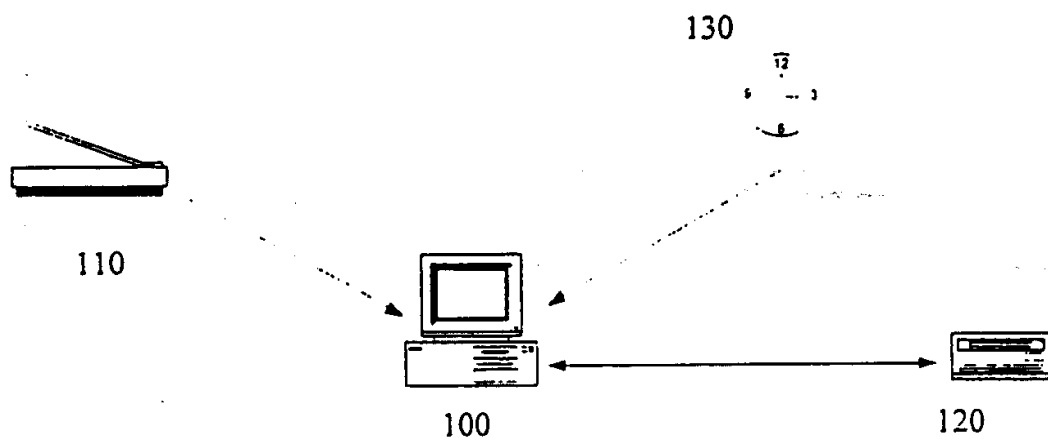


图 1

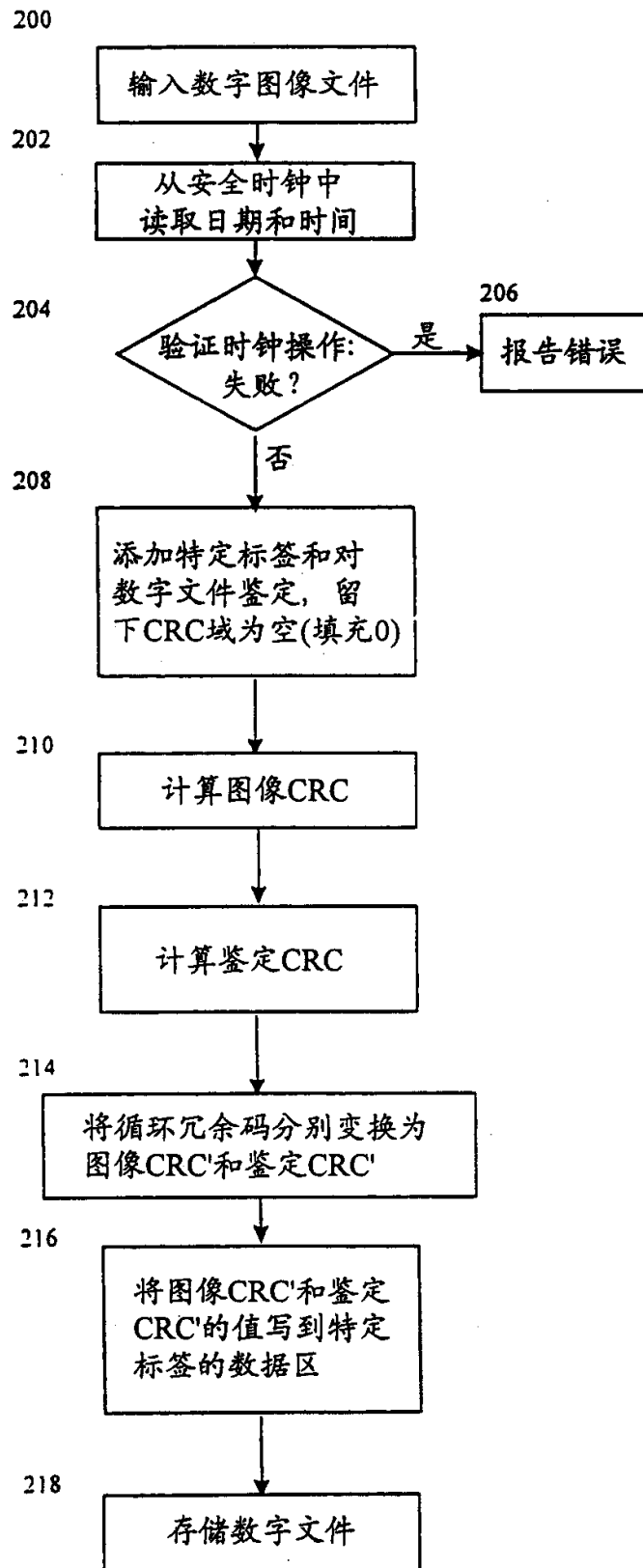


图 2

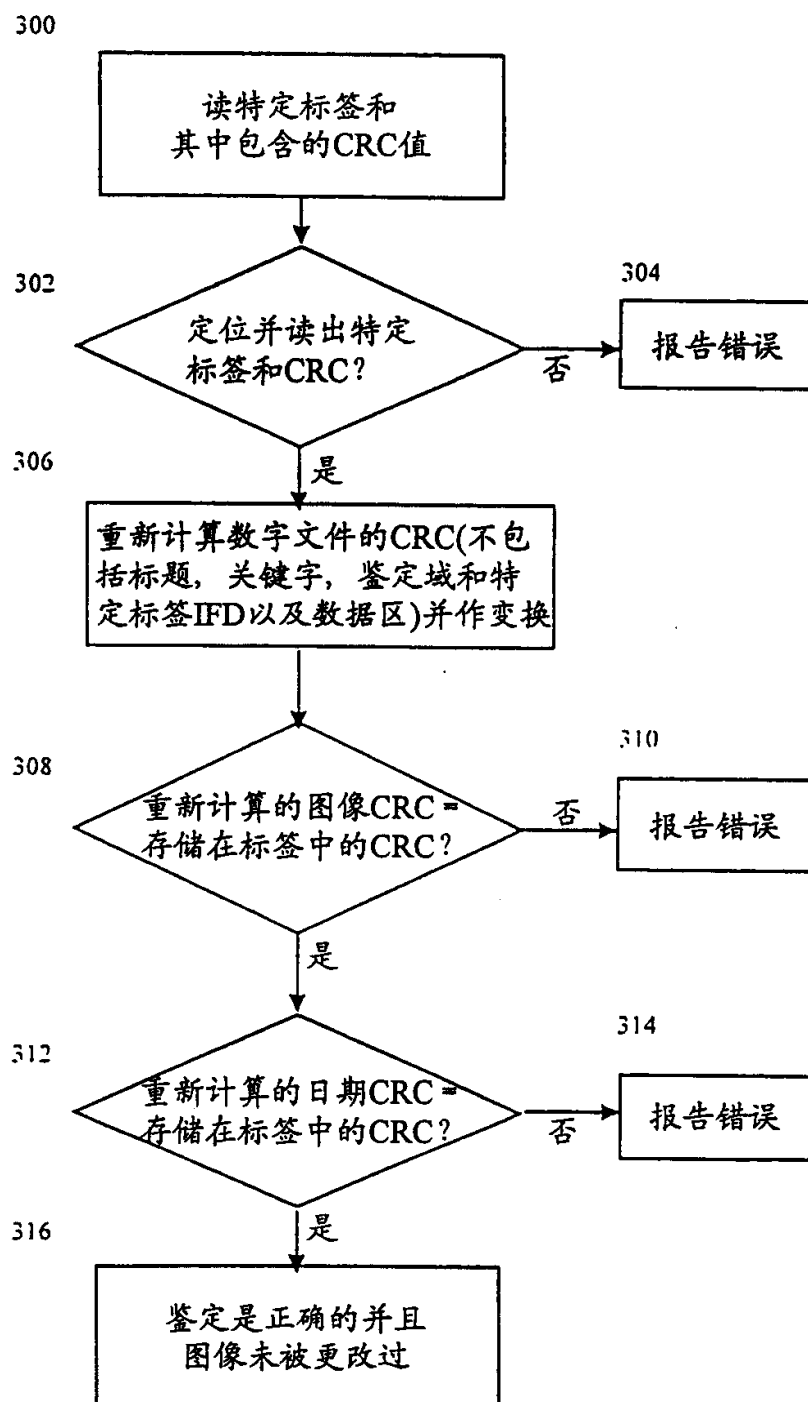


图 3

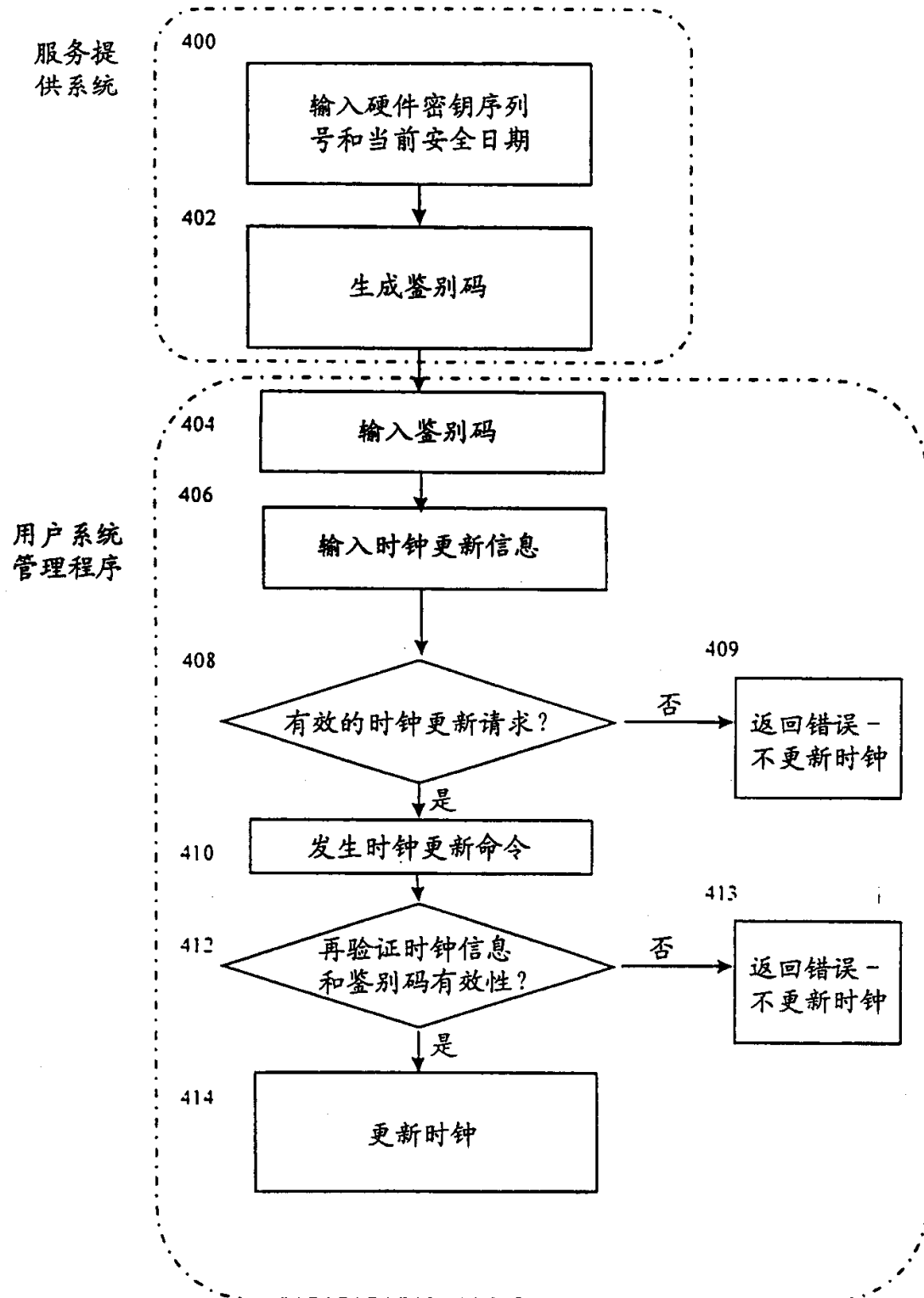


图 4

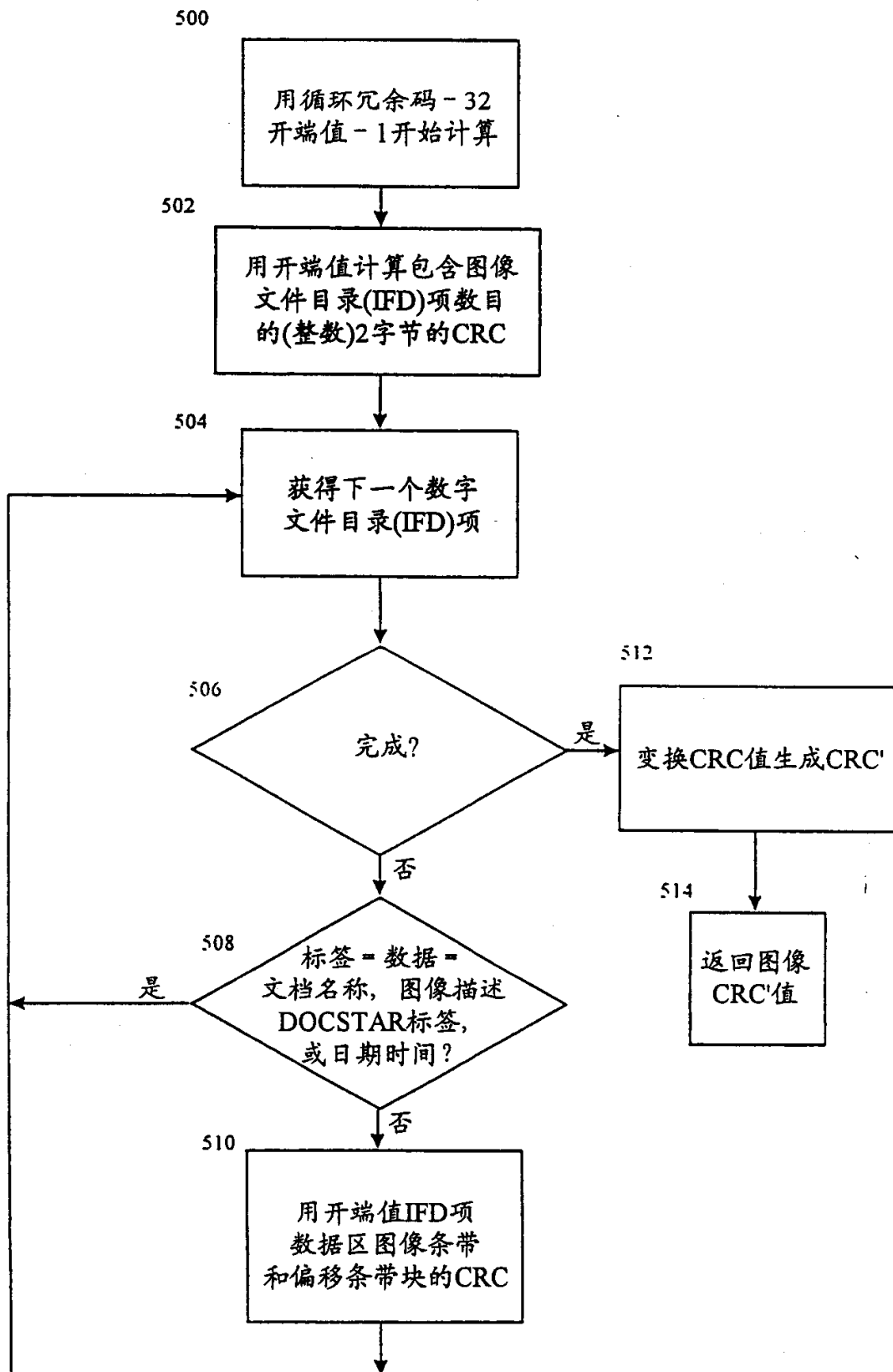


图 5

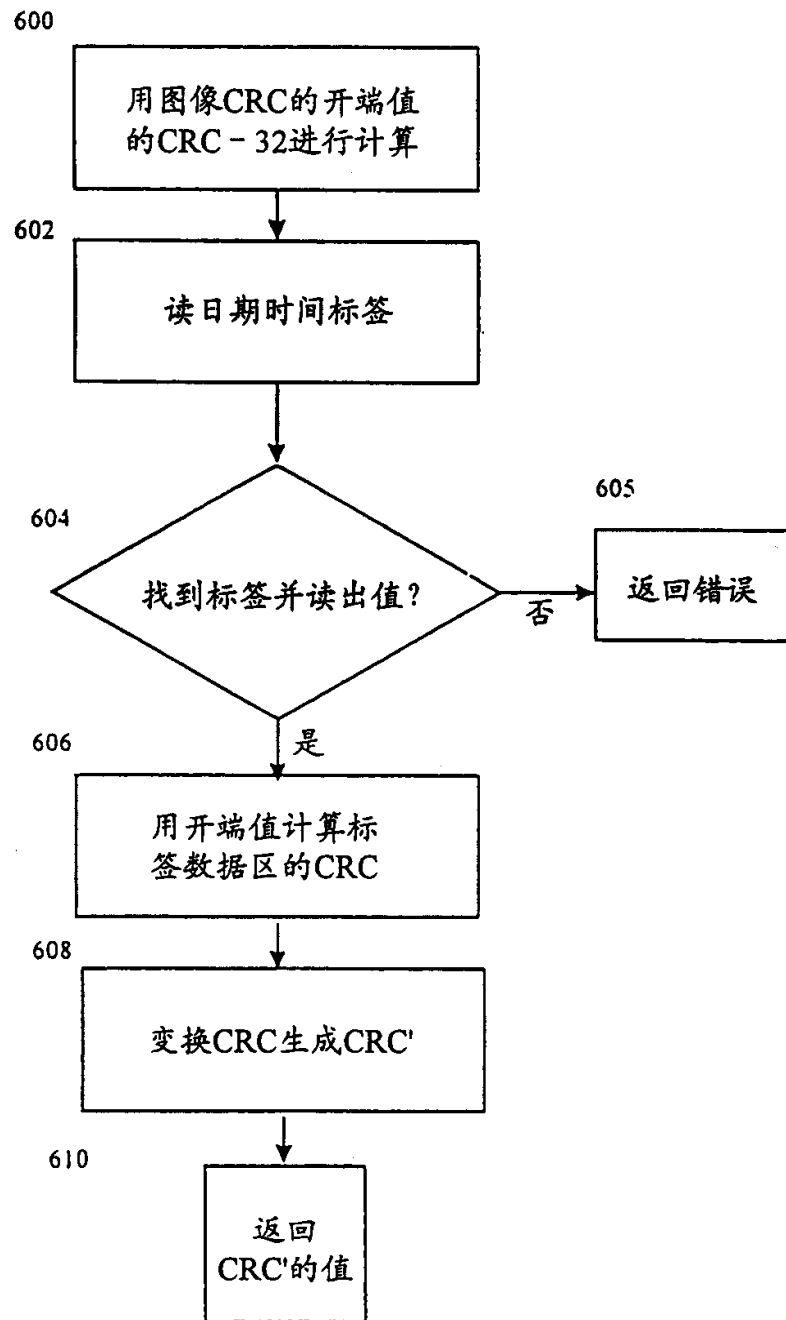


图 6

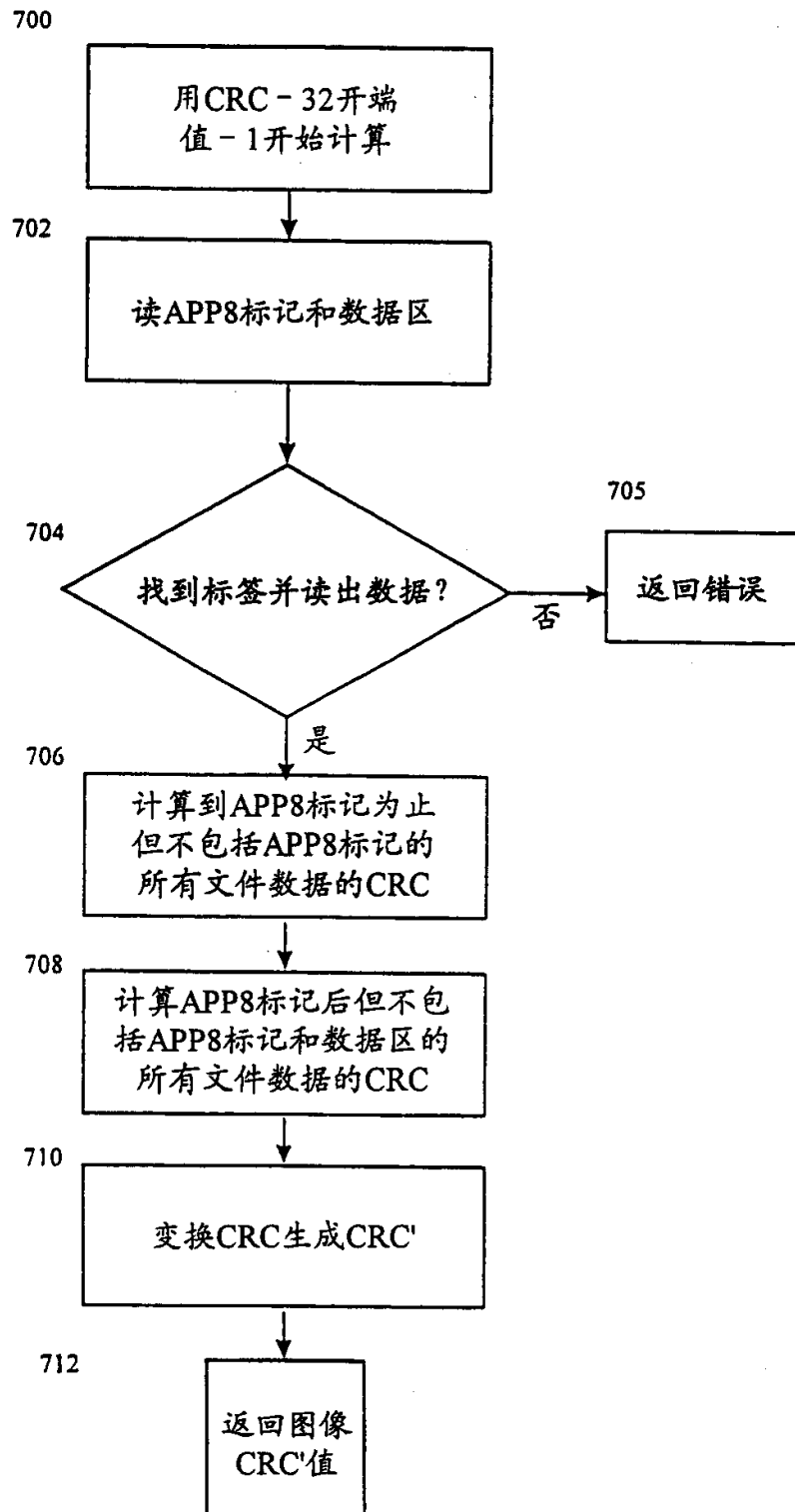


图 7

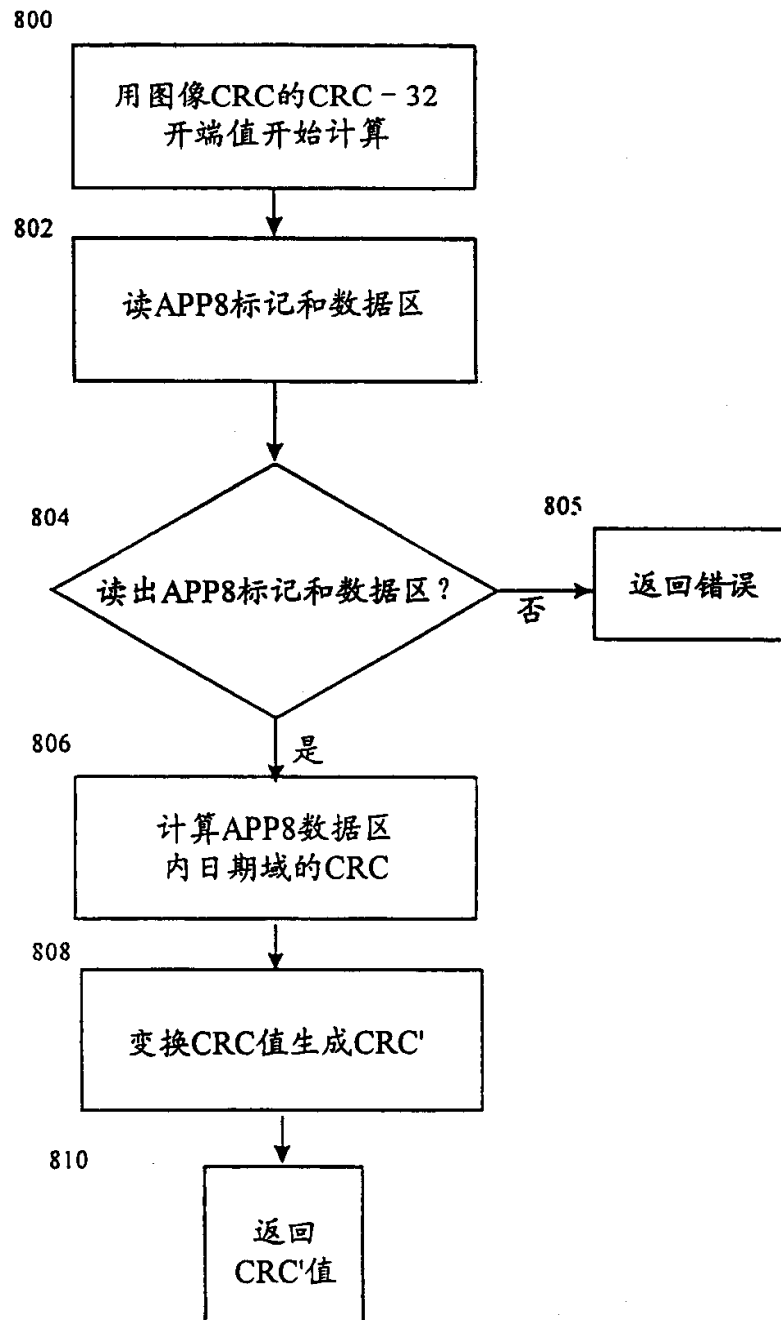


图 8