[19] 中华人民共和国国家知识产权局

[51] Int. Cl[7]
G06T 1/00
H04N 1/32

[12] 发明专利申请公开说明书

[21] 申请号 200380101183.6

[43] 公开日 2005 年 11 月 30 日

[11] 公开号 CN 1703722A

[22] 申请日 2003.10.8
[21] 申请号 200380101183.6
[30] 优先权
[32] 2002.10.9 [33] EP [31] 02079247.9
[86] 国际申请 PCT/IB2003/004400 2003.10.8
[87] 国际公布 WO2004/034325 英 2004.4.22
[85] 进入国家阶段日期 2005.4.8
[71] 申请人 皇家飞利浦电子股份有限公司
地址 荷兰艾恩德霍芬
[72] 发明人 D·K·罗伯特斯

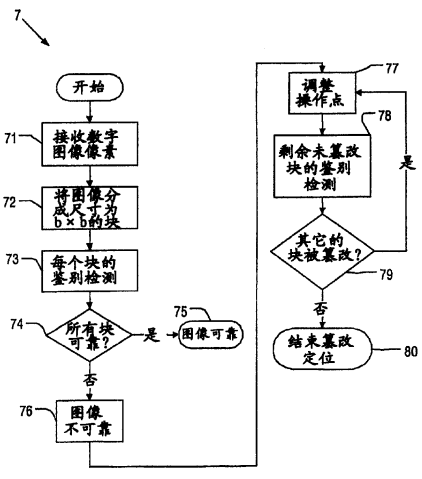
[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 傅 康 王忠忠

权利要求书 4 页 说明书 17 页 附图 7 页

[54] 发明名称 图像篡改的定位

[57] 摘要

本发明提供一种用于校验媒体内容可靠性的方法和装置(8)。根据一实施例提供一种用于数字图像鉴别的准确的篡改定位。典型地，一可疑图像被划分为块。对于每个块，通过计算图像内容的特性生成鉴别位，随后将所述特性的临界值设为‘0’或‘1’。将可疑图像的鉴别位和原始图像的那些鉴别位进行比较。如果存在不符，并且该内容确实被篡改过，那么检测为篡改。由于可允许操作，如压缩，引起的不符被称为错误警告，其应当是被避免的。一所谓的 POC(接收器操作特性)曲线给出了检测概率和错误警告概率之间的关系。更适宜地，用于确定鉴别位的临界值在 ROC 曲线上代表一操作点。根据本发明的一个实施例，最初选择与一低错误警告概率相应的操作点。为了更精确地识别篡改图像区域，利用不同的操作点对相邻块重复执行鉴别决定。该过程继续直到不再进一步发现有篡改块为止。因此提出了一种改进的篡改定位，其对例如通过安全摄像机捕获的图像进行鉴别是有价值的，并且可用于定位任何篡改区域，借此这些图像的值被增加作为例如一法庭中的证据。



ISSN1008-4274

1. 一种验证媒体内容可靠性的方法，所述方法包括步骤：

通过将媒体内容的连续部分中的媒体内容的特性和第二临界值进行比较从所述媒体内容中提取第一鉴别位序列，

接收第二鉴别位序列，所述接收的序列是通过比较媒体内容的所述特性和第一临界值从媒体内容的原始版本中提取出来的，并且

如果接收的第二鉴别位序列与提取的第一鉴别位的序列相匹配则宣布媒体内容可靠，

其特征在于从媒体内容中提取鉴别位的步骤包括根据接收的鉴别位设置第二临界值，从而在所述第一鉴别位序列中提取的鉴别位和在所述第二鉴别位序列中相应接收的鉴别位不匹配的概率和使用第一临界值进行所述提取相比降低了。

2. 如权利要求 1 所述的方法，其中当验证所述媒体内容的可靠性时，错误警告率被减小了。

3. 如权利要求 1 或 2 所述的方法，其中从所述媒体内容中提取鉴别位的步骤包括根据接收的鉴别位控制临界值使得提取的鉴别位和相应接收的鉴别位相匹配的概率高。

4. 如权利要求 1 到 3 中任意一个所述的方法，进一步包括在基于当前不匹配鉴别位在提取鉴别位的步骤期间控制第二临界值，由这样一种方式，根据先前迄今发现的不匹配鉴别位调整可靠性决定处理，从而改进了对所述媒体内容中不可靠部分的定位。

5. 如权利要求 1 到 4 中任意一个所述的方法，包括如果接收的第二鉴别位序列和提取的第一鉴别位序列不匹配，那么宣布媒体内容为整体篡改了。

6. 如权利要求 5 所述的方法，其中接收的第二鉴别位序列和提取的第一鉴别位序列之间的不匹配位包括关于所述媒体内容中至少一第一部分的定位的信息，所述方法进一步包括步骤：

识别和/或标记所述媒体内容中篡改部分的定位以使至少一个篡改部分可见。

7. 如权利要求 6 所述的方法，进一步包括随后的利用一更改过的第二临界值重复执行提取步骤的阶段。

8. 如权利要求 7 所述的方法，其中对所述媒体内容中邻近所述媒体内容的被识别为篡改的部分的部分单独执行所述提取步骤。

9. 如权利要求 1 所述的方法，进一步包括重复执行提取步骤的阶段，根据鉴别位被提取的部分和发现鉴别位与接收的鉴别位不匹配的部分之间的距离控制第二临界值。

10. 如权利要求 1 所述的方法，其中片断是块且媒体内容是数字图像，

其中提取步骤包括对每个块单独做出鉴别决定，并且第二临界值最初从一低错误警告操作点获得，

其中宣布步骤包括如果没有块被宣布为篡改，则宣布该图像可靠，或者如果至少发现一个篡改块，则宣布图像为整体不可靠，

其中邻近那些篡改块的块被宣布比不相邻的块具有较高的篡改概率，并且为剩余块选择新的操作点，所述剩余块是在先前的运行中没有被宣布为篡改的块，重复执行鉴别决定直到不再识别出有篡改块。

11. 如权利要求 10 所述的方法，进一步利用决定边界的改动以便将操作点移动到具有较大检测概率的位置上。

12. 如权利要求 10 或 11 任意一个所述的方法，进一步包括通过标记图像中的篡改块确定篡改图像区域的完整大小或形状。

13. 如权利要求 1 所述的方法，所述第二临界值的所述调整包括根据由相邻决定给出的上下文信息调整操作点或决定边界或已有概率。

14. 如前述权利要求中任一个所述的方法，其中根据以下公式调整第二临界值:

$\lambda_i = \alpha \lambda_1 + (1-\alpha) \lambda_2$，

其中 $\lambda_1 = 1$ 且 $\lambda_2 > 1$ 是决定临界值，α 值被给出为:

$$\alpha = \left(\frac{n}{m}\right)\left(\frac{d - r_m}{d-1}\right)$$，且 $r_m = \min(r, d)$

其中 n 是邻近块 i 的被标记为篡改的块的数目，m 是邻近块 i 的块的总数，r 是以块单位计算的块 i 与最近的篡改块的距离，而 d 是设置篡改块周围的多宽被提高可疑性的最大距离。

其中利用新的第二临界值 λ_i 重新评估后来的鉴别决定，和

如果在后来的鉴别决定中进一步宣布有块篡改，则重复执行调整第二临界值和重新评估块可靠性的步骤，直到不再识别出有块篡改。

15. 如前述权利要求中任一个所述的方法，其中用于确定鉴别位的第二临界值表示为 ROC 曲线上的一个操作点。

16. 根据权利要求 1 所述方法在多媒体鉴别决定中的应用，其中所述多媒体包括图像或视频和/或音频数据。

17. 如权利要求 16 所述的应用，其中所述多媒体鉴别决定被应用于监视系统。

18. 如权利要求 16 或 17 任意一个所述的应用，其中多媒体鉴别决定中决定边界的调整是基于上下文信息进行的。

19. 如权利要求 18 所述的应用，所述上下文信息是基于所述多媒体的篡改定位期间已经被确定为篡改的区域的接近性。

20. 一种用于通过执行权利要求 1 所述方法验证媒体内容可靠性的装置（8），所述装置包括:

通过将媒体内容的连续部分中的媒体内容的特性和第二临界值进行比较而从所述媒体内容中提取第一鉴别位序列的装置（80），

用于接收第二鉴别序列的装置（81），其中通过比较所述媒体内容的特性和第一临界值从媒体内容的原始型式中提取所述接收序列，以及

如果接收的第二鉴别位序列与提取的第一鉴别位序列相匹配则宣布媒体内容可靠的装置（82），

其特征在于从媒体内容中提取鉴别位的装置包括根据接收的鉴别位设置第二临界值的装置（83），从而在所述第一鉴别位序列中提取的鉴别位和在所述第二鉴别位序列中相应接收的鉴别位不匹配的概率和使用第一临界值进行所述提取相比降低了。

21. 一种其上收录有计算机程序的计算机可读媒体（9），所述程序通过执行权利要求 1 所述方法来验证媒体内容可靠性，并且由计算机（94）处理，该计算机程序包括:

通过将位于媒体内容连续部分中的媒体内容的特性和第二临界值进行比较以从所述媒体内容中提取第一鉴别位序列的第一代码段（90），

用于接收第二鉴别位序列的代码段（91），通过比较所述媒体内

容的特性和第一临界值从媒体内容的原始型式中提取所述接收序列，并且

如果接收的第二鉴别位序列与提取的第一鉴别位序列相匹配，则宣布媒体内容可靠的第三代码段（92），

其特征在于从媒体内容中提取鉴别位的代码段（90）包括根据接收的鉴别位设置第二临界值的代码段（93），从而在所述第一鉴别位序列中提取的鉴别位和在所述第二鉴别位序列中相应接收的鉴别位不匹配的概率和使用第一临界值进行所述提取相比降低了。

图像篡改的定位

技术领域

本发明通常属于数字成像领域，具体涉及数字图像和视频的鉴别，更进一步说涉及用于鉴别目的的图像篡改的识别和定位。

背景技术

在数字形式下简易地编辑和改变图像及视频的要求刺激了对那些能够鉴别原始、未改变内容的装置的需要。当判断一个图像已经被改变过时，而还希望具有指示图像区域被改动过的标记。

由于一些图像改变是可接受的事实，比如那些由于衰减压缩引起的改变，因此使鉴别问题复杂化了。这些改变可能引起图像质量的轻微退化，但不会影响图像的判读或预期使用。结果就是采用密码术的传统鉴别技术不再适合了，典型地是当发生篡改时这些方法只能解释图像的仅仅一位的改变。

通常，有两个稳健的方法，也就是，非位感应，图像鉴别，即半易碎水印，以及公知为“指纹”的稳健数字签名。所有这些方法基本上都是根据一组从可疑图像中计算获得的位和从原始图像内容中计算得到的相应位组之间的比较来实现的。通过计算该图像像素值的某些特性，S，并随后将 S 的临界值置为 ‘0’ 或者 ‘1’ 位，从可疑图像中生成鉴别位。计算的特性取决于所使用的水印或指纹方案。典型地，将一个图像划分为多个块并为每个块生成一个鉴别位。典型的块尺寸的例子为 16 × 16 像素或 32 × 32 像素。由于一特殊位中的错误可以与一特殊图像区域的改变相关，因此将数字图像细分到块能够为图像改变提供定位。

对于每个原始鉴别位，必须确定可疑图像是否有可能生成相匹配的鉴别位。这等于判断是否相应的图像块被鉴别或改变过。如果判断一个块被篡改过，并且该图像内容确实被改动过，那么将其称为一次检测。如果，从另一方面讲，当一个块的内容实际上仅仅是经历了一些可允许操作（例如，压缩）而被判断为被篡改时，该判定是错误的，并被称为一错误警告。

原有系统通过将从可疑图像中生成的位与原始鉴别位进行比较来

产生鉴别决定。一个更加成熟的方法是使用‘软确定’信息。在这种情况下利用从可疑图像中计算得到的特性 S 的非临界值判断可靠性。如果位于临界值错误一方的 S 值接近临界值，那么其可能仍然被判断为可信，该临界值用于生成和原始鉴别位相匹配的位。这些为可允许图像操作提供了更多稳健性，从而减少了错误警告发生的可能性。

发明内容

本发明的一个目的是改进改变的图像区域的定位。因此，本发明要解决的一个问题就是提出一种新的，具有改进的篡改定位的图像鉴别方法和装置。本发明克服了现有技术中的上述鉴别缺陷并根据附加权利要求提供的特征至少解决了上述鉴别问题。

根据本发明的一个方面，公开了一种用于验证媒体内容可靠性的方法，装置，和计算机可读媒体。

根据本发明的一个方面，提出一种用于验证媒体内容可靠性的方法。该方法包括以下步骤，通过将媒体内容的连续部分中的媒体内容的特性与一第二临界值进行比较而从媒体内容中提取第一鉴别位序列。其还包括接收第二鉴别位序列，其中通过将所述媒体内容的特性和第一临界值进行比较从媒体内容的一原始文本中提取接收的序列。根据该方法，如果接收的第二鉴别位序列和提取的第一鉴别位序列相匹配，那么该媒体内容被宣布可靠。该方法的特征在于从媒体内容中提取鉴别位的步骤包括根据接收的鉴别位设置第二临界值，从而和利用第一临界值的提取相比，从所述第一鉴别位序列中提取的鉴别位和从所述第二鉴别位序列中相应接收的鉴别位之间不匹配的概率降低了。

根据本发明的另一个方面，通过各附加的独立权利要求提出一种用于验证媒体内容可靠性的装置，该装置执行根据本发明一个方面的上述方法。

根据本发明的再一个方面，通过各附加的独立权利要求提出一种其上包含有计算机程序的计算机可读媒体，该程序由计算机处理，并且其通过执行权利要求 1 所述方法来校验媒体内容可靠性。

根据本发明的一个实施例，在多媒体内容的鉴别决定中使用“上下文”信息，如数字图像或视频。将多媒体内容划分成段，如块，并且为每个块生成“上下文”信息。更具体地，被宣布为篡改的块的数

目和位置，影响着对其他块是否可能被篡改的判定。例如，一篡改块的相邻块和其他较远的块相比具有更大的可疑性。根据本发明的一个实施例，通过调整一个所谓的 ROC 的曲线（接收器操作特性）上的操作点来将这种上下文信息和鉴别决定相结合，这在以下有更详细的阐述。

根据本发明的一个实施例，图像的鉴别校验包括以下步骤：

1. 利用一低错误警告操作点针对每个块独立做出鉴别决定。

2. 如果没有块被宣布为篡改，那么该图像被确定为可靠。

3. 如果发现了一个或多个篡改块，那么认为该图像为整体不可靠。这就是说与那些篡改块相邻的块也有可能被篡改，并且所有其他图像块均可以被假设可能是可靠或篡改。基于以上认知，针对每个块的鉴别决定选择新的操作点。

4. 利用新的决定边界对所有未被宣布为篡改的块的鉴别决定重新进行评估。

5. 如果进一步宣布有块篡改，那么重复执行调整决定边界的操作和重新评估块可靠性的操作。如此继续直到不再识别出有块篡改。

改变决定边界可被使用将操作点移动到具有较大检测概率的位置上。这可能会进一步发现篡改块，并因此有助于确定篡改图像区域的整个大小和形状。

本发明具有高于现有技术的优点，其提出了一种在数字图像的鉴别期间改进的篡改区域定位。

本发明适用于如上所述的不考虑鉴别位是否构成了水印或指纹的情况。

附图说明

以下参考附图所做的有关本发明实施例的描述将使本发明进一步的目的、特征以及优点更显而易见，其中

图 1 是一典型的监视系统的示意图，

图 2 是一表示与篡改检测及错误警告概率有关的 ROC 曲线示例的曲线图，

图 3 是一可靠的非篡改示例图像，

图 4 是一表示图 3 中的示例图像在有区域被篡改时的图像，

图 5 是一表示根据现有的篡改判断技术判断图 4 中的篡改示例图

像有块篡改时的图像，

图 6 是一表示根据本发明判断图 4 中的示例图像有块篡改时的图像，

图 7 是一根据本发明一个方面所述的方法实施例的流程图，

图 8 是一根据本发明的另一方面的实施例的示意图，

图 9 是一根据本发明的再一方面的实施例的示意图，

图 10 是一表示两种不同假设情况下的两个条件概率密度函数的曲线图，

图 11 是说明一 JPEG 图像的错误警告概率的曲线图，以及

图 12 是说明对于每 32 × 32 像素块 1 指纹位的篡改检测概率的曲线图。

具体实施方式

依靠参照一监视系统所描述的实施例对本发明做以下详细描述。但是，本发明并不意味着局限于那些参照所述监视系统描述的典型实施例，并且在附加的独立权利要求范围内本领域技术人员很容易知道其改进和其他应用。

图 1 说明了一典型监视系统 1 的布局。其通常由以下部件组成:

-至少一视频摄像机 10，其具有一通常为模拟格式的视频输出 11，所述模拟格式如 PAL 或 NTSC，

--一数字记录设备 12，其接收来自多个摄像机 10 的视频输入并应用衰减压缩，和

--一提供存储及检索的计算机网络 13，以及

-用于压缩视频的鉴别装置 14

监视系统 1 采用各种压缩方法，包括时间空间技术（如 MPEG），以及静止图像技术（如 JPEG，ADV601）。当采用静止图像压缩技术时，通过例如每 5 秒钟仅保持一个图像，而在时间方向上实现压缩。注意通过数字记录设备 12 进行的衰减压缩产生的视频的失真对于篡改而言必须不能发生错误。

媒体内容篡改的预计类型是数字图像中的像素替换，该类型可以通过本发明所披露的实施例检测及精确定位。例如，可以是通过用“背景”内容替换来除去一个人，所谓背景内容可能是从一较早/较晚的没有人物的图像中复制来的，因此所考虑的图像的所有内容表现为是

正确的，或者任何其他改变所述图像视觉内容的像素修正都是正确的。但是，可允许操作，如为节约存储空间而进行的图像压缩，不会被划分为篡改一类。

篡改区域的最小可检测尺寸的方针为人脸在其中可以被分辨的最小尺寸。这个尺寸对于 PAT/NTSC 视频内容而言大约是 35 像素宽和 50 像素高。

通常，如上所述，通过将从可疑图像获得的鉴别数据和从原始图像获得的相应数据进行比较来执行篡改检测。这可以被分解为两个子问题：

-如何生成适当的鉴别数据，以及

-如何将原始图像的鉴别数据传送到检测可靠性的系统中的点。

对于摄像机 10，它不知道记录设备 12 在压缩期间是否会废除图像。因此必须生成并传送鉴别数据从而使每个图像可以被独立地鉴别，而不用参考在时间上位于其他任何点的图像。

此外，识别可允许的改动和恶意改动的能力一般可由术语半易碎性来表示。通常，根据这种易碎性的存在位置有两种可供选择的鉴别处理办法：

1. 半易碎水印，其中传送原始图像的鉴别数据是这样进行的：即在可允许的改动之后而不是篡改之后其能被正确地重现，以及

2. 半易碎数字签名，其中生成鉴别数据的过程是这样的：即数据对于可允许改动而言是不变量，但对于篡改不是。

半易碎水印通常生成一用于鉴别数据的固定的位模式，随后利用半易碎技术将其嵌入。可靠性检测包括提取水印位和将该水印位与嵌入的模式进行比较。篡改图像区域的位置由提取的鉴别位中的错误来表示。

嵌入位固定模式的使用易于生成明显可靠的篡改图像。例如，从不同的，但是可靠的图像中的同一位置复制的内容可能替换像素。仍然能够成功地提取水印位，并因此判断改动图像是可靠的。

通过生成基于图像内容的鉴别位可以提高安全性。这有助于防止上述的复制攻击的示例。如果水印位相关内容也具有篡改易碎性，那么这种方案即具有半易碎性水印的特性也具有半易碎签名的特性。如果，例如，鉴别数据和水印对于不同类型的图像改动是易碎的，那么

这种方法有助于表示发生了何种类型的篡改。

但是，半易碎水印仅仅能够保护那些被用于嵌入鉴别数据的图像特征（如像素或频率系数）。因此为了保护最重要的感知图像特征就要求把数据嵌入到那些特征中去。这可能会在保证水印不可见性方面产生困难。任何水印位均不能被被不可见地嵌入并且有效检测的图像材料（诸如平面内容）即使没有篡改也会导致位错误。因此没有办法将这些由于零水印容量引起的位错误与由于那些篡改引起的位错误区别开。由平面内容替换原始图像区域可能因此产生一明显可靠的篡改图像。

尝试通过“备份嵌入”来克服这种前述提到的问题。其中，利用两个在空间上相互分离的嵌入位置将每个水印位嵌入两次。但是，这无法保证备份位置也同样不存在零水印容量。对于由于给定嵌入能力的少数鉴别导致的篡改定位能力，或由于增加的嵌入位数目带来的可允许操作的不可见性和稳健性，多次嵌入每个鉴别位必须也具有否定的蕴含式。

通常，数字签名是一组概括图像内容的鉴别位。半易碎签名是这样产生的以便篡改的图像给出一组改变的概括位，但是仅仅经过可允许操作处理的图像是不可以的。一种非位感应类型的签名将被称作为指纹以便提供一种与密码数字签名的明显的区别，同时强调和其他应用之间的相关性。

计算指纹位的图像特征通常在可允许处理的稳健性、篡改的易碎性和计算成本之间选择最恰当的折中。这些特征的例子是 DC 值，时间，边缘，矩形图，压缩不变量，以及噪声模式投影。

通过将从可疑图像中生成的指纹和例如在摄像机中计算的原始指纹进行比较来校验可靠性。典型地，在单独的指纹位和图像位置之间存在直接的关系。例如可将图像分割成块并且对于每块获得一位。因此通过错误的特殊指纹位即可显示篡改图像域的位置。

但是，指纹位的数目和定位能力之间有一个折中。例如，一较小的块尺寸允许较优的篡改区域定位，但是每个图像有更多的块，并因此有更多的指纹位。

在摄像机中已经生成的原始图像的指纹后，还存在着传送这些指纹数据的问题，以便将其用于可靠性验证。

如上所述，一个可能性就是在图像中嵌入指纹位作为水印。水印为传送问题提供了解决方案。通过在图像中不可见地嵌入指纹，这个数据自动被图像携带。明显地，水印对于至少所有可允许的图像处理而言必须稳健。如上所述，如果该水印也是半易碎的，那么这可能有助于识别已经发生的篡改的种类。指纹位的内容相关特性也有助于防止水印内容从一个图像复制到另一个图像而呈现为是可靠的。

指纹防止了用于计算指纹位的图像特征的改动。这些特征可能和那些用于嵌入指纹作为水印的特征不同。对于不可见性和稳健性要求而言，这以最恰当的方式增加了嵌入位的灵活性，并且有助于避免半易碎水印可靠性方案遇到的零水印容量问题。

使用水印传送指纹数据的缺陷在于其有可能限制了篡改定位能力。一个足够稳健的水印一般将具有非常有限的有效载荷大小，其可能在指纹尺寸及定位能力方面带来无法接受的限制。

由于摄像机 10 和记录设备 12 之间的模拟电缆，与视频分开的传送指纹数据是不可能的。这要求在摄像机中生成的鉴别数据为了传送给记录设备必须被嵌入到视频信号本身中。因此形成水印的一种选择是与将图文数据嵌入到电视信号中的方式相同地直接将指纹数据嵌入到像素值中。安全摄像机已经使用这种数据通道传送摄像机参数，控制信息和音频。根据利用的视频线的多少，这些数据通道的数据携带能力可以远远高于水印。如果仅仅在过扫描区域即垂直消隐间隔中使用视频线，那么就可以保持嵌入数据的不可见性。

在将指纹数据以该方式嵌入之前对其进行加密是很重要的。没有加密，以相应于篡改图像的指纹替换原始指纹数据将使伪造看起来可靠。丢失或破坏的鉴别数据必须一直被认定为篡改。

应当基于图像的低频内容来计算指纹。为模拟链接提供复原是很必要的，它严重地限制了视频信号的带宽和损耗压缩，并明显删除了高频内容。

在可允许处理操作被很好地表征的应用中，这一认知可能被用于指纹计算。例如，对于 JPEG 量化是不变量的特性被用于形成指纹。但是，如上所述，由于在监视系统中使用的压缩方法具有极大的多样性，因此这种方法是不可能的。

再者，如上所述，摄像机 10 必须实时地为每个输出图像计算并嵌

入鉴别数据。如果对摄像机成本的影响被最小化，那么这严格约束了计算负担。

仅利用 DC 分量就可以形成低频和低复杂性指纹。图像被划分为块，而块的 DC 值之间的差异，即平均像素亮度，被用来形成指纹。使用 DC 差异为整个图像 DC 分量的变化（例如由于亮度改变带来的）提供不变量。由相邻块的 DC 值之间的差异获知每个块的图像内容与其邻近块内容之间的相关性。根据一特例，按如下公式为第 i 块生成指纹位 b_i：

$$s_i = \sum_{j=1}^{8} (DC_i - DC_j) \qquad (1)$$

如果 $s_i > 0$，$b_i = 1$，否则 $b_i = 0$，

其中 j 表示 8 个与块 i 邻近的块。

块的适当的尺寸和期望在其中进行篡改检测的图像特征的尺寸有关。较小的块具有较大的被检测为改变的可能性，但这是以增加指纹位数目为代价而进行计算和传送的。

检测可靠性最直接的方法是简单的原始和可疑鉴别位的位位比较。但是，由于可允许处理而必然存在一些位错误时，仅这样是不可能令人满意的。

解决这一问题的方法是基于这样的认识：那些由于可允许处理引起的位错误可能轻微地分布在整个图像中，而篡改引起的位错误可能集中在一个有限区域内。因此通过基于该位错误的一后处理操作可以从篡改中将可允许操作识别出来，所述后处理操作如误差松弛（error relaxation），或数学形态学。

通常，可靠性校验提出了比指纹计算更复杂的计算，因为它发生的相对频繁、非实时地、并具有一更强的可用计算平台。

可取的是将这种稳健性更加精密地构造进鉴别决定中而不是应用“事后”的后处理步骤来对可允许处理提供复原。这可以通过在比较可疑图像的指纹和原始指纹位期间使用‘软-确定’信息来实现。这防止了在 s_i 接近零时所表示篡改，并因此有可能发生由于可允许处理带来的指纹位错误。

根据一进一步的实施例，一单独块的鉴别决定被表示为假设值 H_0（即块的图像内容是可靠的）和假设值 H_1（即块的图像内容是篡改的）

之间的一种选择。假设理论的基础在附件中给出，其同样是本说明书的一部分。给定块值 s，其是根据等式 1 计算的，以及原始图像 b_{orig} 的指纹位，则具有最大可能性的假设被选择：

如果 $Pr[H_0 | b_{orig}, s] > Pr[H_1 | b_{orig}, s]$，选择 H_0

但是，根据贝叶斯（Bayes）定理：

$$Pr[H_0 | b_{orig}, s] = \frac{p_{S|H_0 b_{orig}}(s) Pr[H_0]}{p_S(s)}$$

且对于 H_1 也是如此，所以确定法则变为：

如果 $\frac{P_{S|H_0, b_{orig}}(s)}{P_{S|H_1, b_{orig}}(s)} > \frac{Pr[H_1]}{Pr[H_0]}$，选择 H_0 （2）

为每个假设的已有概率赋值是很困难的，因为其相当于表明图像的什么部分被篡改了。因此 Neyman-Pearnon 个人确定法则（如附件中的说明）更加合适。这种方法使针对固定的对于篡改而言是错误的可允许处理的‘错误警告’概率而检测到的篡改概率最大化。事实上，这一结果导致已有概率被一临界值 λ 替换，其被设置为用于实现期望的错误警告率：

如果 $\frac{P_{S|H_0, b_{orig}}(s)}{P_{S|H_1, b_{orig}}(s)} > \lambda$， 则选择 H_0 （3）

如果假设 H_1 为真，那么我们不知道替换内容并且可仅假设等式（1）的结果一般关于图像内容分布，即 $Ps|H_0, b_{orig}(s) = Ps(s)$.

如图 10 所示，从一组图像中估算概率密度函数(PDF) Ps(s)，并使其接近拉普拉斯分布。

如果假设 H_0 为真，那么对于原始图像，等式 1 的结果 S_{orig} 是由 b_{orig} 的值给出的已知符号。因此 S_{orig} 的分布为 Ps(s)的单变模式，也就是指数。那么可允许处理操作引起一错误 E，导致观察值 $S = S_{orig} + E$。应当

针对图像所进行的最严格的可允许处理估算 E 的分布，如最低 JPEG 质量因数。典型地高斯分布为 E 的 PDF 提供最合理的近似。最后，假设独立于 S_{orig} 和 E，下述卷积给出了用于假设测试所需的 PDF:

$$p_{S|H_0,b_{orig}}(s) = \int_{-\infty}^{\infty} p_{S_{orig}}(s-e)p_E(e)de$$

图 10 表示用于与质量因数为 50 且 $b_{orig}=1$ 的 JPEG 压缩相应的 E 的情况下的 PDF 的曲线 101。可以注意到由于 E 该曲线背离了指数形状。这使 S 的非零概率是负的，并因此模拟由于可允许处理引起的指纹位错误。

从图 10 看出，无论临界值 λ 的值是多少，PDF 仅在一个点上相交。因此假设测试降低为一个块值 S 的简单的临界值测试。对于 $b_{orig}=1$ 的临界值 S_T 满足:

$$p_{S|H_0,b_{orig}=1}(s_T) = \lambda p_{S|H_1}(s_T)$$

并且，利用对称性，对于 $b_{orig}=0$ 的临界值为 $-S_T$。

图 11 表示一 JPEG 图像的错误警告概率。从曲线 111 中可以很清楚地看出所期望的是具有较小峰值 PDF 的特性 S。这将减少由于 E 所引起的对位临界值的涂抹，从而获得较少的由于可允许处理引起的指纹位错误。

注意上述推导是假设的 S 值对于不同的块均是独立并且均匀分布的。实际上不是一直如此，邻近块 S 的值之间存在一些相关性。不过，从下面给出的结果中可以看到，该方法是非常有用的。

上述假设测试构架的优点在于其考虑了原始指纹位中的错误概率。这通过根据传送通道的位错误率使 b_{orig} 的值成为一种随机变量分布而实现。

本发明进一步的优点在于通过调整操作点，即临界值 λ 有可能改进篡改区域的定位。通常 λ 被设置来实现期望的低错误警告率。但是，一旦一个或多个块被判定为篡改，那么该图像作为一个整体被认为是

不可靠，并且每个独立的块均被认为可能是篡改的或可靠的。这说明要使用相同的现有概率，即 λ=1，重新评估所有块的可靠性决定。这种方法要进一步考虑篡改块的空间分布。例如，一邻近几个篡改块的块也可能被篡改。可以通过修改现有概率，或等价地，修改 λ 的值，来表示上述观点。试验表明这些操作点的调整和可靠性决定的重新评估有助于极为精确地推断出篡改区域的大小和形状。

准确设置 S 值的哪个范围将被分为可靠的和篡改的，其固定错误警告和检测概率。根据确定边界的位置，可以在检测和错误警告概率之间实现不同的折中。这经常在接收器操作特性（ROC）上有所显示。附图 2 中的曲线图 20 显示了一 ROC 曲线的典型形状。

在图像鉴别中，希望只有较少数的图像被实际上篡改，因此具有一低错误警告概率是很重要的，否则将有大量的可靠图像被宣布为发生了篡改。因此通常选择 ROC 曲线上的操作点以实现可接受的较小的错误警告率。

根据本发明的一个实施例，如图 7 中所示，通过在上述 ROC 曲线上调整操作点而将所述上下文信息并入可靠性决定。根据本发明的该实施例，提出了一种用于检测数字图像的鉴别方法 7，其中方法 7 包括如下步骤。

在步骤 71 中接收一数字图像。方法 7 的目的是如果图像可靠就建立图像，如果图像不可靠则准确定位一个或多个篡改域的空间位置。为了这个目的，根据步骤 72 将图像划分为块，如 b×b 像素大小。在步骤 73 中利用 ROC 曲线上的一低错误警告操作点独立地为每个块做出鉴别决定。在图 2 所示的典型 ROC 中，在曲线图 2 的 ROC 曲线上用“X”21 标出了满足这些条件的示意性操作点。

如果在步骤 74 中没有宣布有块篡改，那么在步骤 75 中认定该图像可靠。如果发现了一个或多个篡改块，那么如步骤 76 所示认定该图像作为一个整体不可靠。这就是说邻近那些在步骤 73 中被检测出来为篡改块的块也有可能被篡改，并且所有其他图像块可以均被假设可能是可靠或篡改。基于这种认知，在步骤 77 中在 ROC 曲线上为每个残余的块的鉴别决定选择新的操作点。在步骤 78 中使用新的决定边界重新评估还没有被宣布为篡改的所有块的鉴别决定。

如果在步骤 78 中进一步宣布有块篡改，那么在步骤 79 中根据做

出的决定再次执行调整决定边界和重新评估块可靠性的步骤。循环执行直到不再发现有篡改块为止。

可在重复步骤 77 中使用决定边界的改动以将操作点移动到具有较大检测概率的位置上。这可能会进一步发现篡改块，并因此有助于确定篡改图像区域的完整尺寸和形状。

如图 2 中所示，选择能够提供低错误警告概率的操作点也可以减小检测概率。这意味着有很多篡改块将不会被检测。假设篡改区域跨越若干鉴别块，那么所有改动块没有被检测的概率就小得多，因此图像是不可靠的，这样的事实仍将是显而易见的。

虽然低错误警告操作点可以达到一良好的检测图像是否已经被改动的概率，但是为图像改动的定位带来了更严重的蕴涵。单独块的低检测概率导致要补充检测哪些图像区域已经发生了变化。这在附图中做如下阐明：图 3 表示原始图像 30，图 4 是改动型式 40；图 5 表示鉴别块被判断为篡改的图像 50（在图像左上角区域内的块）。

在图 5 中可以看到大量图像块被判断为篡改，因此很明显该图像是不可靠的。但是，图 3，4 和 5 之间的比较表明了篡改图像区域的补充检测；改动图像区域的完整尺寸和形状不是非常明显。

如图 4 所示对应用方法 7 的实例，其结果表示在图 6 的图像 60 中。和图 5 所示的检测结果相比，完整得多的篡改区域的覆盖范围和定位是非常清楚的。

利用决定构架，如附件中所述，本发明可以被应用于如下所述的另外的实施例中。

选择具有可接受的低错误警告率的操作点 λ_0。利用该决定临界值评估所有图像块的可靠性。

如果宣布没有块篡改，那么该图像被认定为可靠。

如果发现一个或多个篡改块，那么对于所有其他块 i 确定新的操作点 λ_i。决定临界值的这种调整将考虑发现的篡改块的数目，以及它们和块 i 的接近性。

有许多可用的调整决定临界值的算法。一个非限制示例为：

$$\lambda_i = \alpha \lambda_1 + (1-\alpha) \lambda_2,$$

其中 $\lambda_1=1$，这代表相等的已有概率，$\lambda_2>1$，其给出较高的检测概率，α 值由以下公式给出：

$$\alpha=\left(\frac{n}{8}\right)\left(\frac{d-rm}{d-1}\right)$$ ，且 $r_m=\min(r,d)$

其中 n 是被标记为篡改的邻近块 i 的典型 8 个块的数目，r 是最近篡改块与块 i 的距离（在块单元中），而 d 是设置篡改块周围的多宽应该被提高可疑性的某一最大距离。

利用新的决定边界 λ_i 重新评估鉴别决定。

如果进一步宣布有块篡改，则重复调整决定边界和重新评估块可靠性的步骤，该过程继续直到不再识别有篡改块为止。

该进一步的实施例的典型描述清楚地表明调整操作点就相当于调整篡改块的已有概率。这反过来由块的上下文所证实，即其关于其他篡改区域的位置。

图 8 表示本发明另一方面的进一步的实施例，其中验证媒体内容可靠性的装置 8 包括执行根据本发明一个方面的鉴别方法的装置。

更详细地，装置 8 是一用于验证媒体内容可靠性的装置。装置 8 包括通过将位于媒体内容连续部分中的媒体内容的特性和第二临界值进行比较以从媒体内容中提取第一鉴别位序列的第一装置 80。此外，装置 8 还包括用于接收第二鉴别位序列的装置 81，其中通过比较所述媒体内容的特性和第一临界值，从媒体内容的原始型式中提取所述接收序列。此外，装置 8 包括装置 82，如果接收的第二鉴别位的序列与提取的第一鉴别位的序列相匹配则该装置宣布媒体内容可靠。装置 8 的特征在于从媒体内容中提取鉴别位的装置 80 包括根据接收的鉴别位设置第二临界值的装置 83，从而和使用第一临界值进行所述提取相比，在第一鉴别位序列中提取的鉴别位和在第二鉴别位序列中相应接收的鉴别位不匹配的概率降低了。装置 8 例如结合在鉴别装置 14 中，如图 1 所示。

图 9 所示为本发明的另一个实施例，根据本发明另一个的方面，提出一种其上收录有计算机程序的计算机可读媒体 9，该程序由计算机 94 处理，用于通过执行根据本发明一个方面的方法来校验媒体内容的可靠性。该计算机程序包括几段用于该目的的代码。更详细地，计算机可读媒体 9 上的计算机程序包括用于通过将位于媒体内容连续部分中的媒体内容的特性和第二临界值进行比较以从媒体内容中提取第一鉴别位序列的第一代码段 90。再者该计算机程序包括用于接收第二鉴

别位序列的代码段 91，其中通过比较所述媒体内容的特性和第一临界值从媒体内容的原始型式中提取所述接收的序列。此外，该计算机程序具有一代码段 92，如果接收的第二鉴别位序列与提取的第一鉴别位的序列相匹配则该代码段宣布媒体内容可靠。计算机程序的特征在于从媒体内容中提取鉴别位的代码段 90 包括根据接收的鉴别位设置第二临界值的代码段 93，从而和使用第一临界值提取相比，在第一鉴别位序列中提取的鉴别位和在第二鉴别位序列中相应接收的鉴别位不匹配的概率降低了。

以上计算机程序在例如鉴别装置 14 中运行，如图 1 所示。

可以通过鉴别系统检测篡改的概率和只有一个可允许图像处理被应用时的错误警告概率来衡量该系统的执行情况。只有少数出版物给出过上述信息，通常仅仅是给出一个表示该鉴别方法的示例图像。具体情况下很难评估检测概率，因为它需要对大量图像进行篡改，且以可信方式人工替换图像部分是很耗费时间的。

为克服这一问题，已经通过将第二非相关图像的图像内容混入图像中进行测试的自动处理来评估了检测率。利用不同的测试图像、不同的篡改位置以及不同的替换图像内容执行多次试验。为了获得有关本发明鉴别方法的性能的整个画面，针对不同大小的篡改区域重复执行整个测试。

图 11 和 12 表示作为决定临界值 s_T 的函数的利用该‘模拟篡改’度量的错误警告和检测概率。本结果是针对每 32 × 32 像素块 1 位指纹及 JPEG 质量因数 50 的可允许处理做出的。图 11 表示错误警告概率在 s=0 的指纹位临界值附近显现预期变化，转变锐度是由于对于 JPEG 压缩特性 s 具有较高的稳健性的缘故，因此引起指纹错误的可允许处理具有小概率。图 12 中曲线 121 和 122 表示实验中发现的针对篡改区域两种不同尺寸（分别为 64 × 64 和 100 × 100）的检测概率。很明显地，对于较好的检测率，要求指纹块的尺寸小于希望检测的篡改区域的最小尺寸。

也可以利用在先前部分中获得的概率分布在理论上评估鉴别系统的性能。单独块的检测和错误警告概率为：

$$\Pr(D)=\int_{-\infty}^{s_T} p_{S|H_1}(s)ds=\int_{s_T}^{\infty} p_{S|H_1}(s)ds$$

$$\Pr(FA)=\int_{s_T}^{\infty} p_{S|H_0,b_{orig}=1}(s)ds=\int_{s_T}^{\infty} p_{S|H_0,b_{orig}=0}(s)ds$$

假设个体块决定是独立的，整个图像的错误警告概率被评估为：

$$\Pr(\text{错误警告})=1-(1-\Pr(FA))^N$$

其中 N 为图像中的指纹块数目。这被绘制为图 11 中的曲线 112 并且可以看到显示出和实验结果 111 较好的对应。这证明在实际使用中应用理论方法计算 S_T 的值是正当的，其中需要太低以至于不能在合理的时间内被模拟的错误警告率。

整个图像的检测概率同样可以由以下公式评估：

$$\Pr(\text{检测})=1-(1-\Pr(D))^M$$

但是，设定 M 值，即篡改块的数目，是有问题的，因为它取决于关于指纹块的篡改区域的大小及形状。在图 12 中由以下设置评估检测概率：

$$M=n^2/b^2,$$

其中篡改区域是一 n×n 像素的块，并且利用 b×b 像素的块形成指纹。曲线 123 和 124 表示了对于两种不同尺寸（分别为 64×64 和 100×100）的篡改区域的理论结果。这可以看作是给出了和试验结果的合理匹配，因此是设置决定临界值时有用的检测率评估方法。

上述公开总的来说就是描述了用于安全摄像机视频鉴别的指纹解决方案。基于块 DC 差异的指纹在压缩稳健性、篡改灵敏度和计算成本之间给出了较好的协调。再者，公开了一种可靠性验证的假设测试方法。其具有很多的优点，如容许由可允许处理引起的指纹位错误；在接收的原始指纹中处理位错误的能力；通过已有概率的调整改进篡改的定位。但是，附带的权利要求中限定的这种安全摄像机解决方案仅仅是本发明的一个非限制性实施例。此外，上面借助安全摄像机所述的实施例同样是非限制性实施例。

最后，概括上述，提出了一种用于数字图像鉴别的准确的篡改定位。典型地，将可疑图像划分为块。对于每个块，通过计算图像内容的特性生成鉴别位，随后将所述特性的临界值设为‘0’或‘1’。将可疑图像的鉴别位和原始图像的鉴别位进行比较。如果存在不符，并且该内容确实被篡改过，那么检测为篡改。由可允许操作，如压缩，

引起的不符被称为错误警告，其应当是被避免的。所谓的 ROC(接收器操作特性)曲线给出了检测概率和错误警告概率之间的关系。用于确定鉴别位的临界值在 ROC 曲线上代表一操作点。根据本发明的一个实施例，首先选择与一低错误警告概率相应的操作点。为了更精确地识别一篡改图像区域，利用不同的操作点对相邻块重复执行鉴别决定。该过程继续直到不再进一步发现有窜动块为止。因此提出了一种改进的篡改定位，其对例如通过安全摄像机捕获的鉴别图像进行鉴别是有价值的，并且可用于定位任何篡改区域，借此这些图像的值被增加作为例如一法庭中的证据。

注意调整 ROC 曲线上的操作点，和根据邻近决定重新评估决定的概念不仅在图像或视频或音频鉴别中有意义，其同样可用于必须执行相互相关决定的其他领域。

本发明以上方面的用途和应用领域是多种多样的并包括如上述的应用在监视摄像系统领域中的典型。

以上参考特殊实施例描述了本发明。但是，除了以上优选实施例之外的其他实施例同样可能落在所附权利要求的范围内，例如和以上描述不同的用于生成存储的验鉴别信息的方法，通过硬件或软件执行上述方法，等等。

再者，在该说明书中使用的术语“包括”并不把其他元素或步骤排除在外，术语“一”和“一个”并不把实现权利要求中所述的几个单元或电路的功能的多个和单个处理器或其他单元排除在外。

附件一假设测试

给出针对可疑图像块计算的特性 S 的值，如果相比该块可靠(H_0)的假设而言其具有较大的概率，那么选择该块篡改（H_1）的假设：

如果：$Pr(H_1/S=s)>Pr(H_0/S=s)$，选择 H_1

按照 S 的概率密度函数将其扩展，每个假设的已有概率为：

如果：$\frac{p(s/H_1)Pr(H_1)}{p(s)}>\frac{p(s/H_0)Pr(H_0)}{p(s)}$，选择 H_1

重新整理为：

如果：$\frac{p(s/H_1)}{p(s/H_0)}>\frac{Pr(H_0)}{Pr(H_1)}$，选择 H_1

这种决定处理的困难在于设置已有概率的值，$Pr(H_1)$（任何给出图像均篡改的概率），和 $Pr(H_0)$（任何给出图像均可靠的概率）。不可能知道这些概率，因此可以通过值λ代表它们的比率：

如果：$\frac{p(s/H1)}{p(s/H0)}>\lambda$，选择 H_1

现在决定处理可以被看作通过比较由改变的图像内容生成的值 s 的可能性与由可靠内容生成它的可能性。由λ值确定决定边界。不同的λ值导致不同的错误警告和检测概率从而允许绘制 ROC 曲线。选择能够给出特殊错误警告概率的λ值由此在 ROC 曲线上选择了操作点。该方法已知为 Neyman-Pearson 决定判断标准，其能针对选择的错误警告概率将检测概率最大化。

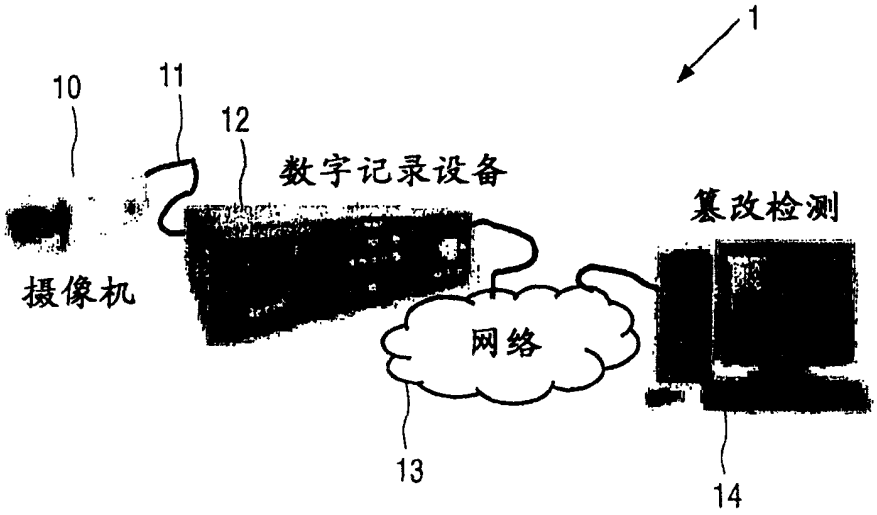

1
10
11
12
数字记录设备
篡改检测
摄像机
网络
13
14

图 1

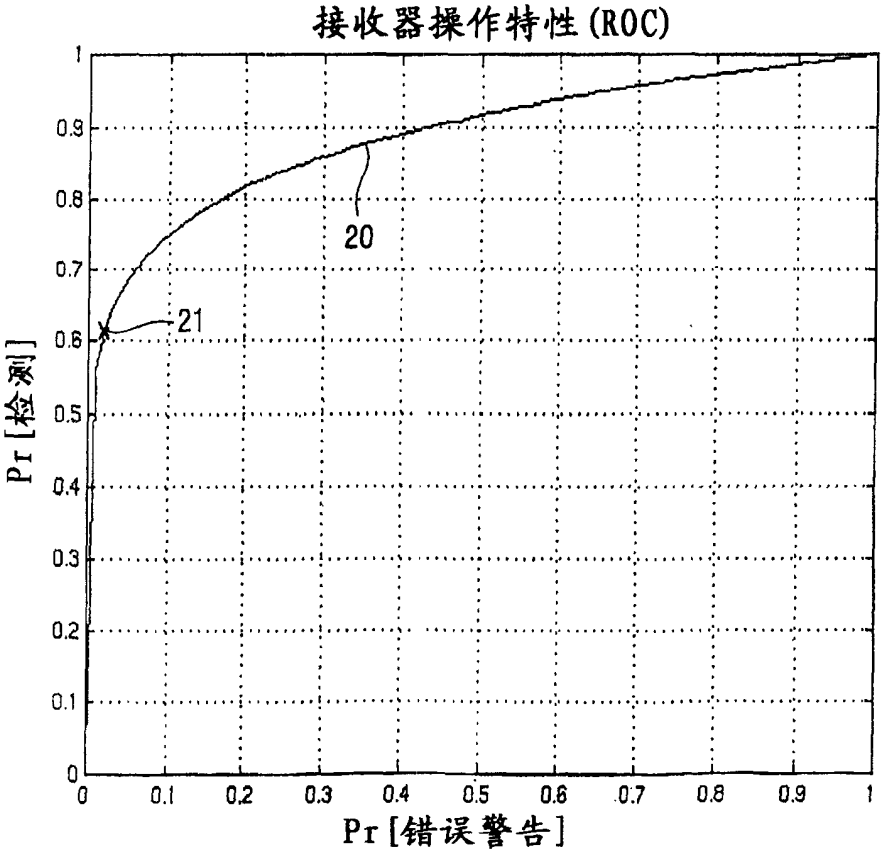

接收器操作特性(ROC)
Pr[检测]
Pr[错误警告]
20
21
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

图 2

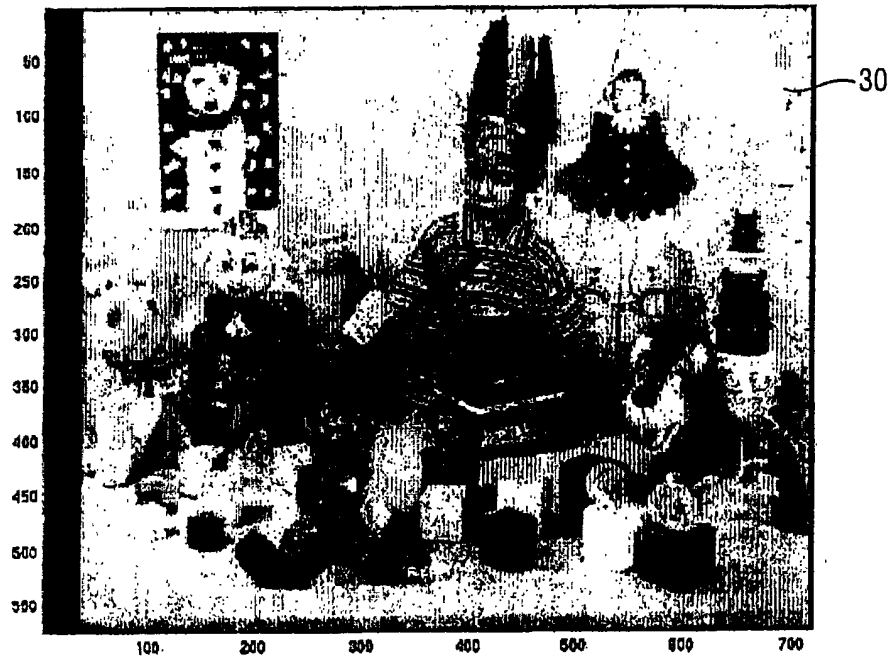

30
50
100
150
200
250
300
350
400
450
500
550
100
200
300
400
500
600
700

图 3

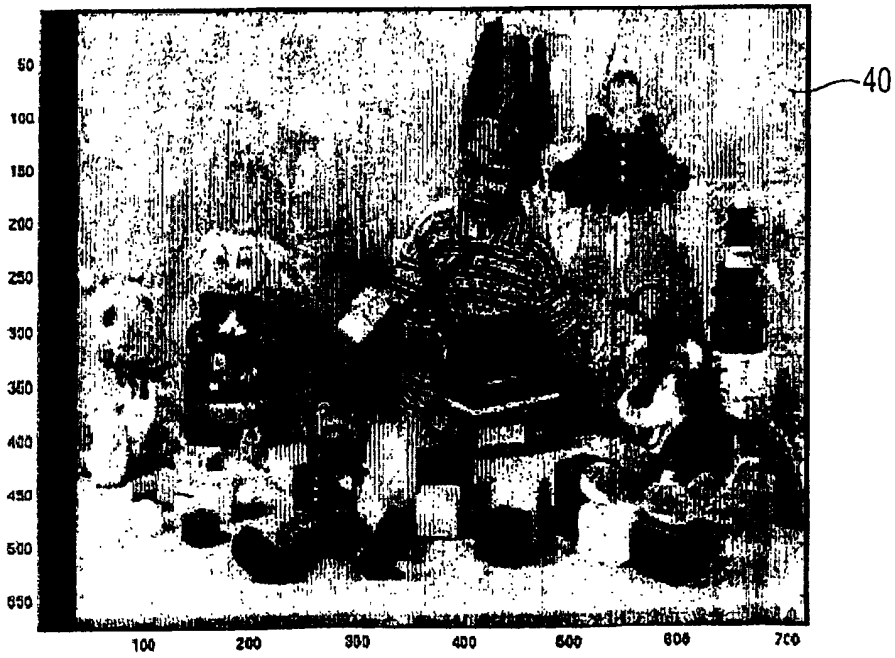

40
50
100
150
200
250
300
350
400
450
500
550
100
200
300
400
500
600
700

图 4

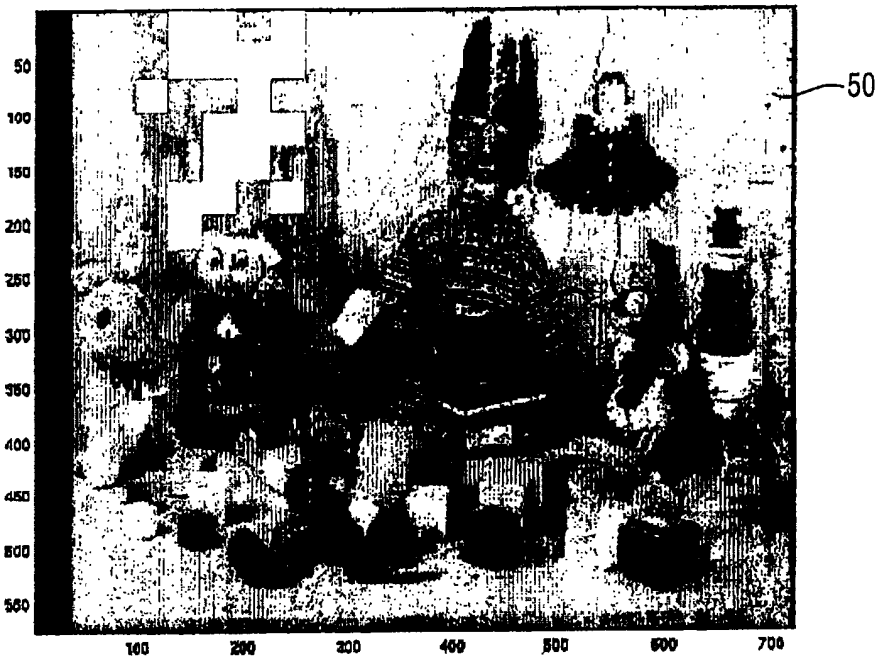

50
100
150
200
250
300
350
400
450
500
550
100
200
300
400
500
600
700
50

图 5


60
50
100
150
200
250
300
350
400
450
500
550
100
200
300
400
500
600
700

图 6

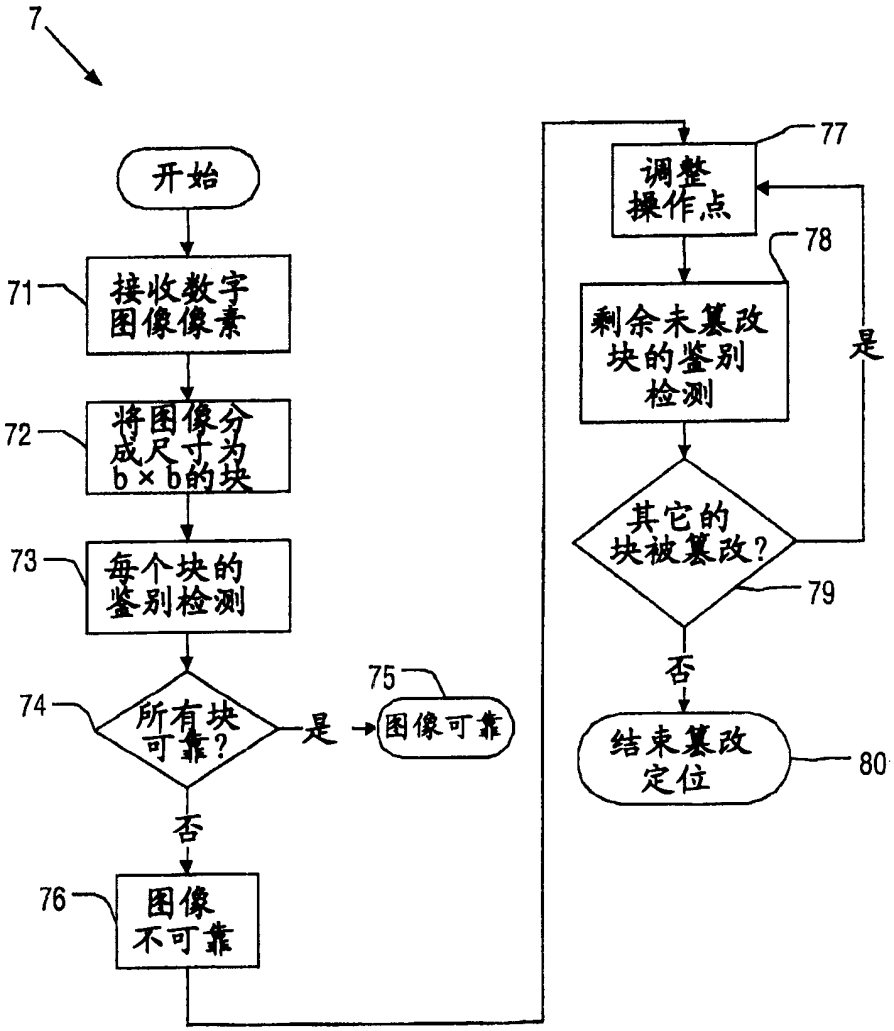

7
开始
71
接收数字图像像素
72
将图像分成尺寸为b×b的块
73
每个块的鉴别检测
74
所有块可靠?
是
75
图像可靠
否
76
图像不可靠
77
调整操作点
78
剩余未篡改块的鉴别检测
其它的块被篡改?
是
79
否
结束篡改定位
80

图 7

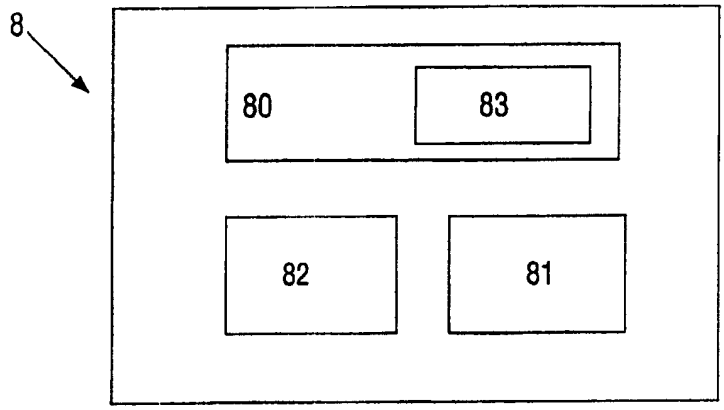


图 8

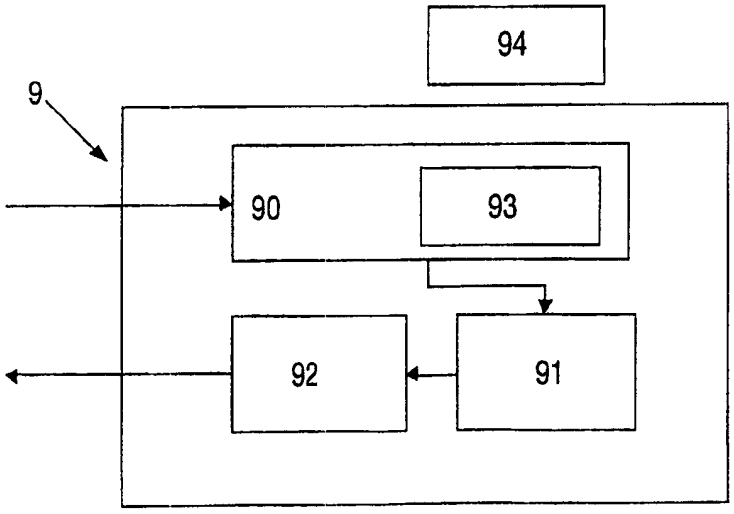


图 9

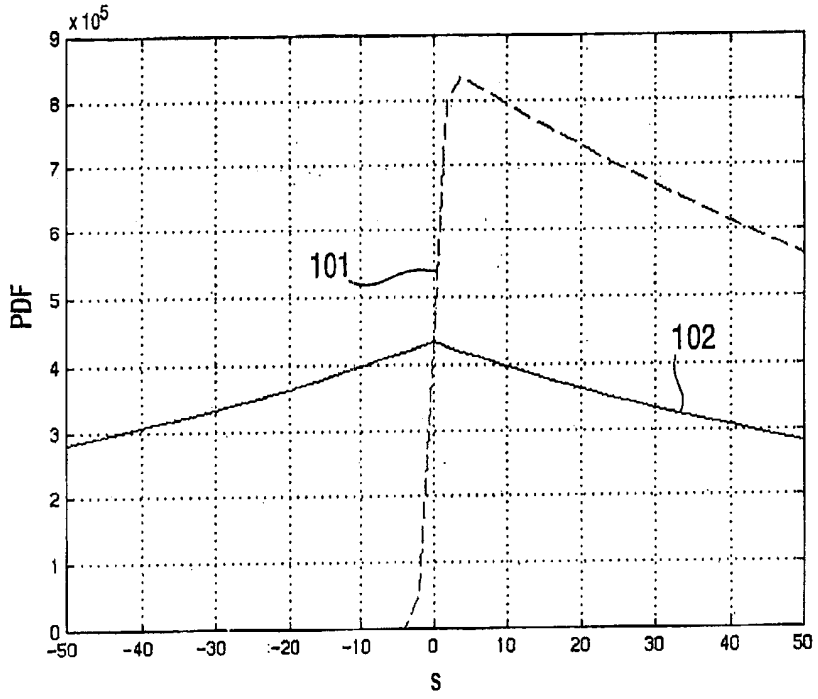


图 10

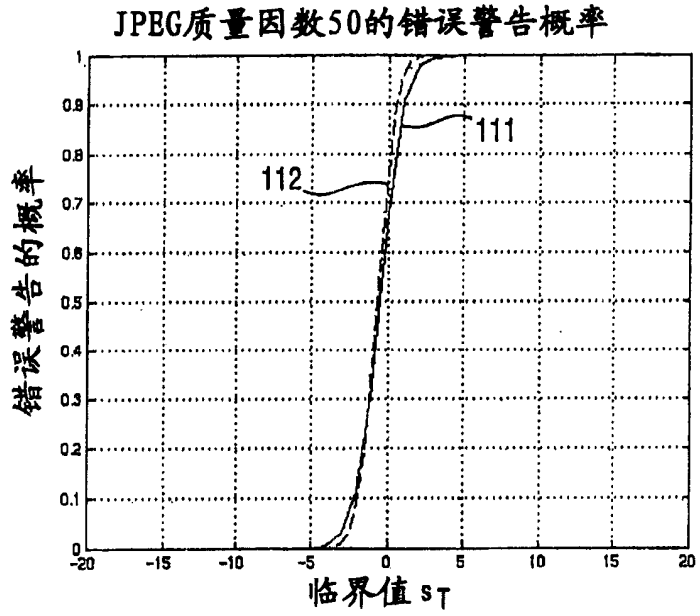

JPEG质量因数50的错误警告概率
错误警告的概率
临界值 s_T
111
112

图 11

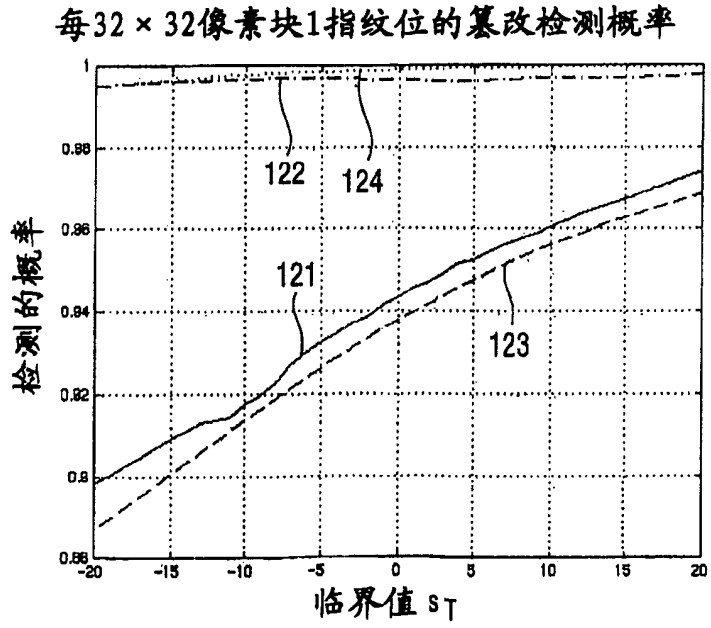

每32 × 32像素块1指纹位的篡改检测概率
检测的概率
临界值 s_T
122
124
121
123

图 12