

[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
E06B 1/02 (2006.01)



[12] 发明专利申请公布说明书

[21] 申请号 200680052783.1

[43] 公开日 2009年2月18日

[11] 公开号 CN 101371001A

[22] 申请日 2006.12.14
[21] 申请号 200680052783.1
[30] 优先权
[32] 2005.12.14 [33] US [31] 11/302,495
[86] 国际申请 PCT/US2006/047665 2006.12.14
[87] 国际公布 WO2007/070622 英 2007.6.21
[85] 进入国家阶段日期 2008.8.13
[71] 申请人 谷歌公司
地址 美国加利福尼亚州
[72] 发明人 迪帕克·吉恩达尔
阿努拉格·阿加瓦尔

[74] 专利代理机构 中原信达知识产权代理有限责
任公司
代理人 钟 强 郑 立

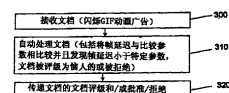
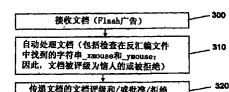
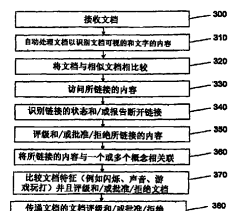
权利要求书4页 说明书29页 附图10页

[54] 发明名称

检测并拒绝恼人文档

[57] 摘要

一种用于为了批准或拒绝和/或评级而评价文档的系统和方法。所述方法包括：将文档与一个或多个标准相比较；确定所述文档是否包含与被确定是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素。



1. 一种批准文档的计算机实现的方法，包括：
将所述文档与一个或多个标准相比较；以及
确定所述文档是否包含与被确定为是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素。
2. 如权利要求 1 所述的方法，进一步包括：基于所述确定的动作做出批准或不批准的决定。
3. 如权利要求 1 所述的方法，其中，基于人类输入将所述可视元素、音频元素或文本元素确定为是令人不愉快的。
4. 如权利要求 1 所述的方法，其中，自动将所述可视元素、音频元素或文本元素确定为是令人不愉快的。
5. 如权利要求 1 所述的方法，其中，所述比较的动作包括确定所述文档是否包含可以生成与被确定为是令人不愉快的一个或多个行为本质上一样的行为的计算机代码。
6. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包含比被确定为是令人愉快的延迟参数的量值更低的延迟参数。
7. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包含无需由所述用户发起而导致一个或多个视频文档的下载的计算机代码。
8. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括无需由所述用户发起而下载一个或多个音频文档的计

计算机代码。

9. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括无需由所述用户发起而下载一个或多个视频或音频文档的计算机代码。

10. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括无需由所述用户发起而下载一个或多个视频以及一个或多个音频文档的计算机代码。

11. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括从除所述文档自身之外的源获得数据的计算机代码。

12. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括打开网络连接的计算机代码。

13. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括以大于先前指定的用于比较的次数循环的计算机代码。

14. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括含有模拟的随机数生成的计算机代码。

15. 如权利要求 5 所述的方法，其中，所述确定的动作包括确定所述文档是否包括调用与被确定为令人不愉快的一个或多个类型的硬件本质上一样的硬件的计算机代码。

16. 如权利要求 15 所述的方法，其中，所述一个或多个类型的硬件包括计算机鼠标。

17. 如权利要求 15 所述的方法，其中，所述一个或多个类型的硬件包括游戏控制器。

18. 如权利要求 1 所述的方法，其中，所述比较的动作包括将所述文档的图像与被确定为令人愉快的质量值相比较。

19. 如权利要求 18 所述的方法，其中，所述一个或多个质量参数包括所述文档的可视内容的边框边界的清晰度。

20. 如权利要求 18 所述的方法，其中，所述一个或多个质量参数包括所述文档的可视内容的边框边界占据所述文档的可视内容的程度。

21. 如权利要求 1 所述的方法，其中，所述比较的动作包括将所述文档的图像的帧的变化对比率与被确定为令人愉快的变化对比率相比较。

22. 如权利要求 1 所述的方法，其中，所述比较的动作包括将图像的帧的色彩变化率与被确定为令人愉快的色彩变化范围相比较。

23. 如权利要求 1 所述的方法，其中，所述比较的动作包括将文档的图像的转换的比率与被确定为令人愉快的转换的比率相比较。

24. 一种用于批准第一文档的系统，包括：

比较模块，将正被评价的所述第一文档的一个或多个特征与至少一个参数相比较；以及

确定模块，确定所述文档是否包含与被确定为是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素。

25. 一种对文档评级的计算机实现的方法，包括：

将所述文档与一个或多个标准相比较；

确定所述文档是否包含与被确定为是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素；以及
基于所述比较结果对所述文档评级。

26. 如权利要求 25 所述的方法，进一步包括：

显示所述文档的所述评级。

27. 一种对文档评级的计算机实现的方法，包括：

将所述文档与一个或多个标准相比较；

确定所述文档是否包含与被确定为是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素；以及
基于所述比较指定所述文档用于审阅。

28. 如权利要求 1 所述的方法，进一步包括：

显示所述文档的批准或拒绝状态。

检测并拒绝恼人文档

技术领域

本发明总体上涉及检测文档的所不希望的特征，诸如广告，并且拒绝此种文档用于分发。

背景技术

因特网为广告提供了有用的场所。然而，因特网广告可能包含通常被发现为对查看广告的人来说是恼人的（annoying）或令人不愉快的（displeasing）特征。例如，广告可能包含冒犯性的语言或诸如闪烁或频闪的恼人的行为或是较差的图像质量。实体希望能够检测恼人或令人不愉快的广告，其中所述实体对于是否允许广告在因特网上显示具有某些控制。

检测恼人或令人不愉快的广告的一种方法是在将广告向受众分发前人工地审阅广告。然而，存在与人工审阅相关联的问题。例如，人工审阅系统和方法的一个显著的缺点是在审阅广告中所涉及的时间和费用。

当前的系统和方法存在这些以及其它的缺点。

发明内容

因此，本发明的各种实施例可以指向一种用于为批准或拒绝和/或评级而评价文档的系统和方法。该方法包括：将文档与一个或多个标准相比较；确定所述文档是否包含与被确定为是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素。基于关于至少一个参数的所存储的信息和所述比较可以对文档作出批准确定。

在另一个示例实施例中，提供了一种用于批准或拒绝文档的方法。处理文档用于分发。将文档与至少一个参数相比较并且就文档是否包含与被确定是令人不愉快的可视元素、音频元素或文本元素中的一个或多个本质上一样的元素作出确定。基于比较的动作，文档被批准或被拒绝用于分发。

附图说明

图 1 描述了根据本发明的实施例的用于批准或拒绝电子文档的系统。

图 2 描述了根据本发明的实施例的用于批准或拒绝电子文档的系统的运行的网络环境。

图 3a 是图示了根据本发明的实施例的用于基于电子文档的特征而批准或拒绝电子文档的示例方法的流程图。

图 3b 是图示了根据本发明的实施例的用于基于广告的特征而批准或拒绝 Flash 广告的示例方法的流程图。

图 3c 是图示了根据本发明的实施例的用于基于电子文档的特征而批准或拒绝动画 GIF 文档的示例方法的流程图。

图 3d 是图示了根据本发明的实施例的用于基于广告的特征而批准或拒绝静态图像广告的示例方法的流程图。

图 4a 描述了根据本发明的实施例的示例文档。

图 4b 描述了根据本发明的实施例的具有抖动广告的示例文档。

图 4c 描述了根据本发明的实施例的具有布局问题的示例文档。

图 5 描述了根据本发明的实施例的提供评级（rating）信息的示例文档。

具体实施方式

在此描述的实施例解决了现有系统和方法的许多问题。因特网内容提供者面临的一个问题是评价大量文档（诸如通过其服务展示的广告）以确定对广大的不同用户来说每个文档是否是恼人的或令人不愉

快的。在此描述的实施例通过识别恼人或令人不愉快参数并且将文档和参数（例如冒犯性的语言或闪烁行为）相比较来处理文档以确定该文档是否是恼人的或令人不愉快的，从而克服这些和其它的问题。所述处理可以自动地发生，即通过机器实现的过程和/或不用人类输入或干涉而发生。

在此描述的实施例允许Flash和动画图像文档（例如广告）。这些类型的广告中的一些是恼人的。本发明的实施例提供上载诸如广告的文档并且将该文档与特定参数相比较。可以通过文档处理器（例如自动地通过图像处理器）将文档与参数相比较。处理器可以处理图像、声音文件和其它的数据来识别广告中的文本、图像（以及口语单词和其它的数据）以及行为。例如，可以使用光学字符识别（OCR）技术在图像中识别文本。通过将文档与特定参数相比较，可以在文档中识别特征并且将特征与文档相关联，并且可以因此对文档进行评级且基于这些特征以及比较参数的评级的状况而批准或拒绝文档。

尽管术语“广告”可被用作说明性的示例，但是应当理解，可以将相同的系统和方法应用于其它形式的文档或电子文档。如在此所使用，术语“文档”和“电子文档”可以涵盖一个或多个广告、内容页面（例如网页）、搜索结果、电子邮件、应用程序、IM消息、音频内容或文件、视频内容或文件、其它文件、可以驻留于计算机系统的一个或多个（例如网络）上的其它数据或应用程序、或其它可定义的概念或内容。

概述及系统体系结构

图 1 描述了根据本发明的实施例用于基于电子文档的特征而批准或拒绝电子文档的系统 100。该系统可以包括：服务器 2、一个或多个提供者 8、一个或多个文档源 12、一个或多个终端用户 10、以及有效地连接到服务器 2 的一个或多个数据库 50。

系统 100 可以使服务器 2 能够处理电子文档的特征评级。文档源 12、提供者 8 以及终端用户 10 可以经由包括因特网通信的电子通信与一个或多个服务器 2 通信。文档源 12、提供者 8 以及终端用户 10 可以包括或具有到用于提供与电子文档相关联的功能的一个或多个服务器 2 的通路（access）。

可以在服务器 2、提供者 8、文档源 12 以及终端用户 10 之间及之中进行通信的信息可以包括以下的一个或多个：文档信息、文档性能信息、文档特征评级信息、敏感性评级信息、适合性标准信息、信任分值信息、提供者信息、链接信息以及其它的信息。文档信息可以包括以下的一个或多个：文档自身、文档中使用的任何语言、长度信息、关于文档中的文件类型的信息（例如 html、doc、zip 等等）、文档类型（广告、教育性文档）、概要信息、音频内容（例如歌词）、可视内容（例如脸部图片）、色情内容、其它冒犯性内容（例如使用潜在冒犯性的词语）、程序设计代码、图像质量、与文档相关联的行为、与年代（age）相关的内容、文档所有者和/或文档创作者的身份、关于文档的预期受众的信息（诸如地理区域、年龄范围、性别、种族、国籍、信仰、其它的人口统计信息）、以及与文档有关或与服务器 2、提供者 8 或文档源 12 有关的任何其它信息。

文档源 12 可以将文档提供给服务器 2，或服务器 2 可以从文档源 12 “拉出（pull）”或检索文档。例如，文档源 12 可以将广告提供给服务器 2，使得服务器 2 然后将广告提供给一个或多个内容提供者 8，并且提供者 8 可以将广告提供给一个或多个终端用户 10（或服务器 2 可以将广告直接提供给终端用户 10）。文档源 12 可以包括任何内容创作者或内容提供者 8，诸如广告收录（listing）提供者或服务器 2。

提供者 8 可以将文档提供给一个或多个终端用户 10a 至 10n。提供者 8 可以包括内容提供者、搜索引擎或在诸如因特网的电子网络上使信息、服务和/或产品可用的其它实体。提供者 8 可以包括以下的一个

或多个，例如：广告收录提供者、电子文档提供者、网站主机、服务器 2、将电子文档提供给用户或其它实体的任何其它实体、或内容的任何其它提供者。

提供者 8 可以是与运行的服务器 2 相关联的实体的合作伙伴。终端用户 10 可以是一个或多个人、计算机、计算机网络或网络上的其它实体。终端用户 10 可以向提供者 8 和/或服务器 2 请求内容并且从其接收内容。基于各种应用可以包括额外的参与者。

服务器 2 可以包括任何服务器 2、集线器、中央处理器、提供者、搜索引擎或网络中的其它实体。与服务器 2 耦接的数据库 50 可以包括一个或多个数据库 50-66。此外，数据库 50-66 可以包括单个数据库 50 的部分。应当理解，数据库 50-66 可以是或不是物理上不同的。服务器 2 及其模块 20-42 可以存储并且访问存储在数据库 50-66 中的信息。

特征数据库 52 可以存储文档和/或与文档有关的数据，诸如部分、图像、行为、文本、以及文档可用的硬件、模式、规则以及程序设计。

可以从文档源 12 和/或提供者 8 接收文档。服务器 2 也可以生成文档。文档可以是已被评级的和/或未被评级的。例如，文档可以具有或不具有与其相关联的特定类型的评级信息。例如，文档可以具有来自模块 28、30，而不是来自终端用户 10 或提供者 12 的评级信息。

特征数据库 52 可以识别某类型的文档。例如，可以根据以下特性中的一个或多个来识别文档：主题、特征评级、合计特征评级、敏感分值、特征类型、语言、地理来源地（例如来源地的国家或城市）、目标受众的地理区域、文档源 12、内容的所有者、内容的创作者、目标人群、行为（诸如图像闪烁）、图像移动、文档可用的硬件（诸如鼠标、游戏控制器、照相机或麦克风）、用户交互是否由文档提供（其可以指示游戏）、文档的程序设计是否涉及随机数生成、或其它标准。

例如，可以根据其冒犯性/适当性特征、关联的关键词、关联的站点（例如从该文档明确地或隐含地链接到的站点，诸如通过嵌入式超文本链接）、关联的站点的状态（例如文档中的链接是否被断开（**broken**）和/或是否指向无效的 URL）、肉体内容（例如人类图像的裸体状态）、色情或其它淫秽内容、成人内容、毒品或酒精相关内容、儿童内容来识别文档。文档也可以包含恼人的行为，诸如闪烁、频闪、重复移动、无限循环的动画、流视频和/或音频的使用、打开网络连接、涉及游戏玩打、或提供者可能希望用来作为批准或拒绝文档的标准的其它行为。也可以由较差的图像质量或其它标准来识别文档。

图像数据数据库 54 可以存储图像数据。可以从文档源 12、图像阅读器模块 28 以及图像文件数据模块 30 接收图像数据。例如，图像阅读器模块 28 可以读取图像数据并且将其存储在图像数据数据库 54 中。

图像数据数据库 54 也可以存储由光学字符识别（“OCR”）（例如 OCR 处理器和/或软件）以及其它图像处理器使用来处理并且识别文本以及图像的广泛多种的图像和数据。例如，图像数据数据库 54 可以存储用于定义并且描述各种图像和图像类型的程序以及文件。程序也可以识别文档中的模式，其可以被用来将文档与其它文档相比较（例如通过将一个文档中的模式与另一个文档中的模式相比较）。图像数据数据库 54 可以存储通用（以及特定的）图像以用于比较。例如，图像数据数据库 54 可以存储苹果的通用图像。比较模块 32 可以处理水果的图像并且将其与所存储的苹果的图像相比较来确定两幅图像是否充分类似并且因此确定是否可以将该图像分类为苹果的图像。

特征评级数据库 58 可以存储特征评级（即“评级信息”）以及由图像数据模块 28 生成的其它信息，以及可能与对文档进行评级有关的任何其它信息。例如，特征评级数据库 58 可以存储这样的信息，所述信息与文档是否涉及闪烁或频闪图像、重复移动、无限循环动画、流

视频和/或音频的使用、打开网络连接、涉及游戏玩打、或其它恼人或令人不愉快的行为有关。每个所存储的特征评级可以与特定的文档（例如电子文档）相关联。

每个特征评级可以包括一个或多个主题评级。

每个特征评级也可以包括一个或多个行为评级。例如，特征评级可以包括对闪烁、频闪、到其它文档的链接（例如到 URL 的链接）、程序代码循环、与流音频或视频的连接、或与诸如鼠标或游戏控制器的硬件装置的连接的评级。

每个特征评级也可以包括一个或多个图像质量评级。例如，特征评级可以包括对与图像尺寸、边缘、清晰度、阴影或透视性有关的质量的评级。

可以将或不将这些独立的评级组合成单个分值（例如用于文档的单个合计评级）。文档的特征评级和/或合计评级可以随着接收到文档的新的评级信息而改变。

特征评级数据库 58 也可以存储文档的一般的批准状态。例如，批准状态可以是“已批准”、“未批准”、“暂停以待进一步审阅”或“已拒绝”。文档可以具有关于多种标准的批准状态。

链接数据库 60 可以存储所链接的文档以及链接本身。例如，服务器可以从文档源 12a 接收图像广告。图像广告可以具有嵌入式超文本链接（例如 <http://www.website.com/about/index>）以及图像中显示的不同链接（例如 www.website.com）。链接数据库 60 可以存储两个链接；链接数据库 60 也可以存储所链接的文档，例如位于所链接的地址的网站的图像或 html 文件。链接数据库 60 也可以存储其它链接相关信息。

其它的数据库 66 可以存储与文档、链接、所链接的文档以及其它信息有关的其它信息。

服务器 2 可以包括一个或多个模块来处理文档和内容、文档评级和其它实体评级、（例如文档源的）信任分值、以及其它数据。服务器 2 的模块可以对包括外部数据、数据库以及其它输入的各种数据源进行存储、访问并且另外与其交互。

文档评级模块 26 也可以处理与由文档执行的行为有关的评级信息，所述行为诸如闪烁、频闪、到其它文档的链接（例如到 URL 的链接）、程序代码循环、流音频或视频文档的下载、或到诸如鼠标或游戏控制器的硬件装置的连接。文档评级模块 26 通过检查用于生成这样的行为的对象、命令、以及其它代码的文档程序代码来完成这样的处理。

例如，涉及鼠标运动跟踪的文档程序代码将涉及用户交互并且很可能涉及游戏玩打，并且因而可能被提供者 8 认为是恼人或令人不愉快的广告特征。类似地，涉及随机数生成的程序代码很可能将涉及游戏玩打并且可以被分配适当的行为评级。无需由用户发起而导致音频或视频行为的程序代码可能被提供者 8 认为是恼人或令人不愉快的广告特征，并且可能被分配适当的行为评级。

文档评级模块 26 也可以检查诸如闪烁、闪动或抖动的行为。Flash 多媒体创作程序（“Flash”）可能已被用来创建文档并且其是基于矢量图形的，并且因此把帧表示成对象的集合而不是位图。SWF 文件由文件头（header）及其后面的一序列的标签组成。存在两种基本类型的标签：定义型标签和控制型标签。定义型标签定义对象，该对象然后由控制型标签操作来形成动画。Flash 播放器处理标签直到其遇到 ShowFrame 标签。在这一点上，在屏幕上示出位于显示列表上的所有对象。播放器然后继续处理新的控制型标签直到其遇到下一个

ShowFrame 标签。动画被以这种方式呈现在屏幕上。Flash 也支持被称为 ActionScript 的脚本语言，ActionScript 支持表达式求值、分支、循环、Javascript 形式的对象模型以及继承。ActionScript 也具有对流音频和视频以及一些网络连接能力的内置支持。

SWF 文件中的一些标签具有 ActionScript 块，其在一些事件后被执行或当显示元素时被执行。代码块通常以常数池（constant pool）开始，该常数池包含 ActionScript 代码中使用的一组字符串。这些字符串可以是字符串常量、变量名或类名。指令从栈中获取一些自变量（argument）并且将结果压入回栈上。例如，ActionScript 中的 Date 对象创建例如：

```
myDate = new Date();
```

在编译后，其将作为以下一组指令出现：

```
action: Constantpool String: "myDate", String: "Date" action: Push  
Lookup:0 Lookup: 1 action:NewObject action: SetVariable
```

进栈操作将两个字符串以特定次序压入栈上。NewObject 命令将一个自变量从栈中弹出并且使用其来例示该类型的对象。作为结果的对象然后被压入回栈上。SetVariable 将两个值从栈中弹出并且将其用作赋值语句的 RHS 和 LHS，即其将栈上的第一值赋值给栈上的第二值（变量）。

为了跟踪广告的点进率（click-through-rate），系统对用于在显示广告时点击的目标 URL 进行评价。对于 Flash，其可以通过设置通常被称为“clickTAG”的变量以传递目标 URL 来完成。广告可以确保在点击后，其即从 clickTAG 变量读取 URL 并且将用户导向到该 URL。系统可以抉择仅接受支持 clickTAG 的 Flash 广告。

GIF 可能已被用来创建就恼人检测和/或评级和/或批准或拒绝而进行评估的文档。GIF 格式通过遮罩（matte）通道来支持透明度。像素的遮罩值指定其是否是透明的。GIF 动画中的每一帧具有页面偏移

量，其指定了该帧相对于动画而被显示的位置。透明度选项连同页面偏移量以及部署（disposal）方法一起经常被用来创建优化的动画。部署方法指定了在显示下一帧之前在前一帧上执行的行为。其支持四个选项：

1. 未指定：替代旧帧
2. 不部署：保持旧帧
3. 恢复到背景：将帧恢复到背景。这里的背景是对 GIF 所指定的特定背景层。
4. 恢复到前一帧：将帧恢复到旧帧之前的状态

GIF 动画中的每一帧包括延迟参数，其指定了以 1/100 秒为单位的对该帧的显示持续时间。NETSCAPE 扩展也允许指定对动画应该进行的迭代次数。

闪烁和抖动广告通常具有较小的延迟参数，因而文档评级模块 26 可以检查已经由提供者 8 量化而被认作是较小的延迟参数。闪烁或闪动广告通常包括快速改变对比度和色彩方案的帧。因此，文档评级模块 26 可以检查不同帧中的相同内容以及不同帧之间色彩改变的显著（如由提供者 8 通过指定的比较率所量化的）强度。抖动广告通常是这样的图像，其每一帧包含相同的对象但是该对象处于稍微不同的位置。因此，文档评级模块 26 可以做逐像素匹配以确定所述帧是否只是单个图像的转换（translation）。为检测抖动广告，文档评级模块 26 也可以通过将包含在图像中的对象的边界对齐或通过计算运动矢量并且检查矢量组是否指向相同的方向来匹配图像。文档评级模块 26 也可以对图像的所有帧计算快速傅立叶变换并且获取系数的绝对差值总和。该差值越接近零，图像越可能是抖动的。如果该差值计算为在由提供者 8 指定的量化参数内，则广告可以被认作是抖动的并且被分配以适当的评级。

文档评级模块 26 也可以通过将文档的质量特征与先前所建立的

质量参数相比较来检查图像质量和布局问题。对于布局问题，文档评级模块 26 可以从图像中提取边界并且检查该边界是否具有锐利的边界或是否占据了图像的大部分，具有锐利的边界或占据了图像的大部分将被认为是所不希望并且导致适当的一个评级或多个评级被分配给该文档。对于图像质量，可以基于由提供者 8 选择的图像质量度量（IQM）来评价文档并且文档将被分配适当的一个评级或多个评级。

图像阅读器模块 28 可以包括配备有一个或多个光学传感器的一个或多个计算机或数据处理器。图像文件数据模块 30 可以处理图像数据，诸如从图像阅读器模块 28 所接收的数据。图像阅读器模块 28 和图像文件数据模块 30 可以被用在对来自文档的图像的诸如图像、文本以及链接的光学数据的处理中，所述处理包括评级。在于 2004 年 5 月 10 日提交的题目为“Method and System for Approving Documents Based on Image Similarity”的美国专利申请 No.10/841,833（代理人卷号为 64557.000027）中进一步详细解释了该处理。该申请通过整体参考而被合并于此。

基于图像阅读器模块 28 处理的光学和/或计算机文件数据，图像文件数据模块 30 可以识别文档中的一个或多个图像。换句话说，模块 30 可以辨别文本、字符和/或图像。基于该信息，图像阅读器模块 28 可以按照这些和其它标准来对文档（例如图像）进行评级。

使用模块 26、28、30 来确定评级信息的优点在于评级信息可以被自动确定（即不用人类干涉）。由于服务器 2 可以接收并且分发数千、数百万和/或数十亿的不同文档，人工审阅每个文档的事务和管理成本可能是不太可能的和/或昂贵的。模块 26、28、30 可以提供对图像或其它文档的充分的初步审阅以确定文档是否被批准（例如可接受的和适当的）用于全面分发或用于分发到一个或多个特定的提供者 8 或终端用户 10。

应当理解，单个文档可以具有一个或多个电子实例以及一个或多个物理实例。例如，单个文档可以具有下列形式：`.pdf` 文件、`.gif` 文件、`.tiff` 文件、其它格式的文件、打印纸张、照片、其它物理对象、或在屏幕或显示器上的可视显示。因而，可以以电子格式（例如`.pdf` 格式）或物理实施例（例如在纸张上所打印的文档）将单个文档从服务器 2 传递到评级实体（诸如图像阅读器模块 28 以及图像文件数据模块 30）。

应当进一步理解，图像阅读器模块 28 和图像文件数据模块 30 也可以被配置为基于声音、动画、视频、弹出能力以及其它视听信息处理并且识别概念。因此，模块 28、30 可以进一步包括扬声器、麦克风以及音频/视频处理器。例如，模块 28、30 可以包括语音识别技术来处理声音文件以确定声音文件中的所说的或所唱的词语。模块 28、30 可以进一步被配置为辨别声音的类型，如大海的声音、爵士乐的声音以及管弦乐的声音。基于音频（以及其它视听）信息，另外的概念可以与文档相关联。

比较模块 32 可以将图像（或其部分）或文件内容与其它图像（或其部分）或文件相比较，所述其他图像（或其部分）或文件例如存储在特征数据库 52 和图像数据数据库 54 中的图像或部分。特别地，比较模块 32 可以将来自一个文档源 12a 的图像与已经被存储在内容数据库 52 或图像数据数据库 54 中来自同一文档源 12a 的一个或多个文档相比较。例如，比较模块 32 可以确定文档与另一个文档是否是一样的（或本质上一样的）。文档比较模块 32 也可以确定在两个或多个文档之间的相似度（例如一个文档与另一个文档 80%相似）。

比较模块 32 也可以识别第一文档具有与第二文档、文件或图像相似的冒犯性的、不适当的内容或恼人的元素。例如，如果模块 32 识别到图像广告包含与已被确定为是色情的网页上的 Janet Jackson 的不适当的图像本质上一样的图像，模块 32 可以标记该图像广告以用于审阅或指示该图像广告可能具有不适当的内容。例如，模块 32 可以将否定

评级信息发送给文档评级模块 26，并且文档评级模块 26 可以标记该广告或改变该广告的评级。

链接模块 34 可以审查文档中的一个或多个链接（例如响应于识别链接而自动进行）。例如，如果图像广告文档包括一个或多个链接（例如如果图像广告在广告上显示了 URL 链接、嵌入了超文本链接或使观看者可以例如通过点击链接而选择相关文档），则模块 28、30 可以访问该链接以及所链接的文档。如果模块 28、30 不能访问该链接或所链接的文档（例如如果处理器的 web 浏览器未能加载所链接的文档或存在其他错误），则处理器可以报告“断开”链接（例如响应于识别“断开”链接而自动进行）。例如，模块 28、30 可以将该信息作为评级信息报告给文档评级模块 26。

链接信息可以包括链接质量评级（例如链接是否工作或是否具有任何问题）及链接的内容（例如内容评级）以及任何所链接的文档（例如所链接的网站）。可以以如在此所描述的评级文档的任何方式来获得信息。任何链接相关信息可以被存储在链接数据库 60 或特征数据库 52 中。

最后，链接相关信息可以被传递给文档评级模块 26，例如，以使得任何所链接的文档（或链接自身）的特征可以作为文档的评级的因素。例如，如果广告链接到与性、毒品或酒精有关的站点或如果其链接到闪烁、包含流音频或视频、包含无限循环动画、涉及游戏玩打等等的文档，则该广告可被评级为不适当的。

链接模块 34 可以周期性地审查文档中的链接，例如，甚至在文档已经被评级后。例如，其可以每个月或用户在链接上点击的每第十次检查该链接，以确保文档具有将用户导向到适当站点的有效（working）链接。

其它模块 42 可以完成与对电子文档评级有关的其它功能。在于 2003 年 12 月 23 日提交的题目为“Method And System For Providing Targeted Graphical Advertisements”美国专利申请 No.10/742,791、于 2004 年 3 月 30 日提交的题目为“System and Method for Rating Electronic Documents”的代理人卷号为 64557.000020 的美国专利申请中、以及在于 2004 年 5 月 10 日提交的题目为“System And Method For Rating Documents Comprising An Image”的代理人卷号为 64557.000024 的美国专利申请中描述了多个另外的服务器 2 以及系统 100 的功能。这些申请通过整体参考而被合并于此。

这些申请描述了使得能够处理文档的评级信息的其它模块 42。例如，文档可以被接收和/或被生成并且被传递给诸如终端用户、提供者以及其它实体的评价者。文档例如可以具有反馈链接或用于对文档进行评级的其他机制，以使得终端用户可以在接收文档后对文档进行评级。服务器 2 可以接收评级并且为每个所评级的文档确定内容评级（例如性内容、暴力内容以及其它适当性或合适性因素）。评级可以被合计并且被存储在特征评级数据库 58 中。合计评级可以指示在诸如成人内容以及暴力内容的各种内容范围中的评级。可以为各种提供者和终端用户确定敏感性信息，所述提供者和终端用户对其接收的内容的类型具有特定的偏好（例如对色情资料的禁止）。基于文档的一个评级或多个评级，文档评级模块 26 可以确定文档是否应该被批准或被拒绝。

说明性系统网络环境

图 2 描述了根据本发明的实施例用于评级、批准或拒绝电子文档的系统的运行的网络环境。在此种环境中，评价者和提供者 8 可以通过网络 14、15 连接到服务器 2（例如使用安全 https 连接）以将文档和评级信息提供给服务器 2，并且从服务器 2 接收文档和评级请求信息。服务器 2 可以将文档、评级以及其它信息存储在数据库 50 中。服务器 2 可以通过各种论坛或馈送（feed）分发文档，包括在印刷媒体中直接分发，在附属于服务器 2 的一个或多个网站上并通过提供者 8 提供文

档。应当注意到，提供者可以包括（例如取决于所希望的安全性通过网络 14 或 15 连接的）服务器 2 的联合合作伙伴（syndication partners）、（例如与内容数据库相关联的）内容系统以及由服务器 2 或提供者 8 操作的搜索引擎系统。

通过这些各种论坛，提供给提供者 8 的文档可以被包括在向终端用户 10 显示的页面（或其它文档）中（通常被称为闪现(impression)）。

服务器 2、提供者 8 以及文档源 12 中的每一个可以包括计算机化的系统，其包括以下系统中的一个或多个：web 服务器 2、数据库服务器 2、代理服务器 2、网络平衡机制和系统、以及使系统能够在因特网或其它网络类型系统上运行的各种软件组件。此外，尽管被描述为 http 网络，但是网络 14 和 15 可以包括其它网络，诸如专线、内联网或任何其它网络。在示例实施例中，在诸如广告提供者的文档源 12 和服务器 2 之间的连接（以及诸如在提供者 8 和服务器 2 之间的其它连接）可以包括安全网络连接，以确保数据不会遭受任何黑客或其它第三方的攻击或破坏。另外，尽管描述了两个文档提供者 12，但是应当理解，在网络中可以提供一个或多个文档提供者 12。类似地，尽管描述了一个数据库 50，但是应当理解，可以提供多个数据库并且可以经由任何类型的网络连接将这样的数据库连接到服务器 2，包括分布式数据库服务器 2 体系结构。

类似地，提供者 8a 可以包括经由任何类型的网络连接到服务器 2 的任何数量的这样的系统，包括 http 或 https 网络。内容提供者 8 可以包括诸如服务器 2 的系统，其提供用于启用通过因特网或其它网络协议的连接的功能。终端用户 10 可以包括任何用户（诸如连接到因特网的用户）并且可以包括计算机化的系统，该计算机化的系统启用通过各种类型的网络的该连接，所述连接包括通过因特网服务提供商、有线公司以及访问因特网上的数据的任何其它方法。提供者 8 可以包括将诸如广告的内容分发给终端用户 10 的任何系统。

说明性过程

图 3a 是图示了根据本发明的实施例用于评级或批准或拒绝电子文档的示例方法的流程图。

在框 300 中，可以接收文档。例如，诸如广告收录提供者的文档源 12a 可以将图像广告传递到服务器，以用于分发给提供者 8 和终端用户 10。以示例的方式，所述文档可以是示出女性模特的图像广告。

在框 310 中，文档可以被处理（例如自动地）以识别文档的可视的以及文字的主题（诸如闪烁、涉及游戏玩打以及具有较差的图像质量的其它特征，稍后将在框 370 中被识别）。例如，图像阅读器模块 28 和图像文件数据模块 30 可以处理如在此所述的文档来识别文档中的文本和/或图像。可以识别女性的图像以及指示对产品背书的文本和到网站的链接（例如 www.bathingsuits.com），其可以被显示在广告中和/或被嵌入广告的超文本中。

在框 320 中，可以将文档与相似文档进行比较。例如，基于框 310 中的图像处理，比较模块 32 可以将所处理的图像信息与其它文档的图像信息相比较并且识别相似文档。诸如与相似文档相关联的关键词的概念可以被识别。例如，通过该比较，被评价的文档可以与裸体、性、暴力、亵渎以及提供者 8 可能希望用作评级或拒绝文档的基础的其它主题相关联。

在框 330 中，例如可以通过链接模块 34 访问所链接的内容。该框 350 仅与包括至少一个链接的文档有关。例如，链接模块 34 可以经由浏览器访问链接在广告中的 www.bathingsuits.com 网站。

在框 340 中，可以识别和/或报告任何链接的状态。例如，链接模块 34 可以识别链接是起作用的并且所链接的网站已被正确加载入浏览

器中。如果链接没有正确地将浏览器导向到正确的网站，或者如果网站未能加载，则可以报告该链接已断开。

在框 350 中，可以对所链接的内容进行评级、拒绝和/或批准。如在框 310 中，可以例如通过模块 28、30 对链接的内容以及任何所链接的文档进行评级和/或批准。评级/批准信息可以与底层（underlying）文档相关联。

如果与文档相关联的链接已断开，则可以标记文档以用于审阅或者将其否决。类似地，如果确定链接与不适当的内容相关联（例如在所链接的网站存在不适当的内容），则可以标记文档以用于审阅。

在框 360 中，所链接的内容可以与一个或多个概念相关联。所链接的内容可以与诸如关键词的概念相关联。如上所述，底层文档的评级和关联性可以反映所链接的内容的评级和关联性。

在框 370 中，可以将文档与参数而不是可视的或文字的内容相比较（可视的以及文字的主题已在在先在的框中被评定），例如文档是否闪烁、是否涉及游戏玩打、图像质量是否较差，以进行评级和/或批准。例如，图像模块 26、28、30 可以识别在女性图像中存在大量肉体内容、文档涉及随机数生成或鼠标移动或指示文档涉及游戏玩打的其它行为、文档闪烁或抖动、或文档具有较差的图像质量，并且因此将文档评级为所不需要的。评级可以基于各种标准，诸如图像内容、文本、移动、声音、链接和所链接的文档以及其它标准。在一些实施例中，可以例如通过模块 26、28、30 自动确定评级/批准信息。可以使用各种技术将广告与特定参数相比较以用于提供者 8 评级和/或批准或拒绝。相关领域的技术人员将认识到，其它的语言和技术可以被用于参数的说明、文档与那些参数的比较、以及在此阐明的不与本发明的精神和范围分开的其它的权利要求元素。为了说明性的目的，在以下段落中描述了评价 Flash 广告、动画 GIF 以及用于评级的静态图像的示例方法。

Flash 广告：可以使用 swftools 库，因为其是鲁棒的（robust）和完全的解决方案。包含在 SWF 文件中的信息可以被用来执行以下任务：

1. 文本提取：在 Flash 动画中使用的大部分文本被存储在 DefineText 标签或其变体之下。在大多数情况下可以通过审查这些标签来提取该文本。该文本将与图像相关联并且可以对其检查关于图像中的文本的策略如语言、商标等等。

SWF 文件支持两种类型的文本-字形文本（Glyph text）和设备文本（Device text）。以确定字符的形状的二次 B 样条的控制点来描述每个字形文本字符。使用标准字体名称来指定用于设备文本的字体，但是也提供字形文本作为应变（fallback）选项。使用 DefineFont 和 DefineFont2 标签来定义字形字体，并且使用 DefineFont 和 DefineFontInfo 标签来定义设备字体。DefineFont 标签包含在动画中作为字形使用的字符列表。没有字符码与字形相关联。此外，在 DefineFont 中列出的字符仅是完整字母表的子集，因为其仅包括在动画中使用的字符并且没有任何附加信息，不太可能将每个形状都与其字符表示相映射，除非其对形状使用如 OCR 的技术。对于字形文本，该附加信息完全缺失，但是对于设备文本，DefineFontInfo 标签包含从字符形状到其字符表示（ASCII 或 Unicode）的映射。

将要被显示的真实的文本存储在 DefineText 或 DefineEditText 标签中。这些标签包含指向在对应于其使用的字体的 DefineFont 标签中的字符记录的指针。作为结果，对文本进行解码要求对 DefineFont 标签和 DefineText 标签两者进行解码。

Swftools 库允许遍历 SWF 文件中的标签。提供者 8 可以首先使用 DefineFont 和 DefineFontInfo 标签对标签进行迭代以将每个字体符号与相应的字符符号相关联。一旦发现字体，则使用相应的字体信息将 DefineText 和 DefineEditText 解码。

2.检查被禁止的行为：执行诸如流音频、流视频以及打开网络连接的某些行为的 Flash 广告可能被提供者 8 指定为是恼人的。为了检查这些行为，可以创建 ActionScript 对象并且可以就这些对象的存在而检查 SWF 文件。

Flash 的安全策略确保了，如果用户 10 正试图加载来自当前域之外的变量但是默许 Flash 电影流出数据或产生网页服务调用，则警告用户 10。即使 Flash 向用户 10 发出了关于电影访问远程数据的警告，但是由于电影打开了警告对话框所以对用户 10 来说其可能是恼人的。因此提供者 8 可能希望检查这样的方法。

对被禁止的方法的检查被分为两部分：（1）检查被禁止的对象以及（2）检查被禁止的方法。被禁止的对象是仅为了访问受限制的网络或一些其它资源而被使用的对象。示例包括如 Camera、Microphone、NetConnection、SharedObject 以及 XMLSocket 的类。为了检查被禁止的对象。提供者 8 可以寻找以下 ActionScript：

```
action: Push < class name > action: NewObject
```

提供者 8 也可以检查关于其它对象的某些网络相关调用。示例包括 Sound.loadSound()以及 Movie.loadMovie()。为了这些检查，提供者 8 将寻找以下 ActionScript 代码：

```
action: Push < method name > action: CallMethod
```

提供者 8 也可以使用作为另一种技术的控制流和数据流分析。

用于检测这样的行为的另一个可能性将是在执行正被审查的文档时监视 Flash 播放器做出的系统调用。提供者 8 可以检查端口打开请求、声音装置打开请求等等。

3.检查游戏：大部分游戏涉及某些类型的鼠标运动跟踪以及随机数生成。Flash 中的鼠标坐标可以通过电影剪辑的_xmouse 和_ymouse 属性的值来确定。可以检查试图做这件事的行为脚本代码。同样，可以检查其它类型的用户 10 交互。ActionScript 中的随机数通常通过“ActionRandomNumber”指令或通过使用内置 Math.random()库函数来生成，并且可以就随机数的存在而检查代码。

检查游戏：游戏典型地涉及鼠标跟踪以及随机数生成。在 Flash 中，可以通过使用电影剪辑（通常是_root 剪辑）的属性_xmouse 和_ymouse 来获得鼠标的位置。在 Flash 中，电影剪辑指的是动画序列，并且这样的动画序列也可以被嵌入到主体动画中。跟踪鼠标的通用方法然后是注册鼠标运动事件，并且在此事件后使用_xmouse 和_ymouse 属性并且采取适当的行为来获得鼠标的位置。电影剪辑、按钮以及 Mouse 对象（也可能是一些更多的对象）支持鼠标运动事件。也存在为事件注册监听程序的多种方法。

许多 Flash 游戏具有随同鼠标光标一起移动的对象（诸如交叉瞄准线（cross-hair））。对于这样的游戏，提供者 8 可以检查反汇编文件中的字符串_xmouse 和_ymouse。通常希望广告使用鼠标位置并且将“_xmouse”和“_ymouse”用作变量或用作字符串常量是罕有的。因此，该技术对于检测游戏来说应该有理由运行良好。可能存在其它的游戏，其中不需要鼠标坐标但是使用诸如鼠标悬停在某些区域上的其它鼠标行为。

通过“ActionRandomNumber”指令或通过 Math.random()调用来获得 ActionScript 中的随机数。对于 ActionRandomNumber 指令，提供者 8 可以只检查相应的 actionscript 标签的存在。对于 Math.random()函数，提供者 8 可以检查以下的一组指令：

```
action: Push "Math" action: GetVariable action: Push "random"
action: CallMethod
```

4.检查声音和视频：分别使用 DefineSound 和 DefineVideoStream 标签将声音和视频并入 SWF 文件中。提供者 8 可能希望允许在广告中使用声音但是不允许仅作为用户 10 行为的结果播放的声音、或提供者 8 可能仅检查 SWF 文件中声音的存在而不考虑其是否正由用户 10 事件发起。这可以使广告接收某一评级、被拒绝、或指示人类批准者（若有的话），对这样的广告进行更全面的检查。

5.检查并且固定 clickTAG：如果 clickTAG 是由提供者 8 使用的，则提供者 8 可能想要确保 Flash 广告具有对 clickTAG 的正确支持并且没有将用户 10 重定向到任何其它 URL。这可以通过检查 SWF 文件中是否存在“clickTAG”字符串以及在文件中不存在 URL 来完成。提供者 8 也能够仅用“clickTAG”替代“clickTag”来固定创意。

为了支持 clickTAG，Flash 广告应当具有按钮，该按钮具有被配置来在通过 clickTAG 提供的自变量上调用 GetURL 的点击行为。一个简单的检查将测试字符串“clickTAG”的存在以及代码中不存在任何 URL。另一个技术是检查“clickTAG”是否正被用作 GetURL2 指令的自变量。提供者 8 可以检查以下一组指令的任一个：

```
//1.getUrl(_root.clickTAG,<target>) 2.getUrl("clickTAG",<target>)  
Push <movie clip> Push "clickTAG" GetVariable GetVariable Push  
"clickTAG" Push <target> GetMember GetUrl2 0 Push <target> GetUrl2 0
```

GetUrl 指令的目标自变量指定了点击应当在其中被打开的目标窗口。另外，提供者 8 也可以检查文件中 URL 的存在。

6.检查和添加 clickTarget：提供者 8 也可能想要控制在其中显示登陆页面的窗口。这可以通过支持类似于 Flash 电影中的 clickTAG 的“clickTarget”自变量来实现。因为这是非标准的特征，提供者 8 可能希望确保所有广告主不用必须改变其广告目录来坚持（adhere to）该特

征，并且因此提供者 8 可能希望增加大部分现有 Flash 文件来支持该特征。

7.检查无限循环动画：该检查可以通过执行 SWF 文件以及注意对 flash 播放器在所允许的最大动画持续时间之后执行的显示装置的任何写入来实现。该技术可以被扩展来执行对网络调用、声音等等的检查。提供者 8 也可以模仿如鼠标点击的用户 10 行为并且检查对 clickTAG 的支持。

动画 GIF：ImageMagik 库可以被用于该目的。对于动画 GIF 最频繁遇到的两种问题是闪烁和抖动的图像广告。这两种类型的广告通常都具有较小的延迟参数并且出于安全可以将该方面用来标记广告的一些子集。

1.闪烁/闪动图像：在该类别下的广告包括具有快速改变对比度和色彩方案的帧。可以通过检查在不同帧中的相同内容以及注意在不同帧之间的色彩强度的改变来检测这样的广告。

闪烁图像在所有的帧中通常具有一组相同的对象并且只有对象的色彩（以及亮度）在帧间改变。基于帧之间的延迟指定比较参数消除了最恼人的图像。例如，在所有帧之间小于 0.5 秒的延迟的比较参数给出图像显现为是恼人图像的较高概率。

提供者 8 也可以识别动画中的相似帧，所述相似帧彼此之间相隔一定距离。该识别可以基于下述观察：具有某类闪烁效果的动画具有相似帧，该相似帧被其中在图片中出现新对象或在图片中存在的对象的色彩改变的帧而穿插。这两者基本上都可以被认为是帧之间的色彩中的急剧改变。

提供者 7 可以首先利用相似帧之间的帧来识别相似帧，然后将穿

插帧与相似帧中的一个相比较。可以计算每像素的对比度变化，作为该变化发生的频率。CIE1976 色差公式可以被用于测量两个图像之间的感知色差。

可以通过考虑诸如色彩的空间分布、色彩沿边缘的改变、掩模（mask）以及人类视觉系统（HVS）的其它特性的因素来改进算法。

2.抖动图像：抖动广告是其每一帧包含相同的对象但是该对象位于稍微不同的位置的动画，以提供抖动对象的闪现。在许多这些广告中，可以显现，帧实际上使用相同的图像于不变的背景，但是该相同的图像在不同的页面偏移量处。可以通过进行逐像素的匹配来检测这些广告。如果所述帧只是单个图像的转换，该帧的快速傅立叶变换(FFT)系数是相同的。另一个技术是通过对齐包含在图像中的对象的边界来匹配图像。

对于由改变相同图像的页面偏移量而创建的抖动动画可以通过比较在动画中使用的实际帧来容易地检测。其它抖动广告包括作为布置在（0，0）页面偏移量处的不同图像的帧。在某些这些广告中，抖动对象被放置于透明的背景上。在此种情况下，可以裁剪每一帧来移除这些透明像素，然后比较所裁剪的图像。

基于快速傅立叶变换(FFT)的技术对抖动动画也执行良好。可以计算所有帧的 FFT 系数。然后可以计算系数的绝对差值总和（SAD）。对于抖动图像，由于帧主要地涉及转换，该差值更接近于零。为了补偿（offset）图像中亮度级别的影响，提供者可以将阈值应用于与帧中的一个的系数的绝对值总和有关的 SAD。通过使用梅林傅立叶变换（Mellin Fourier Transform）可以使该技术更鲁棒。

通常为了检测频闪以及移动对象，提供者 8 可以计算运动矢量并且检查是否存在一组运动矢量指向相同的方向。提供者 8 可以使用基

于块的运动矢量算法。为了在块中充分良好地捕获特征，可以使用尺寸为 12 或更大的较大块。为了避免块的错误对齐的问题，可以使用重叠块。另一个技术是使用用于跟踪 OpenCV 库中的光学流的方法。另一个技术是通过被应用于随后帧中的对象的变换来跟踪 Flash 中的对象移动。如果帧之间的对象被以显著的量变换或缩放，则将指示存在频闪。

静态图像：提供者 8 可能希望检查静态图像中的劣质图像和布局问题。

1.布局问题：提供者 8 可以从图像中提取边界并且检查边界是否具有锐利的边界或是否占据了图像的大部分。可以通过沿着图像的每个边缘扫描并且找到平行于不匹配其的边缘的第一条线来首先识别图像的边界框的边缘。这些线组成图像的边界框。然后对边界框的边缘是否与原始图像中的边缘相对应的检查可以通过沿边缘计算梯度来完成。如果图像具有匹配边界框的边缘或如果边界框与原始图像相比较差别是十分小的，则图像可以被评级为具有布局问题。

2.图像质量：可以使用基于无参考的质量评定。在此种类型的评定中，不存在用来比较给出图像的质量的“理想”图像。在此种示例方法下，提供者 8 可以确定一些图像质量度量标准（IQM），诸如以这些 IQM 来评价给出的图像。同样，OCR 技术可以被用来从图像中提取文本。除文本提取外，其也可以被用来评定图像质量。对于被假定包含文本的广告，如果 ORC 不能从图像中译解任何有意义的文本，则文档可以作为具有较低质量的图像被评级和/或被拒绝。

提供者 8 也可以比较指定此种参数来检测意欲欺骗用户 10 的图像。一些广告可能包含无功能的并且仅被展示为图像的文本框、下拉以及按钮。提供者 8 可以使用边缘/角落检测技术来检测图像中这些对象的存在。

如果文档基于其特征（例如如果广告包含诸如闪烁或抖动的所不希望的行为）而被标记或不被批准，则文档可以被传递给一个或多个评价者以用于人类审阅。在多个人类审阅员批准文档后，文档可以稍后被批准。

在框 380 中，文档、评级以及批准或拒绝状态被传递给例如提供者 8。

本领域的普通技术人员应当理解，所述动作可由硬件、软件或其组合来执行，需要或无需人类干涉，如可以在与诸如提供者、评价者、数据库以及终端用户 10 耦接的诸如服务器 2 系统的一个或多个计算系统中具体化。此外，应当理解，不是所有的框都必须被完成。例如，在一个实施例中，方法可以在框 310 开始并且在框 370 结束。同样，不必以图 3 中示出的顺序来执行每个框的行为。可以考虑任何执行顺序。

图 3b 是图示了根据本发明的实施例用于评级或批准或拒绝 Flash 广告的示例方法的流程图。

在框 300 中，可以接收 Flash 广告文档。以示例的方式，在该示例中的 Flash 广告是游戏，其中用户 10 必须使用鼠标来玩打该游戏。

在框 310 中，文档可以被处理（例如自动地）。在反汇编文件中，字符串 `_xmouse` 和 `_ymouse` 被检查。找到该字符串并且将文档评级为恼人广告或将其拒绝。

在框 320 中，文档、评级和/或拒绝状态被传递给例如提供者 8。

图 3c 是图示了根据本发明的实施例用于批准或拒绝闪烁动画 GIF

文档的示例方法的流程图。

在框 300 中，可以接收动画 GIF 文档。以示例的方式，动画 GIF 文档是闪烁动画。

在框 310 中，文档可以被处理（例如自动地）。检查帧之间的延迟并且发现在所有帧之间其是 0.2 秒。然后将该所确定的延迟与 0.5 秒的延迟比较参数相比较。由于发现所确定的延迟小于比较参数，所以文档被认为是闪烁广告并且被评级为闪烁广告或被拒绝。

在框 320 中，文档、评级和/或拒绝状态被传递给例如提供者 8。

图 3d 是图示了根据本发明的实施例用于评级或批准或拒绝静态图像文档的示例方法的流程图。

在框 300 中，可以接收静态图像。以示例的方式，静态图像是可以具有围绕边框的可乐罐的图像。

在框 310 中，文档可以被处理（例如自动地）。从图像中提取边框并且确定边框占据图像的百分比。发现边框占据了图像的 50%。将 50% 值与 30% 的比较参数相比较。由于所确定的值（50%）大于比较参数（30%），所以文档被认为具有布局问题并且文档被评级为恼人广告或被拒绝。

在框 320 中，文档、评级和/或拒绝状态被传递给例如提供者 8。

说明性用户界面和结果

图 4a 描述了根据本发明的实施例的示例图像广告文档 1。图像广告 1 可以包括横幅广告、可以在网页上显示的另一个广告、或可以经由电子装置显示的另一个图形广告。图像广告 1 示出了特定图像 1A、

广告文本 1B 以及一个或多个链接 1C。特定图像 1A 包括电影放映机和 DVD、VHS，以及数字电影容器的图像。文本 1B 包括广告消息，例如产品或服务的描述、对潜在消费者的建议和 / 或其它的广告文本。链接 1C 可以包括指向诸如广告主的网页 URL（或其部分）的其他文档的链接。例如，链接 1C 可以包括嵌入式超文本链接，并且该嵌入式链接可以与显示在文档 1 的图像中的链接 1C 相关联。在一些实施例中，在（例如网页浏览器中）查看文档 1 的同时选择（例如在其上点击）所显示的 URL 或其它链接可以将用于查看文档的查看者的机制（例如网页浏览器）导向到与该链接相关联的内容（例如广告主的网页）。

文档可以具有各种类型的内容。例如，文档可以具有文字、图像、声音和其它信息，以及可以动态产生文字、图像、声音以及其它信息的功能或程序。每个文档可以具有不同量的各种类型的内容，诸如性内容、暴力内容、毒品或酒精相关内容、财经内容、成人相关内容、儿童相关内容以及其它内容。

文档 1 可以明确地显示链接 1C。同样，链接 1C 可以被嵌入到文档中（例如在文档的程序设计中）或其部分中，以使得链接 1C 是不可见的。在此，选择（例如在其上点击）文档 1、图像 1A、文本 1B 或文档的其他部分可以将用户 10 的文档查看机制导向到所链接的文档（或多个文档）。文档 1 自身、图像 1A 以及文本 1C 也可以包括一个或多个链接 1C。例如，登有多个产品的广告可以包括关于每个产品的链接，其中选择（例如在其上点击）与特定产品有关的图像 1A、图标 1A 或文本 1B 可以将网页浏览器（或其它文档查看机制）导向到与特定产品相关联的商家的站点的页面（或导向到另一个文档）。文档 1 可以迅速改变色彩、涉及不由用户 10 发起的流音频或视频、或可能被提供者 8 认作是恼人的或所不希望的其它行为。

文档可以具有各种类型的内容。例如，文档可以具有文字、图像、声音和其它信息，以及可以动态产生文字、图像、声音以及其它信息

的功能或程序。每个文档可以具有不同量的各种类型的内容，诸如性内容、暴力内容、毒品或酒精相关内容、财经内容、成人相关内容、儿童相关内容以及其它内容。

图 4b 描述了根据本发明的实施例的示例视频广告文档 500（其是抖动广告）的几个连续帧 10、20、30、40、50、60、70、80、90、110、120。帧 10、20、30、40、50、60、70、80、90、110、120 描述了可乐罐 11、21、31、41、51、61、71、81、91、111、121，其在帧 10、20、30、40、50、60、70、80、90、110、120 内的不同位置处变换，使得当连续显示所述帧时，可乐罐 11、21、31、41、51、61、71、81、91、111、121 在视频广告文档 500 中出现抖动。帧 10 描述了竖直站立的可乐罐 11，帧 20 描述了稍微倾向右侧站立的可乐罐 21，帧 30 描述了再次竖直站立的可乐罐 31，帧 40 描述了稍微倾向左侧站立的可乐罐 41，并且在帧 50 再次开始该次序。本发明将检测这个广告为恼人的，并且在一个实施例中标记文档来通知提供者 8 该广告需要被审查。在另一个实施例中本发明将自动拒绝该广告。

图 4c 描述了根据本发明的实施例的示例图像广告文档 1（其是静态广告）。图像广告 1 示出了可乐罐 2 的图像以及图像四周的边框 6。图像广告 1 具有不良的布局；具体而言，广告的尺寸导致较大的区域被广告 1 的边框 6 占据。本发明的实施例将检测这个布局问题并且通知提供者 8 广告该需要被审查。在另一个实施例中本发明将自动拒绝该广告。

图 5 描述了根据本发明的实施例的示例文档 1。图 5 描述了在监视器上向提供者 8 提供关于广告的评级以及批准/拒绝信息的文档。将正被评级以及被批准或被拒绝的文档广告 3 提供给提供者 8。所提供的评级包括内容评级 5、行为评级 7、链接评级 9、合计评级 11 以及批准/拒绝状态 13。

应当理解，在此所述的服务器、处理器以及模块可以自动地或经由自动化系统执行其功能（例如将文档与特定参数相比较和/或确定评级信息）。如在此所使用，术语“自动地”涉及由任何机器可执行的过程来执行的行为，例如不需要人类干涉或输入的过程。

本发明的实施例不局限于在此所述的特定实施例的范围中。例如，尽管已经参考图像广告而描述了在此公开的多个实施例，但是这里的原理同样适用于诸如网站的其它文档。事实上，根据前述描述和附图，除在此所述的内容外，本发明的实施例的各种修改对本领域的普通技术人员来说将是显而易见的。因而，这样的修改意在落入所附的权利要求书的范围内。此外，尽管本发明的实施例为了特定的目的在特定的环境中在特定实施方式的情况下已经在此被描述，但是本领域的普通技术人员应当认识到其有用性不限于此，另外本发明的实施例可以为任何数量的目的在任何数量的环境中被有利地实现。因此，应当考虑到如在此所述的本发明的实施例的完整灵魂和精神来解释在下面阐述的权利要求。

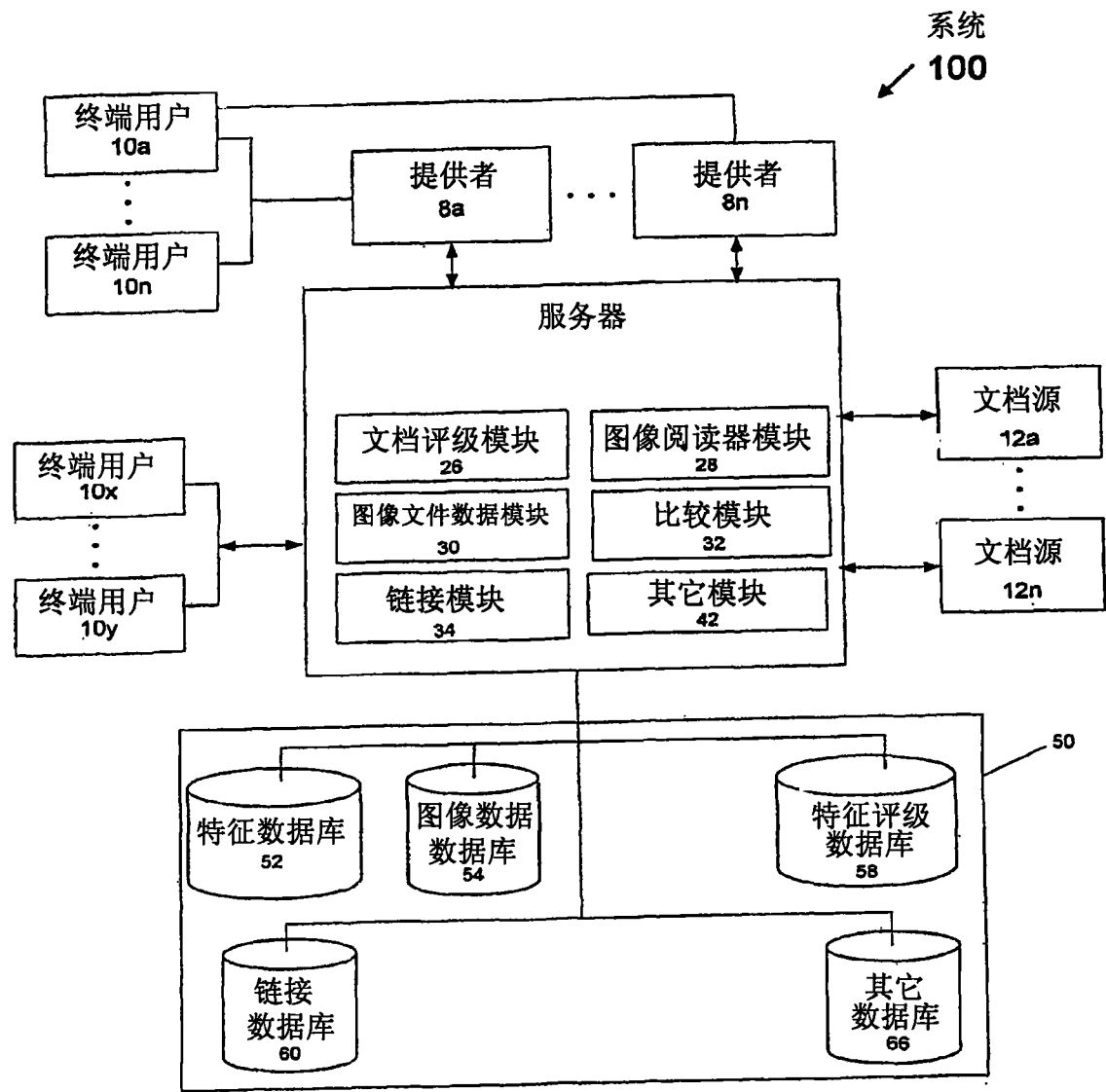


图1

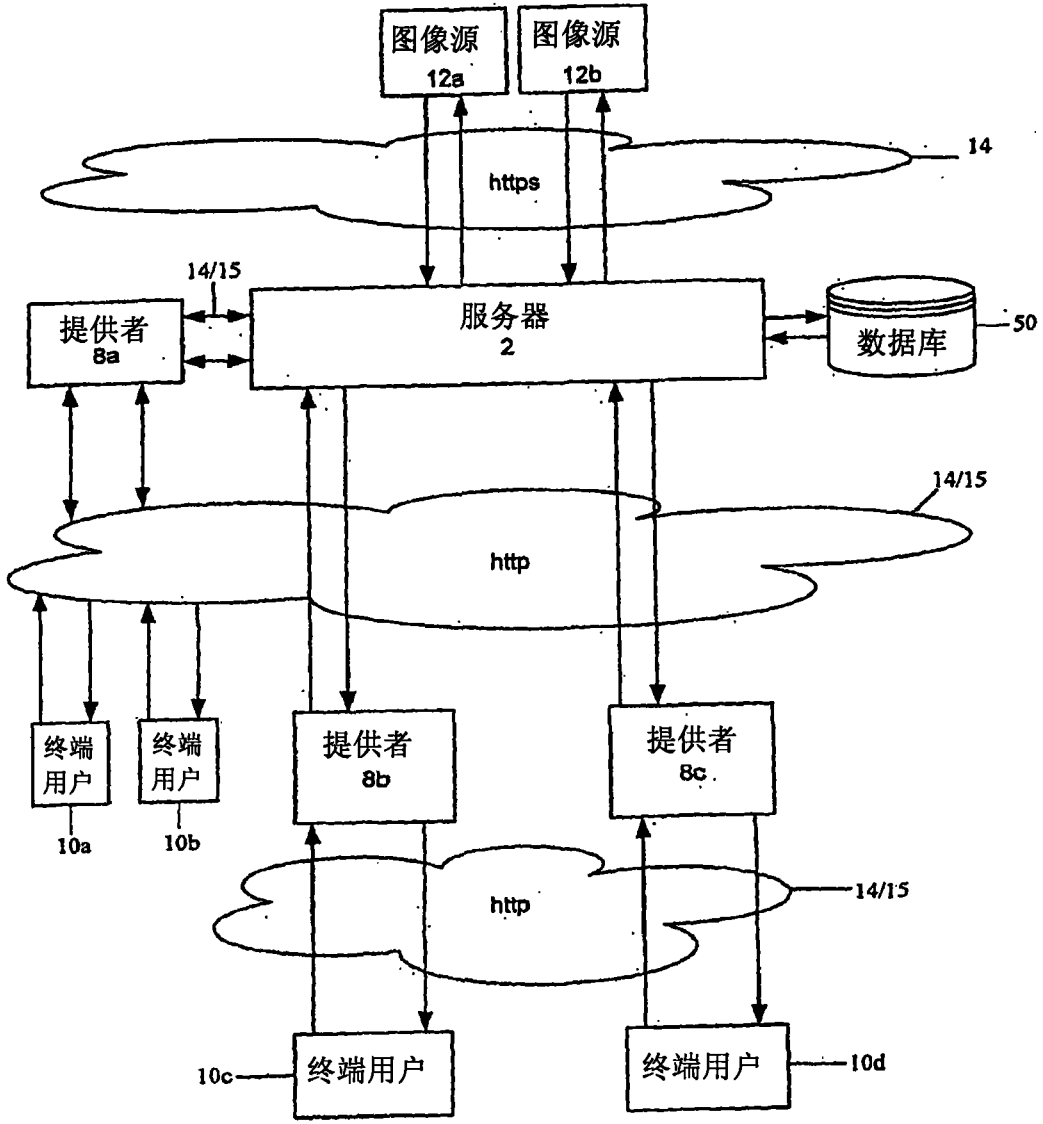


图2

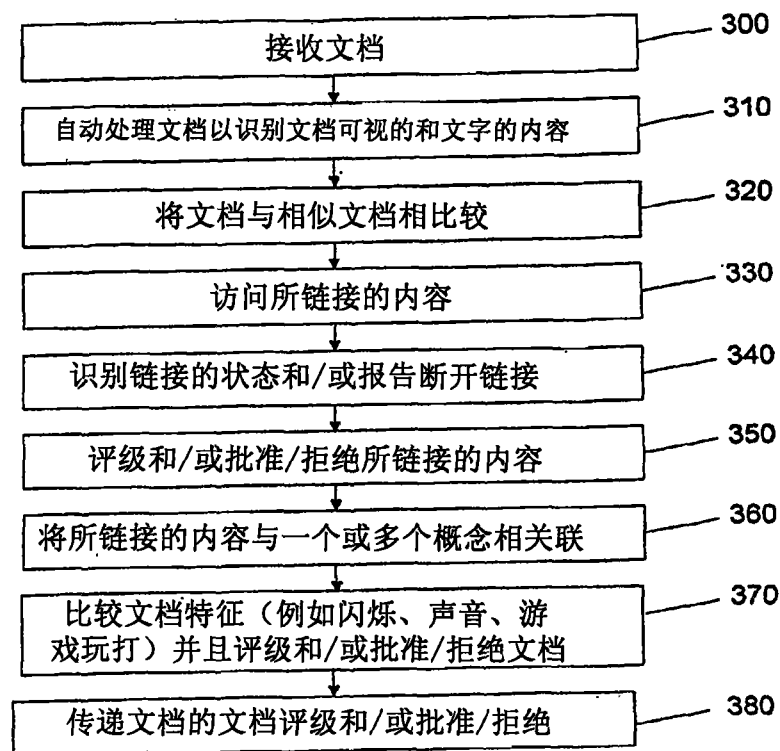


图3a

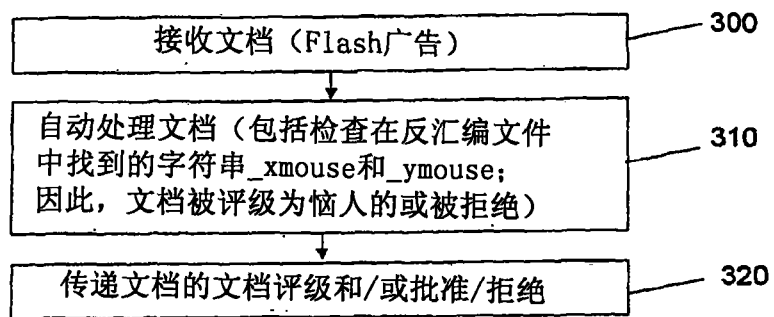


图3b

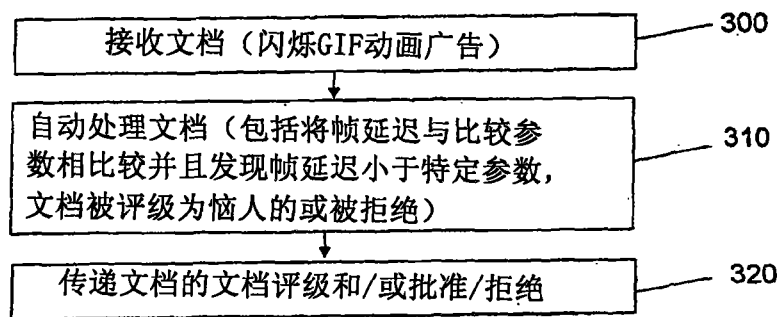


图3c

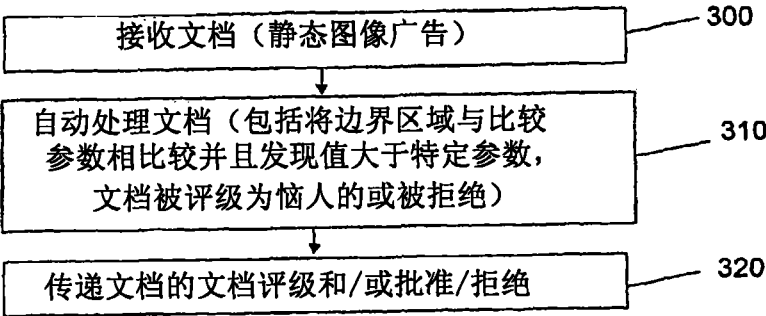


图3d

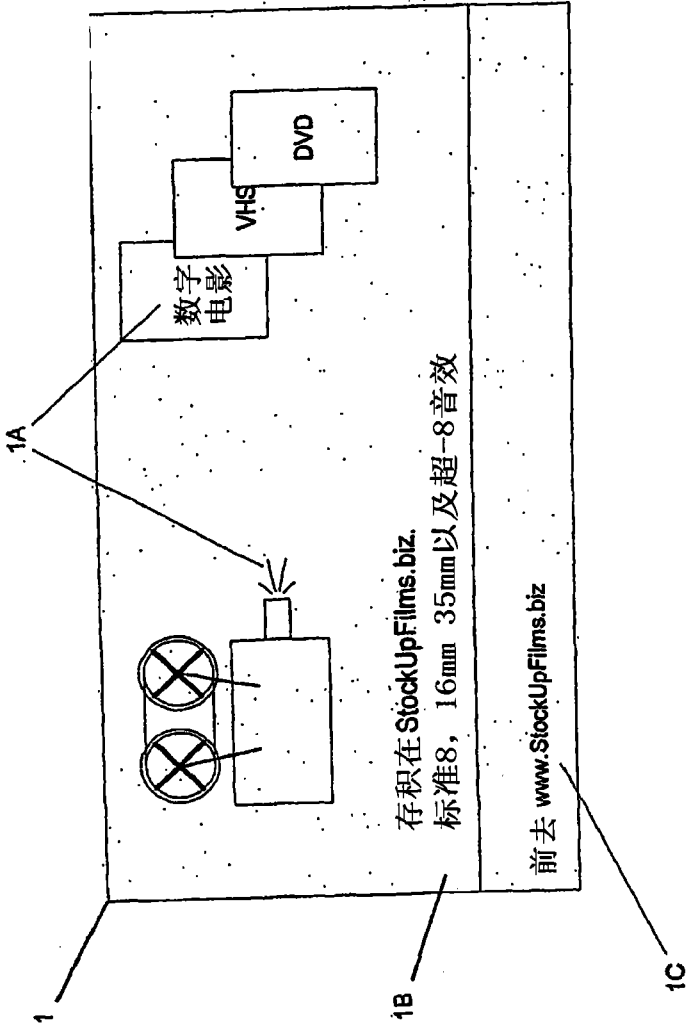


图4a

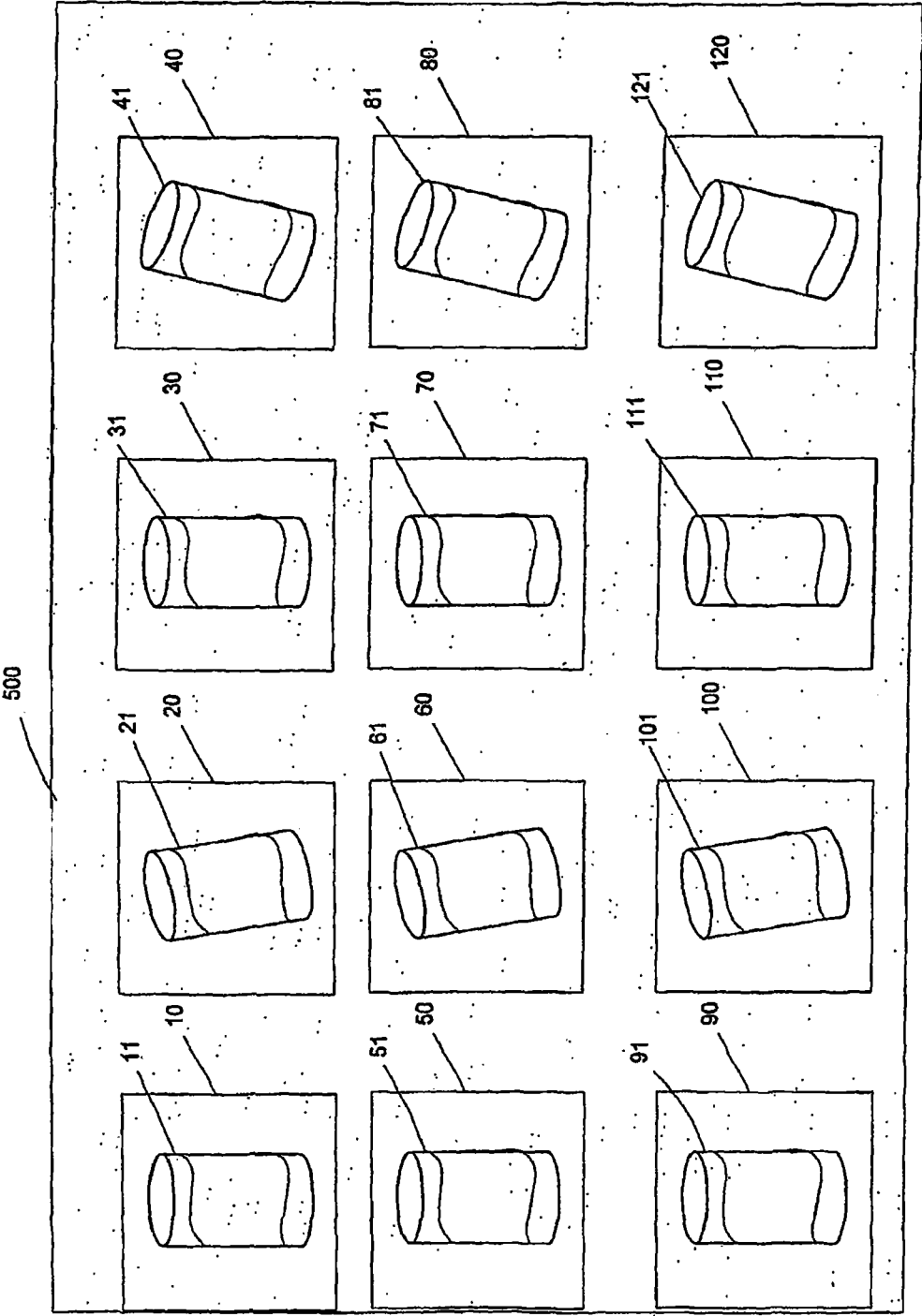


图4b

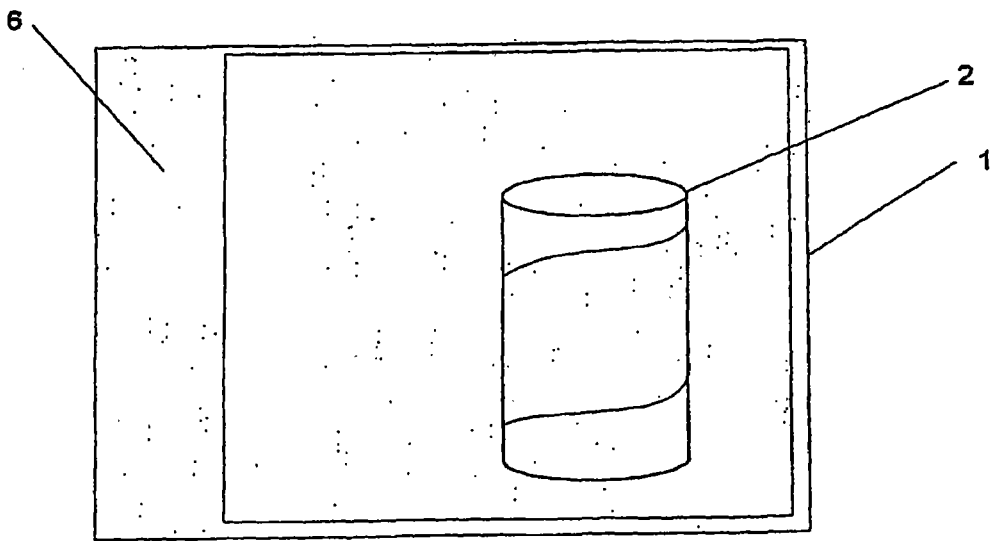


图4c

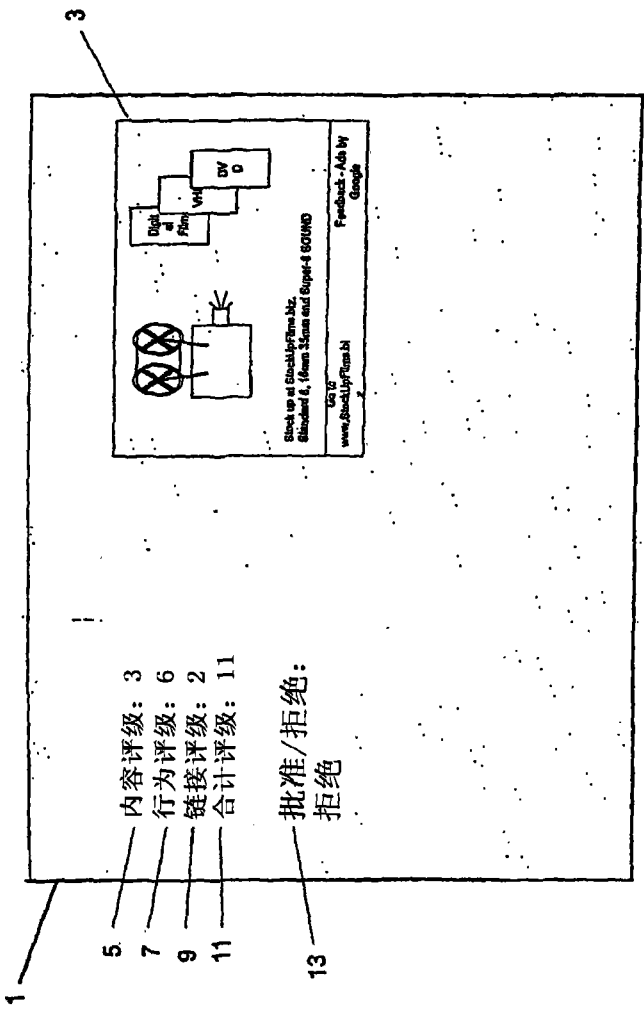


图5