

[19] 中华人民共和国国家知识产权局



[12] 发明专利申请公开说明书

[21] 申请号 200480006709.7

[51] Int. Cl.

G06F 17/30 (2006.01)

G06Q 10/00 (2006.01)

[43] 公开日 2006 年 4 月 12 日

[11] 公开号 CN 1759396A

[22] 申请日 2004.3.3

[21] 申请号 200480006709.7

[30] 优先权

[32] 2003.3.13 [33] EP [31] 03100639.8

[86] 国际申请 PCT/IB2004/050195 2004.3.3

[87] 国际公布 WO2004/081817 英 2004.9.23

[85] 进入国家阶段日期 2005.9.12

[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 H·A·A·M·德鲁伊特

J·A·海特斯马

A·J·L·M·马安多克斯

[74] 专利代理机构 中国专利代理(香港)有限公司

代理人 吴立明 王 勇

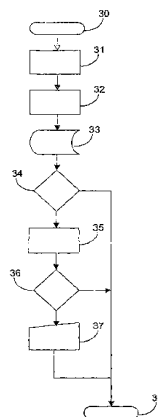
权利要求书 1 页 说明书 13 页 附图 4 页

[54] 发明名称

改进的数据检索方法和系统

[57] 摘要

一种获得与内容项关联的数据的方法，包含步骤：获得(32)内容项的标识符，利用该标识符执行(33)数据库查找以得到数据以及如果数据库查找没能获得该数据则将内容项提交(37)给输出端(210)以便由人类(200)来处理，其特征在于：自动地将内容项分类到若干类别之一中的步骤(35)，以及基于内容项的分类而有条件地执行至少一个其它步骤。在一个实施例中，只有当(36)数据库查找没能获得该数据并且将内容项分类到若干预定类别之一中时才将内容项提交给输出端。服务器(300)和计算机程序产品用于执行该方法。



1. 一种获得与内容项关联的数据的方法，包括步骤：获得(32)内容项的标识符，利用标识符来执行(33)数据库查找以获得数据，以及如果数据库查找没能获得该数据则将内容项提交(37)给输出端
5 (210)以便由人类(200)处理，其特征在于：自动地将该内容项分类到若干类别之一中的步骤(35)，以及基于内容项的分类而有条件地执行至少一个其它步骤。

2. 权利要求1的方法，包括：当且仅当(44)将内容项分类到若干预定类别之一中时，执行数据库查找。

10 3. 权利要求1的方法，包括：当且仅当(36)数据库查找没能获得数据并且将内容项分类到若干预定类别之一中时，将内容项提交(37)给输出端。

4. 权利要求1、2或3的方法，包括：基于将内容项分类到其中的类别而从多个输出端中选择输出端。

15 5. 权利要求1的方法，包括：通过计算内容项的指纹而获得该标识符。

6. 权利要求1的方法，包括：通过从嵌入内容项中的水印的有效负载中提取标识符而获得标识符。

20 7. 权利要求5或6的方法，包括：当且仅当将内容项分类到若干预定类别之一中时，获得该标识符。

8. 一种用于获得与内容项关联的数据的系统(300)，包括：用于获得内容项标识符的装置(302, 322)，用于利用标识符来执行数据库查找以获得数据的装置(303, 310)以及用于当数据库查找没能获得该数据时将内容项提交给输出端(210)以便由人类(200)处理的装置
25 (210)，其特征在于：至少一个装置(302; 322; 303, 310; 210)适用于根据用于将内容项分类到若干类别之一中的装置的输出端而操作。

9. 权利要求8的系统(300)，其中当用于执行数据库查找的装置(303, 310)没能获得数据并且用于分类的装置(309)将内容项分类到若干预定类别之一中时，用于提交内容项的装置(210)被配置成
30 有条件地提交该内容项。

10. 一种计算机程序产品，被设置成使处理器执行权利要求1的方法。

改进的数据检索方法和系统

本发明涉及一种获得与内容项关联的数据的方法，包含步骤：获得内容项的标识符，利用该标识符来执行数据库查找以获得数据以及如果数据库查找没能获得该数据则将内容项提交给输出端以便由人类来处理。

本发明进一步涉及一种用于获得与内容项关联的数据的系统，包含用于获得内容项的标识符的装置，用于利用该标识符来执行数据库查找以获得数据的装置，以及如果数据库查找没能获得该数据时用于将内容项提交给输出端以便由人类处理的装置。

本发明进一步涉及一种计算机程序产品。

由于越来越多的内容都在变得可利用，因此，自动广播监控器，即无线电台或电视台播放表的自动生成变得越来越重要。用于自动内容标识的已知技术常常是基于水印或指纹的。基于水印的系统从嵌入在内容项中的水印的有效负载中提取内容项的标识符。基于指纹的系统计算项的最相关感知特征并表示并利用其作为标识符。将若干内容项的标识符与诸如名称、艺术家、流派等等这样的它们的关联数据一起保存在数据库中。通过获得其标识符并且利用该标识符作为查询关键字或查询参数来在数据库中执行查找或查询而检索特定内容项的数据。然后该查找返回与标识符关联的数据。

当正在广播标识符在数据库中的歌曲、视频剪辑、影片或其他内容时，这种系统进行自动识别。然而，不管数据库多大，总是会广播那些其标识符不在数据库中的内容。例如，还不能识别标识符没有添加到数据库中的新发行的歌曲。而且由于系统的成本随着数据库的尺寸而线性增长，因此拥有很大型数据库也很没有成本效率。此外，在一个国家中由无线电台广播的98%的歌曲是属于仅来自于一个小的歌曲集合（典型地是20,000到30,000首）中。

假定它们想识别每个广播的内容项，当前的广播监控器提供者使人类收听或观看所有未识别的内容。由于这是一种手工操作，因此，提供者承受了大量的成本。

一种音频指纹的应用是其中消费者可利用他的移动电话来识别他不知晓名称的歌曲的服务。为了最佳的消费者满意度，必不可少的是消费者想识别的歌曲的指纹是处于数据库中的可能性。因此，将所有对该指纹服务的电话呼叫记录到音频文件中并且例如是一周一次地手动识别所有（或某个百分比）的这些文件。这样做是为了最佳化指纹数据库的内容，因此最大化了消费者想识别的歌曲的指纹存在于数据库中的可能性。用于视频的类似应用也是可能的。

美国专利5,862,223中公开了一种专家匹配方法和装置，其中将用户请求指定给人类专家以便由那些专家来回答。当收到请求时，由人类专家为类似请求搜索数据库来以避免重复工作。如果未发现类似请求，利用在该请求中发现利用的关键字或主题基于请求的分类而为适当的专家执行搜索。

美国专利申请2003/0037010公开了一种用于检测数字作品的未经授权传输的方法。由所关心的如MP3、AVI、ASF或OGG这样的文件类型来辨认并识别所关心的作品。查询数据库以确定所讨论的作品是否匹配数据库中的内容。通过数据库查找获得内容项的元数据。如果数据库查找失败，则手动地提供该元数据。利用作品标识符来搜索数据库。如果数据库搜索显示该作品被识别为版权保护，则采取适当的处理（例如块传输）。

本发明的一个目的是提供一种根据前文的方法，其降低了现有技术所需的手工劳动。

根据本发明通过这样一种方法来实现这个目的，该方法的特征在于：自动将内容项分类到若干类别之一中的步骤，以及基于内容项的分类而有条件地执行至少一个其它步骤。通过将自动分类技术与内容识别技术进行组合，显著地降低了必须手动处理数据库中的失败查找的成本。本发明建立在这样一个理解的基础上，即自动分类步骤允许理性判断数据库查找和/或向人类提交内容项是否将用于有用目的。

在以上提及的现有技术文献中，没有公开或暗示这种理解。在美国专利5,862,223中，对分类既不有条件地执行数据库搜索也不执行专家搜索。即不考虑将请求所分类的类别而执行这两个步骤。该类

别用于促进搜索，而不是为了决定没有必要搜索。在美国专利申请2003/0037010中，辨认内容为MP3或AVI的步骤仅用于选择适当的内容标识符模块，而不是决定数据库查找不必要。当辨认出内容为MP3或AVI时，查询元数据数据库或具有注册内容的数据库不是有条件的。

5 基于所有要处理的内容都是可识别的这个隐含假定而建立这些系统。因此，永远不会跳过数据库查找，并且如果数据库查找失败，总是提交该内容以便人工识别是有意义的。因而，即使所讨论的内容只不过是是不可能识别的随机噪声，这些系统将也总是执行数据库查找并提交该内容以便人工识别。

10 在本发明的第一方面，该方法包含当且仅当将内容项分类到若干预定类别之一中时执行数据库查找。避免了保证会失败的这种查找方式。例如，如果将内容项分类到“音乐”类别中，则在具有音乐的数据库中的查找可能会成功，但是不会发现分类为“噪声”的内容项，因此可以将其忽略。

15 在本发明的第二方面，该方法包含当且仅当数据库查找没能获得数据并且将内容项分类到若干预定类别之一时向输出端提交内容项。这降低了操作员必须收听或观看的内容数量。例如在广播音频监听的情况中，采用一种在“音乐”和“非音乐”间进行区分的简单的两类分类器。在这个情况中，仅必须手动识别那些未被指纹识别的并且分类为“音乐”的音频。由于大多数未识别的音频由语音组成，因此可实现手工劳动的显著降低。

20 还提供将传递给操作员的数据量最小化的优点。必须将服务器未辨认出的所有内容传送到操作员可以从中收听或观看它们的终端。这意味着潜在大量的音频或视频内容必须传送给操作员。如果操作员位于一个物理上很远的地点，那么所需带宽可能很昂贵。根据本发明，不需要传送那些分类为不可识别的内容，这降低了所需带宽。

25 在进一步的实施例，该方法包含根据内容项分类到其中的那个类别而从多个输出端中选择该输出端。可利用指定流派（流行的、古典的等）来标记未识别的内容的更成熟的分类器增加了自动将未识别的内容分给在相应流派中具有专门技术的适当人员的额外可能性。

在进一步的实施例中,该方法包含通过计算内容项的指纹而获得标识符。如音频或视频剪辑这样的内容项的指纹是所讨论的项最相关感知特征的表示。有时将这种指纹也称作是“(鲁棒)散列”。

5 在进一步的实施例中,该方法包含通过从嵌入在内容项中的水印的有效负载中提取标识符而获得该标识符。水印检测可能需要大量的处理,尤其是在视频水印检测的情况中。

根据本发明的第三方面,如果要通过指纹计算或水印有效负载提取而获得该标识符,那么当且仅当将内容项分类到若干预定类别之一中时执行这些操作是有益的。

10 在水印的情况中,通过分类并忽略被辨认为不含有水印的内容,降低了所需的处理量。例如是广告或随机噪声这样的可被分类在表示未出现水印的类别中的内容现在不必受到水印检测。

在指纹计算的情况中,在某些结构中,在物理上远离执行数据库查找地点的地点进行指纹计算。在这种结构中,也计算如语音或噪声这样的不可识别的内容的指纹、或不必识别的如商业广告节目或新闻的内容的指纹。通过应用分类器以“剔除”这种不可识别的内容,降低了需要传送到数据库查找部分中的指纹数量。这也降低了要传送的数据量。

20 本发明的另一个目的是提供一种根据前文的系统,其降低了在先技术所需的手工劳动。

根据本发明通过这样一种系统来实现这个目的,该系统的特征在于:至少一个装置适用于依靠将内容项分类到若干类别之一中的装置的输出端而进行操作。

25 在一个实施例中,当在执行数据库查找的装置没能获得该数据且用于分类的装置将内容项分类到若干预定类别之一中时,用于提交内容项以便由人类来处理的装置被设置成有条件地提交该内容项。

30 参考附图中所示的实施例本发明的这些及其他方面是很显然的并且参考附图中所示的实施例来阐明本发明的这些及其他方面,其中:

图1示意性地示出一种用于利用指纹作为标识符而获得与内容项关联的数据的系统;

图2示意性地示出一种用于利用从水印有效负载中提取的标识符而获得与内容项关联的数据的系统；

图3显示了说明根据本发明方法的实施例的流程图；

图4显示了说明根据本发明方法的另一个实施例的流程图；以及

5 图5显示了说明根据本发明方法的又一个实施例的流程图。

全部附图中，相同的参考标记表示类似或相应的特征。在软件，以及本身代表软件实体，如软件模块或对象中典型地实现附图中表示的一些特征。

10 图1示意性地显示了客户机101、服务器300和指纹数据库310。客户机101可以是如收音机这样的音频装置或如电视机这样的视频信号源。它还可以是移动电话。客户机101通常获得从另一个源中再现的内容项。例如收音机从空中或从电缆连接中拾取广播传输并从中产生/再现可听到的信号。电话可以利用其内置话筒而接收音频或
15 利用其内置照相机而接收视频。

出于简洁的考虑，参考音频剪辑来讨论图1的实施例，虽然本发明同样可以用视频剪辑来工作。

此时服务器300包含输入模块301、指纹模块302、数据库管理系统（DBMS）后端模块303、以及响应模块304。服务器300的任务是获得与由客户机101送给它的音频剪辑关联的数据。通常这个数据是诸如音频剪辑的名称或艺术家这样的元数据，但是它也可以是象人类可在其中购买在音频剪辑中广告了的产品的互联网站这样的数据。
20

输入模块301从客户机101接收音频剪辑。接着将音频剪辑送到指纹模块302。指纹模块302从所接收的音频剪辑中计算指纹。在国际专利申请W002/065782（代理人案号PHNL010110）中描述了一种用于计算鲁棒指纹的方法，不过当然可采用用于计算指纹的任何方法。
25 然后指纹模块302将所计算的指纹作为标识符提供给DBMS后端模块303。

DBMS后端模块303在数据库310上执行查询以从数据库310中检索
30 出一组与所接收的标识符关联的元数据。如图1所示，数据库310包含指纹FP1、FP2、FP3、FP4和FP5形式的标识符以及各自的关联元数据组MDS1、MDS2、MDS3 MDS4和MDS5。以上提及的国际专利申请W0

02/065782也描述了一种将代表未知信号的指纹与保存在数据库中的已识别信号的多个指纹进行匹配以识别该未知信号的有效方法。

可按照多种方式来组织数据库310以最佳化查询时间和/或数据结构。在设计数据库310中的表时,应当考虑指纹模块302的输出。

- 5 在图1所示实施例中,数据库310包含具有包含相应指纹和元数据组的条目(记录)的简单的表。

实现数据库310的另一个方式是建立几个表。第一个表包含多个唯一标识符(主要关键字),其每个都与各自的元数据组相关联。这个表可从各种音乐识别源中获得。艺术家、名称和发行年的组合
10 可以合并以形成唯一标识符,但是这不保证是唯一的,因此最好是采用真正的全球性唯一值。

接着,采用为每个内容项而包含第一个表中的指纹和唯一标识符的条目来建立第二个表。用这种方法,不用必须复制元数据即可使多个指纹与一个元数据组相关联。如果对于一个内容项来说多个指
15 纹是可能的,那么将所有这些指纹保存在第二个表中,所有指纹都与那个内容项的唯一标识符相关联。

然后DBMS后端模块303将所接收的指纹与第二个表中的指纹进行匹配,获得标识符并且将该标识符与第一个表相匹配以获得元数据。如果数据库310是SQL数据库,那么两个表可在该标识符上相连
20 结。DBMS后端模块303将查询结果馈送给将所发现的元数据传送回客户机101的响应模块304。

如果客户机101是移动电话,那么元数据可以例如作为SMS消息或电子邮件消息而传送。如果通过移动电话来发送由输入模块301接收到的音频剪辑,那么通过主叫号码识别或自动号码识别或类似手段
25 来获得电话号。然后输入模块301将呼叫号码提供给响应模块304,以便向同一个号码发送SMS消息。

可替换地,输入模块301可以接收识别用户的另一方式,如当连接到服务器300时,由用户提供用户名或电子邮件地址。为了利用该服务需要注册,然后可通过例如是在用户提供的用户名的基础上检
30 验用户的登记细节而获得目的地址。

还可替换地,所发现的元数据最好是与客户机101的标识符和记录该条目的时间戳一起而被记录在日志文件中。用这种方法,日志

文件含有所处理的内容项的准确报告。因此，这个日志文件可用作例如通过特定信道广播了些什么的证明。可由如美国作曲家、作家和出版商协会（ASCAP）或荷兰BUMA/Stemra这样的版权资料交换所来采用这个日志文件以确定应当给特定版权所有者多少版税。这种版税经常是根据特定歌曲广播的次数而估计的，并且这个列表提供了公正的第三方的准确估计。电台可能在过低估计或过高估计其广播特定内容项的数目，或者可能不愿意提供足够的细节。

当然这个列表还对许多其他目的很重要。如果以这种方法识别的内容项包含广告或宣传消息，那么可将该列表用于证明（或反驳）在某个时间广播特定的广告或消息。用这种方法，电台可显示它满足了对广告客户的合同义务。

在国际专利申请序号PCT/IB03/00260（代理人案号PHNL020101）中描述了进一步的改进。根据这个文献，服务器300除了处理来自客户机101的请求之外还将监测一个或多个广播信道。与这些信道上的内容关联的元数据将被复制到次数据库中。因此，该次数据库含有少量的条目。从而在次数据库上匹配会比在第一数据库上匹配更快。仅当在次数据库中发现无匹配时才执行在主数据库中的匹配。因为，可以预料对于在所监测的传送信道上传送的内容项来说会产生许多请求，于是仅利用更小且更快的次数据库来应答许多请求。因此，平均说来，降低了匹配指纹所需的时间。

图2示意性地显示了一种图1系统的改变，其中指纹模块302替换为水印提取模块322。现在把由接收模块301所接收的剪辑传送到这个企图检测和提取在剪辑中的水印数据的水印提取模块301中。然后将所提取数据用作可被传送给DBMS后端模块303以用于数据库查找的标识符以便获得与该标识符相关的元数据。

如图2所示，数据库320包含标识符ID1、ID2、ID3、ID4和ID5以及各自的元数据关联组MDS1、MDS2、MDS3、MDS4和MDS5。服务器300的其余操作与参考图1在以上所讨论的相同。

在服务器300中，是否利用指纹或水印技术或其它类型的识别机制来操作，可能会发生在数据库310、320中不能发现特定标识符。在这种情况下可将所讨论的内容项提交给输出端以便操作员200能够手动地检查该内容项并且试图手动辨认。输出端可以例如是终

端、扬声器、显示器或到网络的连接。可将在数据库中未发现的内容通过电子邮件或用其它方式被传送给远程地点。因此，远程地点处的操作员能手动地检查该内容。

5 在图1和2中，在这种情况下将在终端210处利用扬声器播放该内容项，并且期望操作员200收听该项并且输入或选择元数据。接着将这个元数据提供给响应模块304以便可将它发送回给客户机101或记录在日志文件里，或者以其它方式来使用。如果操作员200不实时输入元数据，或没能认出内容项，那么应当由响应模块304向客户机101发送一个适当的消息。

10 本发明是建立在对服务器300将受益于用于将内容项分类到若干类别之一中的自动分类技术的使用的理解的基础之上的。如果依据内容项的分类而有条件地执行上面描述的至少一个步骤，那么将改进服务器300的操作。

15 为此，服务器300具有自动分类器309。本说明书结尾处会论述若干自动分类方法。用于自动分类的典型系统包括特征提取级，其后是将特征映射到内容的一个或多个类别的分类级。

20 分类器309分析内容项并且将内容项分类到若干类别之一中。例如分类可能是象“音乐”和“非音乐”(如“语音”或“噪声”)或者“电影”对“商业广告节目”那样简单。也可利用更详细的分类，如流派分类、特定内容的自动检测加亮以及自动的讲话者辨别。这种分类方法提供将音频分类到象“经典”、“摇滚”、“语音”、“爵士”、“说唱”等这样的类别中、或者将视频分类到象“电影”、“商业广告节目”、“新闻”等这样的类别中。

25 根据本发明的一个方面，如果数据库查找没能获得该数据并且将内容项分类到若干预定类别之一中，则提交内容项以便由人类来处理。例如，在广播监控的情况中，可采用在“音乐”和“非音乐”之间进行区分的简单的两类分类器。在这种情况下，仅需要手动识别那些未自动识别的并且分类为“音乐”的音频。由于大多数未识别的音频由语音组成，因此可实现手工劳动的显著降低。

30 在一个实施例中，根据内容项分类到其中的那个类别而从多个输出端中选出将内容项提交给其以进行处理的输出端。如果采用多类分类器，那么人类也可在较高级别上管理指纹数据库的内容。用特

定流派（流行的、经典的等等）来标记未识别的音乐的更成熟的分类器增加了可将未识别的音频自动分配给具有相应流派专门技术的适当人员的额外可能性。

5 为此，对服务器300进行编程以将内容项提交给多个输出端之一。每个输出端与一个特定类别相关联。例如，当操作员在特定终端登录时可手动或自动地形成这种关联。操作员可输入这个信息，或者可在数据库或这个操作员的用户需要文件中注册该信息以便服务器300能够自动获悉该信息。然后服务器300将该内容项提交给与将内容分类到其中的那个类别相关联的特定终端。

10 这种成熟的分类也允许调整数据库310、320的内容。例如，由于清楚地指示了许多对元数据的请求涉及到爵士音乐并且目前数据库并未含有足够的爵士音乐，因此如果把为了人工识别而提交的大多数内容项分类为爵士乐，那么应当将更多的爵士音乐添加到数据库310、320中。

15 根据本发明的另一个方面，如果将内容项分类到若干预定类别之一中则执行数据库查找。例如，如果数据库310、320仅分别含有音乐或广告，只有当将内容项分类成“音乐”或“广告”时才执行数据库查找。避免了保证会失败的这种查找方式。

20 在移动电话业务的情况中，很大部分的所有所记录的未识别的音频剪辑仅包含噪声。这通常会在移动电话距音频源太远时发生。当采用两类分类器时，可以忽略分类为“非音乐”或“噪声”的所有记录。因此，操作员200仅必须收听并且识别分类为“音乐”的录音。因此，可更有效地完成未识别的音频剪辑的识别。

25 然而，在这个实施例，重要的是选择一种具有极低的错误否定率的分类技术。即错误地没有分类到一个预定类别中的项的数目应当非常低。发明人发现，实际上对于短音频剪辑（5到10秒）来说，典型的两类分类器具有约90%的分类性能。这意味着10%的音乐被列为非音乐。所组合的分类/目标识别系统的综合性能将因此而降低到90%以下，这明显是不希望有的。

30 在又一个方面（未显示），依据由分类器309将内容项分类到若干类别之一中，指纹模块302或水印提取模块322用于有条件地操作。在这个实施例中，在如“噪声”或“非音乐”这样的类别中所分类

的内容不是必须受到水印有效负载提取或指纹计算。

值得注意的是，有可能组合上述的实施例。例如依据分类可有条件地进行指纹计算，以及基于分类还可能自动选择由哪个输出端提交未识别的内容。否则，两类分类器可用于确定获得标识符是否有意义，以及多类分类器可用于确定是否将内容提交给输出端，或者可用于决定要将内容提交给哪个输出端。

图3显示了一个说明根据本发明方法的实施例的流程图。该方法起始于步骤30。在步骤31中，接收如音频或视频剪辑这样的内容项。在步骤32，例如通过从嵌入内容项中的水印的有效负载中提取标识符或者通过在内容项上计算指纹而获得内容项的标识符。

在步骤33中，执行数据库查找以检索与该标识符关联的数据。如果在步骤34中判定查找成功，则获得所请求的数据并在步骤39结束该方法。如果查找失败，则该方法继续到步骤35。

在步骤35中，将内容项分类到若干类别之一中，例如“音乐”或“噪声”。然后在步骤36中判断是否基于分类而将内容项提交给操作员200。例如，如果将内容分类为“音乐”，则将它提交给操作员200。

如果判定是要提交内容项，那么在步骤37接收操作员200输入的或从数据库中选出的数据。然后该方法还是在步骤39结束。如果不是必须提交该内容项，那么该方法直接在步骤39结束。在这种情况下，向客户机101提供某种出错消息。

图4显示了说明根据本发明方法的另一个实施例的流程图。此处不再重复与图3的步骤相同的步骤。

步骤44中判断是否基于步骤35中所获得的分类而执行数据库查找以检索与该标识符关联的数据。例如，如果将内容项分类为“噪声”，则数据库查找不会检索到任何匹配并因此可被跳过。如果希望数据库查找，那么该方法转到步骤33，否则在步骤39结束。如果在步骤46中，判定了步骤33的查找失败，则该方法转到步骤37，否则在步骤39该方法结束。

图5显示了说明根据本发明方法的又一个实施例的流程图。此处不再重复与图3或4的步骤相同的步骤。

在步骤54判断是否基于步骤35中所获得的分类而获得内容项的标识符。例如，如果将内容项分类为“噪声”，则没有水印要检测，

在噪声上计算指纹也不会得到从内容项的有意义标识符。因此，在这种情况下，该方法立即结束。在其它情况下，该方法转到步骤32以获得标识符，以及转到步骤33以执行数据库查找。从那以后步骤与图4中的相同。

5 为了使所属领域技术人员能够根据本发明而构造服务器，下面给出一些对现有分类技术的参考。值得注意的是本发明不依赖于一个特定的分类技术。选择要采用哪个特定技术取决于情况，例如两类或者多类分类器是否必要、人类是否希望仅向服务器提供“音乐”及“噪声”等等。

10 一些两类分类器技术为：

E. Scheirer和M. Slaney, 1997年，德国慕尼黑，Proc. ICASSP, 第1331-1334页，名为Construction and evaluation of a robust multifeature speech/music discriminator的文章。

15 G. Lu和T. Hankinson, 1998年，北京，4th int. Conference on signal processing, 名为A technique towards automatic audio classification and retrieval的文章。

20 R. Jarina、N. Murphy、N. O'connor和S. Marlow, 2001年，Malta, WSES International Conference on Speech, Signal and Image Processing, 名为Speech-music discrimination from MPEG-1 bitstream的文章。

一些多类分类器技术为：

M. Zhang, K. Tan和M. H. Er, 1998年，J. Audio. Eng. Soc., 146: 836-844, 名为Three-dimensional sound synthesis based on head related transfer functions的文章。

25 T. Zhang和C. C. J. Kuo, 2001年，IEEE Transactions on speech and audio processing, 9: 441-457, 名为Audio content analysis for online audiovisual data segmentation and classification的文章。

30 J. Foote, 1997年，Proc. AAAI 1997 Spring Symposium on Intelligent Integration and Use of Text, Image, Video, and Audio Corpora, 名为 A similarity measure for automatic audio classification的文章。

M. S. Spina和V. W. Zue, 1997年, 希腊Rhodes, Proceedings of Eurospeech, 名为Automatic transcription of general audio data: Effect of environment segmentation on phonetic recognition的文章。

- 5 G. Tzanetakis、G. Essl和P. Cook, Princeton, NJ, Proceedings International Symposium for Audio Information Retrieval (ISMIR), 名为Automatic musical genre classification of audio signals的文章。

- 10 D. Pye, 2000年, 在ICASSP 2000, 第四卷, 第2437-2440页, 名为Content-based methods for the management of digital music的文章。

- 15 D. N. Jiang、H. J. Zhang、J. H. Tao和L. H. Cai, 2002年瑞士洛桑, Proceedings of ICME : 2002 IEEE international conference on multimedia and expo , 名为Music type classification by spectral contrast feature的文章。

值得注意的是, 以上提及的实施例说明而不是限制本发明, 并且所属领域技术人员在不脱离附加权利要求范围时可设计出许多替换的实施例。

- 20 例如, 连接到个人计算机的麦克风可用作客户机101。然后计算机从麦克风录音, 并例如经由互联网作为电子邮件消息或利用FTP、HTTP文件上载或类似机制而将该记录传送给服务器300。具有记录装置的便携式设备也可以用来产生这种记录。然后将该便携式设备经由电话线或网络接线而连接到服务器上。然后由于以数字格式传送该项, 因此如网络收音机这样的其他传送信道允许直接录音以及内容项的传输。

在权利要求中, 放入括号中的所有参考符号不应看作是对权利要求的限制。单词“包含”不排除除了权利要求中列举的那些元件或步骤之外的元件或步骤的存在。元件前面的单词“一”或“一个”不排除多个这种元件的存在。

- 30 可以借助于包含几个不同的元件的硬件、以及借助于适当地程序控制计算机而实现本发明。在设备中, 权利要求列举了几个装置, 可通过硬件的同一个对象而体现几个这些装置。在互相不同的从属

权利要求中叙述某种手段的仅有事实不表示这些手段的组合不能用作优点。

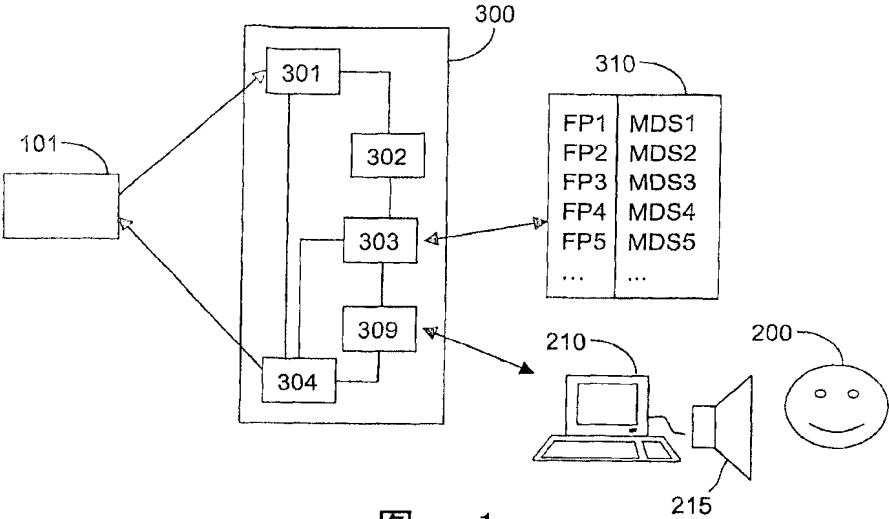


图 1

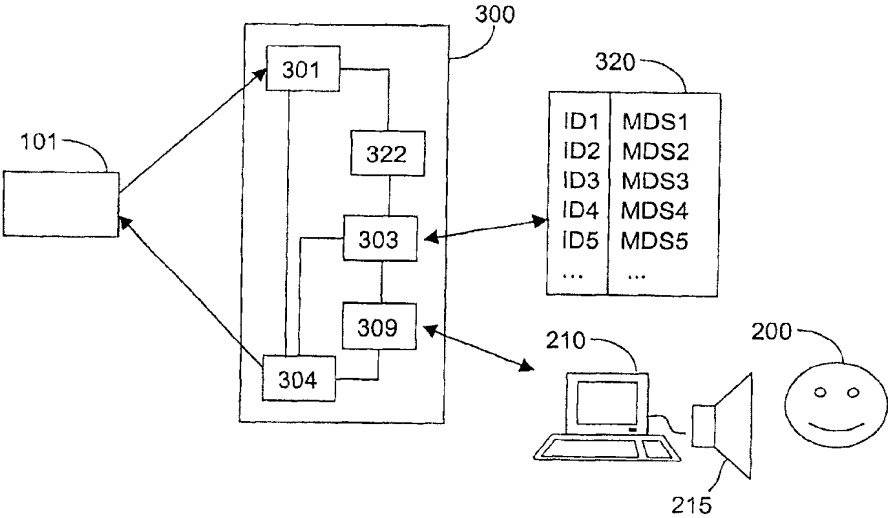


图 2

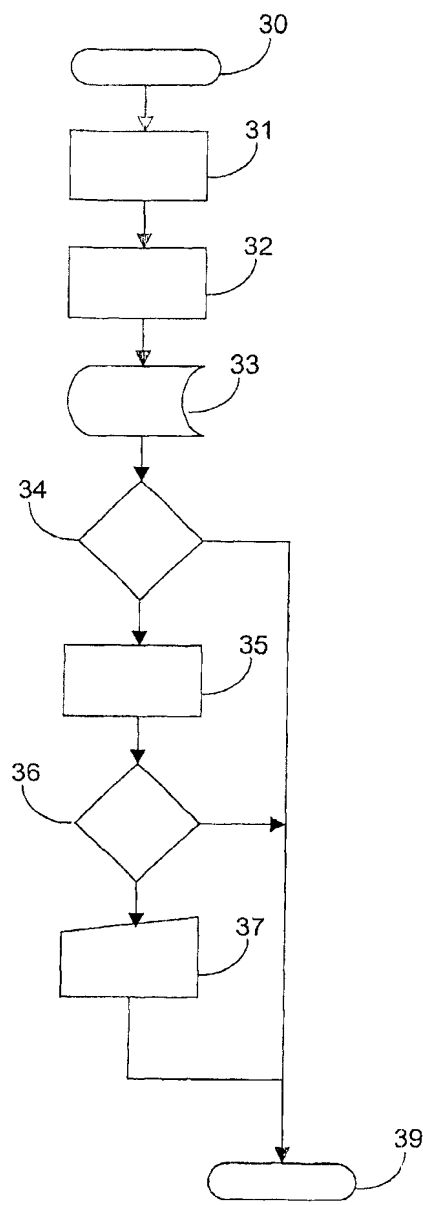


图 3

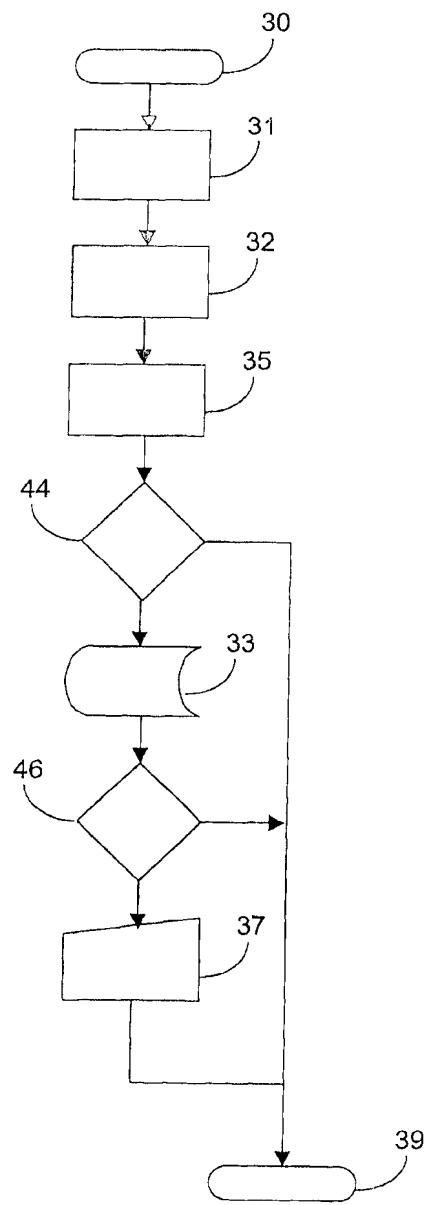


图 4

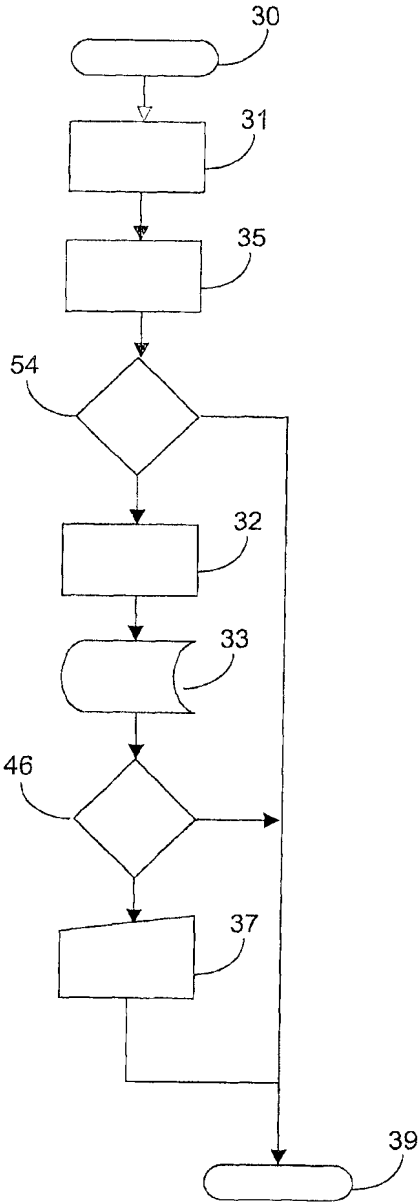


图 5