[19] 中华人民共和国国家知识产权局

[51] Int. Cl.
H04N 7/00 (2006.01)
G11B 27/00 (2006.01)

[12] 发明专利申请公布说明书

[21] 申请号 200480037954.4

[43] 公开日 2007 年 1 月 10 日

[11] 公开号 CN 1894964A

[22] 申请日 2004.12.7
[21] 申请号 200480037954.4
[30] 优先权
[32] 2003.12.18 [33] EP [31] 03104799.6
[86] 国际申请 PCT/IB2004/052695 2004.12.7
[87] 国际公布 WO2005/062610 英 2005.7.7
[85] 进入国家阶段日期 2006.6.19
[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 M·巴比里 G·E·梅肯坎普
B·P·G·赫特 B·梅里亚多

[74] 专利代理机构 中国专利代理(香港)有限公司
代理人 程天正 王 勇

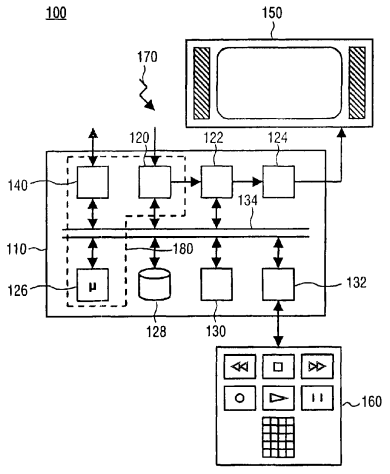
权利要求书 2 页 说明书 8 页 附图 3 页

[54] 发明名称

创建视听数据流的多媒体概要的方法和电路

[57] 摘要

随着消费者可检索到的视听数据量的快速增长，对于视听数据(例如电影)的适当概括的需求也在增加。为此，本发明提供一种创建视听数据流(例如电影)的多媒体概要的方法。首先，检索文本概要(204)。接着，对视听数据流进行分段(208)并且从视听数据流(210)和文本概要(206)中提取信息。最后，选择承载着与由文本概要所承载的信息相匹配的信息的分段(212)。大量的电影和电视剧的概要可在因特网上获得，所述概要是由爱好者制作并且是为爱好者制作的，这为创建多媒体概要提供了可靠的来源。



1、创建视听数据流的多媒体概要的方法，包括如下步骤：

a）从外部源获得（204）该视听数据流的已制作好的文本概要；

b）对该文本概要进行分析（206）以提取信息；

c）对该视听数据流进行分段（208）和分析（210）以提取信息；

d）从该视听数据流中选择（212）包括与从该文本概要中提取的信息相匹配的信息的分段；以及

e）对所选择的分段进行组合（214），从而形成多媒体概要。

2、如权利要求1所述的方法，其中所述外部源是下列各项的至少其中之一：

a）图文电视；

b）电子节目指南；或者

c）因特网服务器。

3、如权利要求1所述的方法，其中

a）该视听数据流包括承载着对应于该视听数据流的字幕的子流；以及

b）从该视听数据流中提取的信息是通过对字幕进行分析而从该视听数据流中提取的。

4、如权利要求3所述的方法，其中所述子流承载着：

a）闭路字幕数据；

b）图文电视字幕数据；和/或

c）图形格式的字幕。

5、如权利要求1所述的方法，其中从所述文本概要中提取的信息是关键词。

6、如权利要求5所述的方法，其中所述关键词是包含在文本概要中的名词、形容词和/或动词。

7、如权利要求1所述的方法，其中利用与从所述文本概要中提取的信息相关的信息来扩展从所述文本概要中提取的信息。

8、如权利要求6所述的方法，其中从所述文本概要中提取的信息是名词、形容词和/或动词，并且利用与从所述文本概要中提取的名词相关的其他名词、形容词和/或动词来扩展所提取的信息。

9、如权利要求7所述的方法，其中所述其他名词、形容词和/或动

词是从所述文本概要中提取的名词、形容词和/或动词的同义词。

10、如权利要求 5 所述的方法，其中：

a）该视听数据流包括承载着字幕的子流；以及

b）通过对字幕进行分析从该视听数据流中提取信息；以及

c）从该视听数据流中选择包括与从所述文本概要中提取的信息相匹配的信息的分段的步骤包括选择其中所述字幕包括至少一个关键词的至少一个分段。

11、如权利要求 1 所述的方法，其中从该视听数据流和所述文本概要中提取的信息包括单词，并且当从该视听数据流中提取的至少一个第一单词与从所述文本概要中提取的至少一个第二单词匹配时，选择该视听数据流的一个分段。

12、如权利要求 1 所述的方法，其中在所述多媒体概要被重放时组合所述分段。

13、用于创建视听数据流的多媒体概要的电路（180），包括：

a）通信单元（140，120），用于从外部源获得该视听数据流的已制作好的文本概要；以及

b）处理单元（126），其被构建为：

i.）分析该文本概要以提取信息；

ii.）对该视听数据流进行分段和分析以提取数据；

iii.）从该视听数据流中选择包含与从该文本概要中提取的信息相匹配的信息的分段；以及

iv.）组合所述分段，从而形成多媒体概要。

14、用于处理视听数据的设备（110），其包括根据权利要求 10 的电路。

15、计算机程序产品，其包括对处理单元（126，304）进行编程以执行如权利要求 1 所述的方法的代码。

16、承载着如权利要求13所述的计算机程序产品的数据载体（130，310）。

创建视听数据流的多媒体概要的方法和电路

本发明涉及一种创建视听数据流的多媒体概要的方法。

本发明还涉及一种创建视听数据流的多媒体概要的电路。

本发明还涉及一种包括这种电路的用于处理视听数据的设备。

此外，本发明还涉及一种计算机程序产品，其包括对处理单元进行编程的代码。

此外，本发明还涉及一种承载这种计算机程序产品的数据载体。

据报道，很长一段时间以来，消费者可得到的存储量以及消费者使用的存储量正在增加。并且，呈现给消费者的内容量以及消费者可得到的内容量正在不断地增长。为了对由消费者存储或者为消费者存储的所有内容提供适当的概述，适当的概要是必不可少的，特别是对于视听数据流（例如电影）来说更是如此。

对于消费者来说，亲自对他或她可获得的每一部电影做概括是不可能的。因此，非常希望这个概括电影的处理是自动的。

专利申请 US 2002/0083471 公开了一种提供视频节目的多媒体概要的系统和方法。创建多媒体概要的处理是从根据 WO 02/041634 所公开的方法自动创建文本概要开始的。虽然自动创建文本概要不需要用户交互，但是其需要大量的处理能力并因此导致昂贵的电路。此外，自动创建文本概要还很容易由于选择了视频节目的错误部分而导致失败。出现这种情况的原因在于：用于自动创建文本概要的电路根据几个规则来工作，而所述规则可能不适用于每一个视频节目。

本发明的一个目的是提供一种需要更少处理能力的创建多媒体概要的方法和电路。为了实现该目的，本发明提供一种创建视听数据流的多媒体概要的方法，该方法包括如下步骤：从外部源获得该视听数据流的已制作好（ready-made）的文本概要；对该文本概要进行分析以提取信息；对视听数据流进行分段和分析以提取信息；从视听数据流中选择包括与从文本概要中提取的信息相匹配的信息的分段；以及将所选择的分段进行组合从而形成多媒体概要。

本发明是基于这样一种认识而建立的：可以获得具有电影和电视剧等视频节目的已制作好的文本概要的许多数据库。用于通过例如因特网检索这些文本概要的电路非常多并且可以以很低的价格获得，而且只需要最低的处理能力。此外，文本概要通常可以免费获得。

此外，这些概要通常是由电影评论家、电影爱好者或电视剧爱好者作出的，这些人了解电影和风格，并且他们知道电影或某集电视剧的亮点在哪。通过这种方式，可利用专门的智力规则来创建文本概要。通过这种方式，比起应用对于人脑所使用的规则而言几乎是原始的规则的电路，提供了更为准确的文本概要。

在根据本发明的一个实施例中，视听数据流包括承载对应于该视听数据流的字幕的子流；以及从该视听数据流中提取的信息是通过对字幕进行分析而从视听数据流中提取的。

这个实施例的一个好处是：字幕很容易提取，因为不必从待概括的其他视频数据（例如电影）中提取字幕。

在根据本发明的另一个实施例中，从文本概要中提取的信息是关键词。

这个实施例的一个好处是：（可以在所述子流中获得的）单词很容易处理，因为可以将所述单词转换成字母数字式数据并将其作为字母数字式数据进行处理。

根据本发明的另一个实施例，利用与从文本概要中提取的信息相关的信息来扩展所述从文本概要中提取的信息。

这个实施例的一个好处是，短文本概要可以通过这种方式来提供更多信息或更详细的信息。特别地，由图文电视提供的概要是相当短的，因为它们通常必须显示在一页上。通过扩展从该概要中提取的信息，可以获得附加的信息以用来搜索待概括的视听数据流中的匹配分段。

在根据本发明的又一个实施例中，在多媒体概要被重放的时候组合所述分段。

这个实施例的一个好处是：不需要大量的额外存储空间来存储整个多媒体概要，这是因为可以从原始视听数据流中重放分段。多媒体概要的建立可以在多媒体概要的重放之前离线地（off-line）进行。其结果可以是关于待概括的原始视听数据流的播放列表。

根据本发明的用于创建视听数据流的多媒体概要的电路包括：通信单元，用于从外部源获得视听数据流的已制作好的文本概要；以及处理单元，其被构建为：分析文本概要以提取信息；对视听数据流进行分段和分析以提取信息；从视听数据流中选择包含与从文本概要提取的信息相匹配的信息的分段；以及组合这些分段，从而形成多媒体概要。

根据本发明的视听数据处理设备包括这种电路。

根据本发明的计算机程序产品包括对处理单元进行编程以执行根据本发明的方法的代码。

承载根据本发明的计算机程序产品的数据载体承载着这样的计算机程序产品。

现在将通过附图对本发明的各实施例进行详细的说明，其中：

图 1 示出了根据本发明的设备的一个实施例；

图 2 是一个流程图，其描述了根据本发明的方法的一个实施例；

图 3 示出了根据本发明的数据载体的一个实施例。

图 1 示出了一个消费者电子系统 100，其包括视频记录器 110、电视机 150 以及控制装置 160，其中该视频记录器作为根据本发明的设备的一个实施例。视频记录器 110 被设置成接收并记录视听数据流以及由信号 170 载送的与这些视听数据流有关的交互式应用。

为此，视频记录器 110 包括用来接收信号 170 的接收器 120、多路分解器 122、视频处理器 124、例如微处理器 126 的用于控制该视频记录器 110 所包括的部件的中央处理单元、作为存储装置的硬盘驱动器 128、程序代码存储器 130、用于从控制装置 160 接收信号的用户命令接收器 132 以及用于连接该视频记录器 110 所包括的部件的中央总线 134。

该视频记录器进一步包括网络接口单元 140，其用于连接至网络（例如因特网或 LAN）。该网络接口单元 140 可以被具体实现成模拟调制解调器、ISDN、DSL 或有线电视调制解调器或者 UTP/以太网/TCP-IP 网络接口。

接收器 120 被设置成调谐到一个广播（音频或视频）频道，并从

信号 170 中导出该广播频道的数据。可以通过任何已知的方法（有线电视、地面、卫星、宽带网络连接）或其他任何分发视听数据的方法来接收信号 170。信号 170 甚至可以从另一个消费者电子设备的输出中导出。接收器 120 输出载送至少一个视听数据流的基带信号。

多路分解器 122 被设置成从可能包含在由接收器 120 输出的基带信号中的其他数据中多路分解出视听数据。视频处理器 124 被设置成以可以由电视机 150 呈现的方式来呈现由多路分解器 122 输出的视听数据。可以以各种模拟格式（例如 SECAM 和 PAL）或者以数字格式来提供所述输出。

存储在程序代码存储器 130 中的数据使得微处理器 126 能够执行根据本发明的方法。程序代码存储器 130 可以被具体实现成闪速 EEPROM、ROM、光盘或其他任何类型的数据承载介质。

所述存储装置还可以被具体实现成光盘驱动器（例如 DVD 或蓝光驱动器），并且适于存储由接收器 120 或网络接口单元 140 接收的内容，以便将来重现在电视机 150 上或者将来通过网络接口单元 140 进行散播。所述内容可以在存储之前被处理。

为了给视频记录器 110 的用户提供针对存储在硬盘驱动器 128 内的所有数据的良好的概述，微处理器 126 为视听数据流创建概要，其中所述视听数据流例如是电影、电视节目或者其他存储在硬盘驱动器 128 内或者由接收器 140 接收的数据。这可以自动完成，也可以由用户发起。

图 2 是一个流程图 200，其对根据本发明的创建视听数据流的概要的方法的一个实施例进行了描述。下表 1 中提供了各个块中的处理步骤。下面将结合图 1 对该处理进行说明。

附图标记	处理步骤
202	启动概要处理
204	检索已制作好的文本概要
206	分析检索到的概要
208	对将要概括的流进行分段
210	对将要概括的流的分段进行分析
212	选择具有与从文本概要提取出的信息相匹配的信息的分段

214	对所选择的分段进行组合
216	返回概要

表 1

在处理步骤 202 中，该处理被自动地（由微处理器 126 所运行的代理）启动或者由用户活动（例如操作控制装置 160）来启动。

接着，在处理步骤 204 中，检索待概括流的已制作好的文本概要。电影的概要可以在许多地方获得，例如在因特网址 http://www.cinema.nl 处获得。此外，图文电视和电子节目指南（EPG）也提供电影和其他节目（例如电视剧）的文本概要。特别是对于肥皂剧，在广播了一集之后，其概要就提供其完整情节。

在一个有利的实施例中，通过网络接口单元 140 从因特网服务器检索概要。在本发明的另一个实施例中，从图文电视数据中检索概要，其中图文电视数据被多路复用在广播信号中，并在多路分解器 122 中被从广播信号中导出。对于模拟电视信号，图文电视数据被多路复用在垂直消隐间隔中。在数字电视的情况中，图文电视数据可以在单独的流中与视听数据流一起提供。图文电视数据还可以通过因特网（例如http://teletekst.nos.nl/）获得，并且可以通过网络接口单元 140 来检索。

虽然在很多情况下图文电视数据和 EPG 数据是与视听数据流一起接收的并因此实际上可在视频记录器 110 中获得，但是在这个应用的情境中，其被认为是从外部源检索到的，这是因为通过这些措施检索到的文本概要的产生与视听数据流的创建（例如电影的拍摄）是分离的。

在本发明又一个实施例中，从电子节目指南中获得概要。这个节目指南可以以和检索图文电视数据相同的方式从广播信号或者从因特网中获得。

以这种方法获得概要的一个主要优点是：不必从待概括的视听数据流中作出概要，相反，概要已经是可获得的。

在检索了概要之后，在步骤 206 中分析概要以提取信息。在一个优选实施例中，从该概要中提取出关键词。这些关键词可以是不止出现一次的、或者是在电影等的标题中出现的动词、名词或形容词。

在另一个实施例中，该信息提取处理搜索与从文本概要中提取的

关键词相关的单词。这些相关单词可以是同义词，但是我们还可以考虑其他关系，例如“传真”与“电话”有关，“汽车”与“驾驶”有关。在一个实施例中，与所提取的信息有关的信息是利用网络接口单元 140 从外部数据库中检索的。在另一个实施例中，用于搜索附加的相关信息的数据库被存储在硬盘驱动器 128 中。

所述数据库还可以包括不被认为是关键词的单词。例如“to be”或者其他使用频率很高的动词的所有变化形式。

然后，在处理步骤 208 中，利用例如在相同申请人的申请 WO02/093929 中公开的已知方法对视听数据流进行分段。

在对多媒体数据对象进行了分段之后，在处理步骤 210 中对这些分段进行分析以便提取信息。本文中提出了用于从分段中提取信息的本发明的各种实施例。当多媒体数据对象是电影并且在所述电影本身中提供有字幕时，可以从其他视频数据中提取出字幕，并且可以利用 OCR 算法来读取字幕。

当把字幕作为附加的数据（例如图文电视或闭路字幕）以字母数字格式来提供时，可以以一种简单的方式自动地提取信息。

前面段落中所讨论的两个选项的中间选项也是可能的。在 DVD 上，字幕可以由内容供应商在单独的流中以图形格式来提供。为了提取信息，可以容易地将字幕转换成字幕数字式字符，因为不必从所述字幕所针对的视听数据流中的视频数据中提取所述字幕。

在本发明的另一个实施例中，电影中的人物的语音是利用语音识别算法来提取的。虽然这种处理需要大量处理能力，但是在未来的几年内微处理器的处理能力预计将进一步提高。这使得可以在操作过程中（on the fly）使用廉价的商用微处理器来进行语音识别。

从字幕或转换后的语音文本中提取名词、动词和/或形容词，正如在处理步骤 206 中从概要中提取数据一样。

除了文本之外，也可以从视听数据流中提取其他信息，例如爆炸、动作场景、对话或者主要人物的面部（通过面部识别的方式）。

当视听数据流已经被分段并且已经从文本概要和视听数据流中提取出信息时，在处理步骤 212 中为多媒体概要选择分段。这是通过分析从文本概要中提取的信息以及搜索包含匹配信息的分段来进行的。在本发明的一个实施例中，当一个分段包含至少一个包含在从文本概

要中提取出的信息中的关键词时，为多媒体概要选择该分段。

在本发明的另一个实施例中，当一个分段包含相关的关键词的组合（例如“警察”和“逮捕”，或者“荷兰”和“木屐”时），为多媒体概要选择该分段。这样的组合也被看作是包含在从视听数据流中提取的信息中的单词和包含在从文本概要中提取的信息中的单词之间的匹配。

承载着对于理解视听数据流所表示的故事的情节来说很重要的、不同于（口头）文本的其他信息也可以被包括在概要中。这样的分段例如是具有动作场景和爆炸的分段。

在本发明的一个实施例中，除了由分段所承载的信息之外，场景必须满足其他要求才能被选择在多媒体概要中。这种要求是场景的长度和各种场景的位置，因为在大多数情况下，希望用于概要的分段是从视听数据流的整个长度选择的，而不希望出现所选场景的 90％是从该流的前 10％选择的情况。

在视听数据流的适当分段已经被选择后，这些分段被组合在一个新的视听数据流中，从而形成必须为之制作概要的原始视听数据流的多媒体概要。这是在处理步骤 214 中进行的。优选地，这些分段按照其在原始视听数据流中出现的顺序而被组合。

然而，在本发明的又一个实施例中，这些分段按照包含在所述分段中的信息出现在文本概要中的顺序而被组合。在本发明的再一个实施例中，所述分段是以时间顺序排列在多媒体概要中的。这意味着当原始视听数据流包括例如电影中的角色的倒叙（flash-back）时，这些倒叙首先被放入多媒体概要中，接着是其他分段。

在本发明的另一个实施例中，该方法返回一个具有指向原始视听数据流中的各场景的指针的播放列表。这个实施例的一个好处是，不必为多媒体概要存储单独的流。

最后，在处理步骤 216 中返回多媒体概要。该多媒体概要可以被存储在硬盘驱动器 128 中。

本领域技术人员可以认识到：流程图 200 所描述的处理的各处理步骤不必按照所呈现的顺序来执行。例如，还可以在视听数据流已被分段并且已经从其中提取信息之后检索概要。此外，可以同时执行各步骤。

对本领域技术人员来说，对上面的说明中所述的实施例进行各种修改是显而易见的。并且可以在不脱离本发明范围的情况下改变各实施例的特征。

例如，可以不扩展从文本概要中提取的信息，而是扩展从视听数据流中提取的信息或者扩展从上述两个信息源中提取的信息。

此外，虽然根据本发明的方法的各实施例被描述为主要由单个处理单元（微处理器126（图1））以及在较低程度上由接收器120（图1）和网络接口单元140（图1）（所有的三个部件形成了作为根据本发明的电路的一个实施例的电路180）来执行，但是本发明的其他实施例也是可能的，其中一个或多个分离的步骤由分离的部件（比如ASIC之类的专用电路）来执行。

本发明可以被具体实现为计算机程序产品，其使得通用计算机（例如图3中的个人计算机300）能够执行根据本发明的方法。

图3还示出了包括对个人计算机300进行编程以执行根据本发明的方法的数据的数据载体310。

为此，将数据载体310插入到包括在个人计算机300中的盘驱动器302中。该盘驱动器302从数据载体310中检索数据，并且将其传送到微处理器304以对微处理器304进行编程。接着，编程后的微处理器304执行根据本发明的方法。

个人计算机300包括通信单元306，以便获得待概括的视听数据流的文本概要。该通信单元306可以被具体实现为模拟的、有线电视或DSL调制解调器、被具体实现为网络接口（UTP、以太网、TCP-IP）或者本领域技术人员所知的其他任何类型的通信单元。

总之，本发明涉及如下内容：

随着消费者可检索到的视听数据量的快速增长，对于视听数据（例如电影）的适当概括的需求也在增加。为此，本发明提供一种创建视听数据流（例如电影）的多媒体概要的方法。首先，检索文本概要（204）。接着，对视听数据流进行分段（208）并且从视听数据流（210）和文本概要（206）中提取信息。最后，选择承载着与由文本概要所承载的信息相匹配的信息的分段（212）。大量的电影和电视剧的概要可在因特网上获得，所述概要是由爱好者制作并且是为爱好者制作的，这为创建多媒体概要提供了可靠的来源。

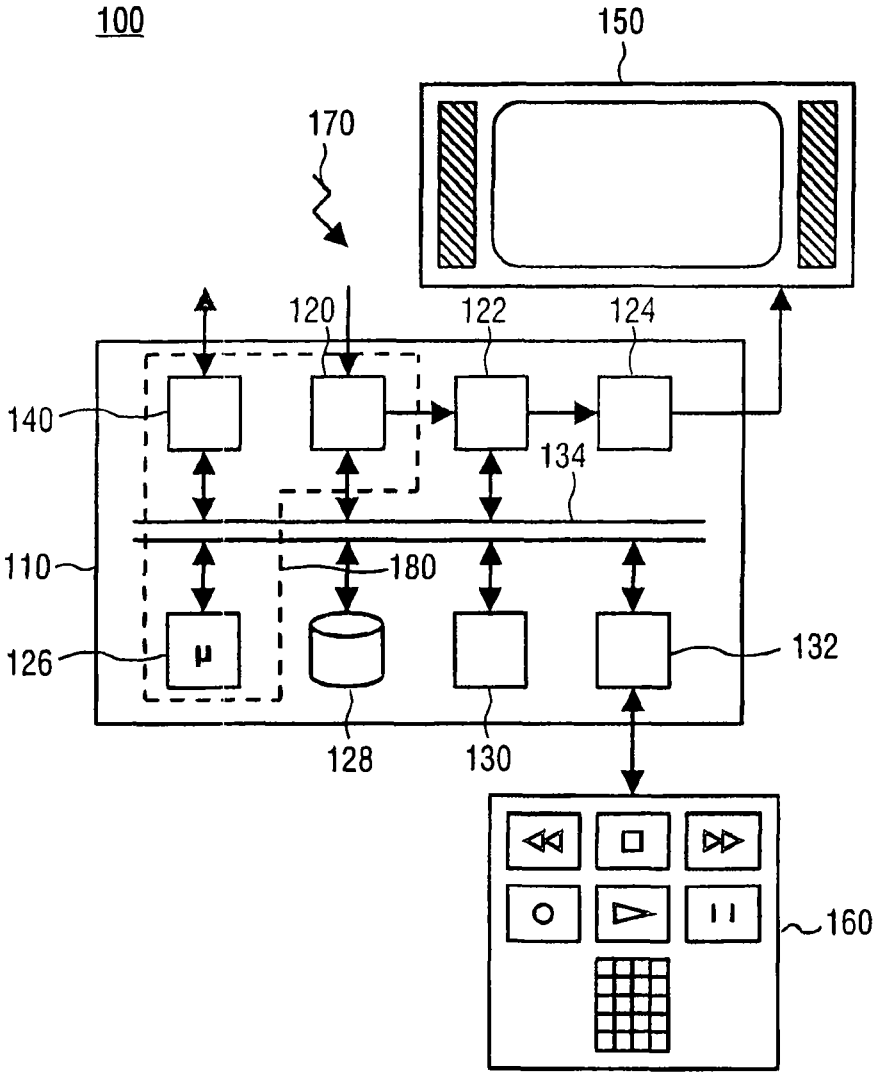


图 1

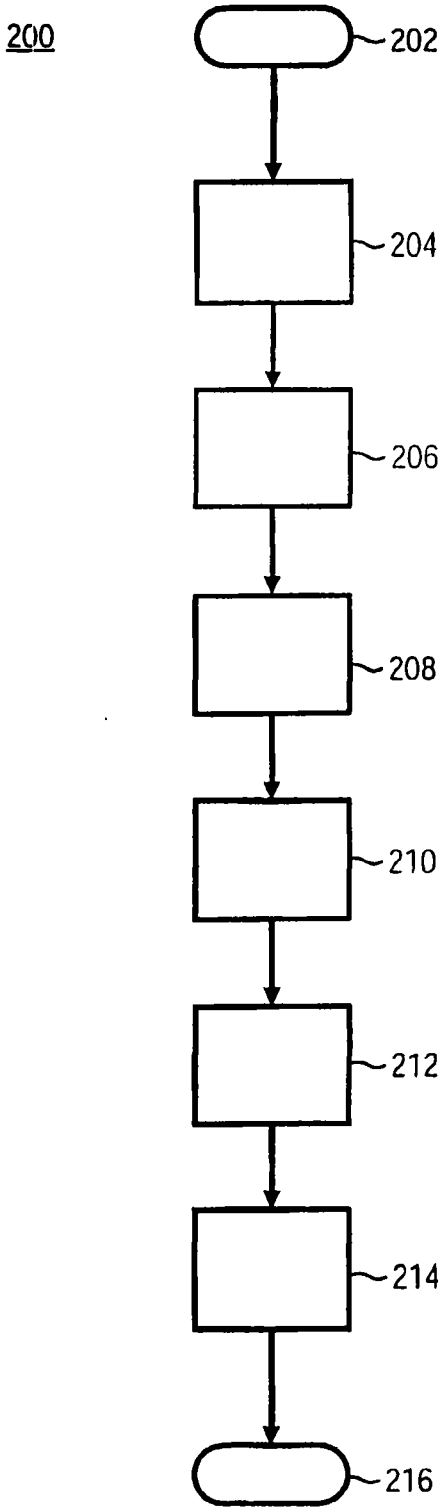


图 2

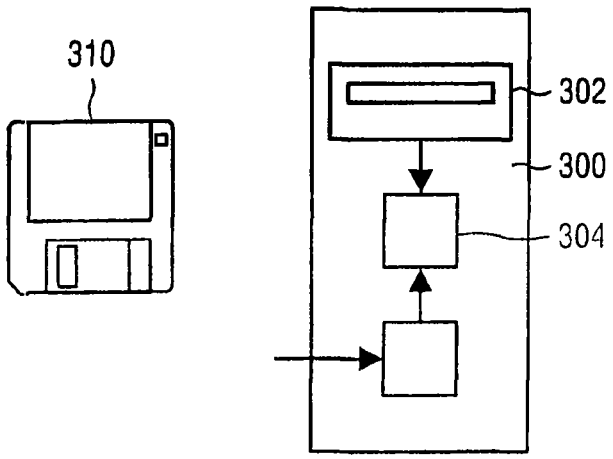


图 3