



[12] 发明专利申请公开说明书

[21] 申请号 200380102504.4

[43] 公开日 2005 年 12 月 14 日

[11] 公开号 CN 1708978A

[22] 申请日 2003.10.3

[21] 申请号 200380102504.4

[30] 优先权

[32] 2002.10.30 [33] US [31] 10/284,219

[86] 国际申请 PCT/IB2003/004465 2003.10.3

[87] 国际公布 WO2004/040905 英 2004.5.13

[85] 进入国家阶段日期 2005.4.29

[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 N·迪米特罗瓦 R·S·贾辛斯奇

[74] 专利代理机构 中国专利代理(香港)有限公司

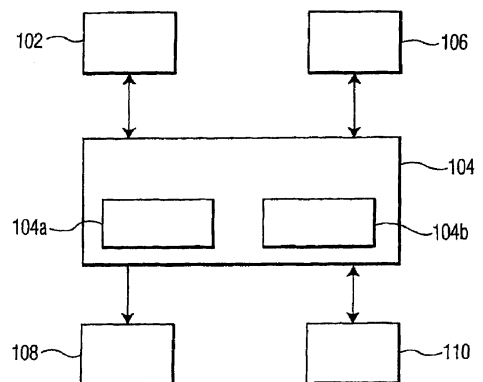
代理人 程天正 王忠忠

权利要求书 4 页 说明书 7 页 附图 2 页

[54] 发明名称 用于编辑源视频的方法和设备

[57] 摘要

用于对已经拍摄好的源视频进行编辑以稳定该视频信号中的图像的方法和设备。为了从视频中消除跳跃运动，首先要检测镜头的变化。然后，对该镜头中的视频内的任何跳动进行分类，并且根据这一分类将视频进一步分段成较小的段。消除选定段中的跳动。然后将经过修正的包括多个帧的镜头加到在前的镜头上，直到视频的所有镜头都已经针对跳动进行了适当的修正。为了帮助用户识别所要编辑的镜头，对这些镜头的关键帧或瞬态图像进行显示，从而使得用户能够判定镜头的处理是否是期望的和应当将哪些镜头加到最终的视频中。



1. 一种编辑源视频信号的方法，包括：
接收源视频信号；
通过数据处理器对源视频进行分析，以识别源视频信号中的镜头，源视频信号中的镜头被界定为在源视频信号的帧实质上不同于其紧跟的在前帧时的两个实例之间；
通过数据处理器对镜头进行分析，以识别该镜头期间的任何跳跃图像运动；
通过数据处理器将所识别出的跳跃运动分类为多种类型的跳跃运动之一，一种类型是没有跳跃运动；和
考虑所分类的跳跃运动类型对经过分类的镜头进行稳定，这种稳定足以产生经过修正的镜头输出视频信号，在所述镜头输出视频信号中，从源视频信号中的镜头中基本上消除了跳跃运动。
2. 按照权利要求 1 所述的方法，还包括，在对经过分类的镜头进行稳定之前：
产生适用于向用户指示对源视频信号中的镜头所分类的跳跃动作类型的指示信号；和
接收重分类信号，该重分类信号指示用户对源视频信号中的镜头进行了跳跃运动类型的重新分类；
其中考虑对镜头重新分类的跳跃运动类型而对镜头进行稳定。
3. 按照权利要求 2 所述的方法，还包括：
选择镜头中的帧；
产生适用于显示器的关键帧信号，以在其上显示帧。
4. 按照权利要求 1 所述的方法，还包括将经过修正的镜头输出视频信号加到另一个视频信号上。
5. 按照权利要求 2 所述的方法，还包括将经过修正的镜头输出视频信号加到另一个视频信号上。
6. 按照权利要求 3 所述的方法，还包括将经过修正的镜头输出视频信号加到另一个视频信号上。
7. 按照权利要求 5 所述的方法，其中所述另一个视频信号是另一个经过修正的镜头输出视频信号。
8. 按照权利要求 6 所述的方法，其中所述另一个视频信号是另

一个经过修正的镜头输出视频信号。

9. 按照权利要求 7 所述的方法, 其中, 按照与源视频信号中出现的镜头相同的顺序, 将经过修正的镜头输出视频信号相互加在一起。

5 10. 按照权利要求 8 所述的方法, 其中, 按照与源视频信号中出现的镜头相同的顺序, 将经过修正的镜头输出视频信号相互加在一起。

10 11. 按照权利要求 7 所述的方法, 其中, 按照与源视频信号中出现的镜头不同的顺序, 将经过修正的镜头输出视频信号相互加在一起。

12. 按照权利要求 8 所述的方法, 其中, 按照与源视频信号中出现的镜头不同的顺序, 将经过修正的镜头输出视频信号相互加在一起。

15 13. 按照权利要求 8 所述的方法, 其中, 通过用户在显示器上将经过修正的镜头的帧拖拉成期望的顺序, 将经过修正的镜头输出视频信号按照期望的顺序相互加在一起。

14. 按照权利要求 3 所述的方法, 还包括采用关键帧信号在显示器上显示帧, 其中显示器使用指示信号来显示对源视频信号的部分进行分类的跳跃运动的类型。

20 15. 按照权利要求 1 所述的方法, 还包括, 在对经过分类的镜头进行稳定之前, 将该镜头分成多个段, 每个段具有单独一种识别出的跳跃运动的类型; 并且

其中所述对经过分类的镜头进行稳定的步骤包括对各个段进行稳定。

25 16. 按照权利要求 1 所述的方法, 还包括, 在对经过分类的镜头进行稳定之前, 将该镜头分成多个段, 每个段具有小于预定阈值持续时间的持续时间; 并且

其中所述对经过分类的镜头进行稳定的步骤包括对各个段进行稳定。

30 17. 按照权利要求 15 所述的方法, 还包括, 在对经过分类的镜头进行稳定之前, 将该镜头分成多个段, 每个段具有小于预定阈值持续时间的持续时间; 并且

其中所述对经过分类的镜头进行稳定的步骤包括对各个段进行稳定。

18. 按照权利要求 8 所述的方法, 还包括将加好的经过修正的镜头输出视频信号存储到存储介质上。

5 19. 一种用于编辑源视频信号的设备, 包括:

用于接收源视频信号的接收装置;

处理器, 用于对源视频信号进行分析, 以识别源视频信号中的镜头, 源视频信号中的镜头被界定为在源视频信号的帧实质上不同于其紧跟的在前帧时的两个实例之间, 用于对镜头进行分析, 以识别该镜头期间的任何跳跃图像运动, 并且用于将所识别出的跳跃运动分类为多种类型的跳跃运动之一, 一种类型是没有跳跃运动, 该处理器产生指示所分类的跳跃运动的类型的信号; 和

10 视频稳定器, 该视频稳定器接收经过分类的镜头和指示所分类的跳跃运动类型的信号, 该视频稳定器考虑所分类的跳跃运动类型、采用来自处理器的信号对经过分类的镜头进行稳定, 这种稳定足以产生经过修正的镜头输出视频信号, 在所述镜头输出视频信号中, 基本上消除了来自源视频信号中的镜头中的跳跃运动。

20. 按照权利要求 19 所述的设备, 还包括:

15 产生装置, 该产生装置从处理器接收指示所分类的跳跃运动类型的信号并且对其进行响应产生适用于向用户指示对源视频信号中的镜头所分类的跳跃运动的类型的指示信号; 和

接收装置, 用于接收表示用户对源视频信号中的镜头进行了跳跃运动类型的再分类的再分类信号;

25 其中视频稳定器接收再分类信号, 并且其中考虑了对镜头进行跳跃运动再分类的类型来对镜头进行稳定。

21. 按照权利要求 20 所述的设备, 还包括关键帧选取装置, 用于在所接收到的源视频信号中的镜头中选取一个帧, 并且用于产生适用于显示器的关键帧信号, 以在其上显示该帧。

30 22. 按照权利要求 19 所述的设备, 还包括用于将镜头分成多个段的装置, 每个段具有单独一种识别得到的跳跃运动类型; 和

其中所述视频稳定器通过对各个段进行稳定来稳定经过分类的镜头。

23. 按照权利要求 19 所述的设备,还包括用于将镜头分成多个段的装置,每个段具有小于预定阈值持续时间的持续时间;和
其中所述视频稳定器通过对各个段进行稳定来稳定经过分类的镜头。
- 5 24. 按照权利要求 22 所述的设备,还包括用于将镜头分成多个段的装置,每个段具有小于预定阈值持续时间的持续时间;和
其中所述视频稳定器通过对各个段进行稳定来稳定经过分类的镜头。
- 10 25. 按照权利要求 21 所述的设备,还包括显示器,该显示器接收关键帧信号并且适用于在其上显示该帧。
26. 按照权利要求 19 所述的设备,还包括用于使得用户能够以期望的顺序添加经过修正的镜头输出视频信号的装置。
27. 按照权利要求 26 所述的设备,还包括用于接收和存储所添加的经过修正的镜头输出视频信号的装置。

用于编辑源视频的方法和设备

技术领域

5 本发明涉及一种用于编辑源视频的方法和设备。

背景技术

人们目前使用视频编辑器程序来编辑已有的源视频，以创作个性化的视频，例如，仅包含一个或多个视频中用户感兴趣的部份的家庭影片。视频编辑器程序将用户选定的帧或图像部分与其它的帧或图像连接起来，以创作经过编辑的或个性化视频。

虽然视频编辑处理为用户提供了编辑已有视频和由此创作新的视频的灵活性，但是这种处理一般来说很耗费时间并且使人疲惫。在视频上采集的图像通常会出现抖动和跳动，因为在最初拍摄视频时摄像机处于运动状态。这种摄像机运动造成视频帧中的图像表现出不平滑或跳跃运动，并且在放大所拍摄的图像时表现出不平滑的变焦。

当前的视频器材，例如，摄录像机和个人录像机能够在最初拍摄视频时在一定程度上对图像进行稳定，并且在拍摄视频的时候将目标定位在镜头（shot）之内。不过，这些特征目前并没有在对已经拍摄好的视频进行操作的视频编辑器程序中得到实现。这样，如果视频最初是使用不具有从所记录的视频中消除摄像机抖动的能力或者在拍摄视频的时候实现无抖动变焦的能力的录像机拍摄的，那么随后将很难从视频中消除这些瑕疵。

发明内容

本发明致力于一种用于对已经拍摄好的源视频进行编辑、以稳定视频中的图像的方法和设备。按照本发明，要从视频中消除跳跃运动，首先要检测镜头的变化。然后，对该镜头的视频中的任何跳动进行分类。如果该镜头具有不止一种跳动的分类，如果可以的话，则将该镜头进一步分段成较小的段，以致每个段具有单独一种跳动的类型。如果任何一段都长于预定的阈值长度，则将该段分解成更短的段。然后消除各段中的跳动。然后将经过修正的每个都包括多个帧的段加到在

前的段上，直到对视频的所有段都已经针对跳动进行了适当的修正。为了帮助用户识别所编辑的段，对这些段的关键帧或瞬态图像进行显示，从而使得用户能够判定段的处理是否是期望的和应当将哪些段加到最终的视频中。

5 本发明视频处理系统包括视频源、智能视频编辑器、输入装置、显示器和存储装置。视频源提供所要编辑的视频。输入装置接收来自用户的命令，并且将它们传送给智能视频编辑器。智能视频编辑器从视频源接收所要编辑的视频，并且依照存储在智能视频编辑器的存储器中的视频处理程序和依照通过输入装置从用户那里接收到的指令对
10 输入视频进行处理。将来自智能视频编辑器的电子视频信号传送给显示器，以显示给用户。存储装置将经过编辑之后的视频存储在存储介质中。

通过下述结合附图考虑的详细说明，本发明的其它目的和特征将会变得显而易见。不过，应当理解，附图是为了图解说明的目的而单独设计的，并不作限制本发明之用，对本发明的限定应当参照所附的
15 权利要求书。此外还应理解，并没有必要按比例画出附图，除非另有说明，它们仅仅用于从概念上解释说明本文所介绍的结构和流程。

附图说明

20 在附图中，相同的附图标记表示相同的元件：

附图 1 是本发明的视频处理系统的框图；和

附图 2 是使用本发明的视频处理系统编辑视频的一种实施方式的流程图。

具体实施方式

25 附图 1 表示本发明的视频处理系统，该系统包括视频源 102、智能视频编辑器 104、输入装置 106、显示器 108 和存储装置 110。视频源 102 提供所要编辑的视频，并且该视频源 102 可以是多种源设备中的任何一种，例如，摄录像机、VCR、DVD 播放器、个人计算机、计算机网
30 络、因特网或者手持装置，比如与网络和与视频处理系统无线连接的移动电话。

输入装置 106 接收来自用户的命令并且将它们传送到智能视频编

辑器 104。该输入装置 106 可以是，例如，遥控器、键盘、鼠标、个人数字助理或者移动电话。输入装置 106 与智能视频编辑器 104 或者是硬连线连接的，或者是无线连接的。智能视频编辑器 104 从视频源 102 接收所要编辑的视频。智能视频编辑器 104 的处理器 104a 依照存储在存储器 104b 中的视频处理程序并且依照通过输入装置 106 从用户那里接收到的指令对输入视频进行处理。存储器 104b 还可用于暂时存储输入视频信号，直到编辑完成，以下载到存储装置 110 中。虽然所示出的处理器 104a 仅包含一个处理器和一个存储器，但是按照另外一种可选方案，它可以由多个各自具有自己功能的装置或模块构成。智能视频编辑器 104 可以位于接近用户的地方，该智能视频编辑器 104 为专用硬件装置（比如机顶盒）、为可插接在其它用户装置上的模块，或者为运行在数据处理装置或计算机上的软件。按照另外一种可选方案，可以将智能视频编辑器 104 远离用户地定位在可通过电话、电缆、移动电话或卫星连接或者可通过因特网接达用户的远程服务器上。从智能视频编辑器 104 送出的电子视频信号被传送到显示器 108，以显示给用户。显示器 108 可以是计算机监视器、电视机或者任何其它类型的具有显示器的装置，比如摄录像机、移动电话、个人数字助理等。存储装置 110 将经过编辑之后的视频存储到存储介质中。存储装置 110 可以是，例如，将经过编辑的视频存储到录像带上的 VCR、计算机硬盘驱动器（在本地位于接近智能视频编辑器 104 和显示器 108 的位置，或者在远程位置上位于与视频编辑器 104 相连的服务器上）、将经过编辑的视频存储到 DVD 或 CD 上的 DVD 或 CD 写入器或者任何其它类型的数据存储装置。

按照本发明，为了从视频中消除跳跃动作，首先要检测视频中的渐进的和突发的变化或转换，即，识别各个独立的镜头。在文献中，这样的突发转换一般称为“切换”。视频中的渐进镜头变化有这样几种类型：渐隐渐显、淡入淡出、翻转或者任何其它类型的特效。然后，检测该镜头的视频中的任何跳动。如果一个镜头中有超过一种类型的跳动，则如果可以的话，将该镜头分段，使得各段具有一种类型的跳动。如果任何一段特别长，就是说，其长度大于预定的或用户选定的阈值长度（比如，30 秒钟），则根据需要对该段进行进一步分段。然后依据跳动的类型对这些段进行分类。长镜头会频繁出现在个人家庭

录像中，一般来说可能持续好多分钟。例如，在生日聚会上，摄像机可能会静止相对较长的时间段，然后用户在四处走动的同时进行拍摄。在这种实例情况下，会有两段：静止段和晃动段（即，在用户走动的时候进行的拍摄）。然后使用公知的图像稳定算法（下面将做讨论）对选定段进行跳跃动作消除。然后将经过修正的包括多个帧的段加到在前段上，直到针对跳动对整个视频都进行了适当的修正。

附图 2 表示使用本发明的智能视频编辑器 104 进行的视频编辑的一种实施方式的流程图。当智能视频编辑器 104 最初接收到视频时，最开始是对各个独立的镜头进行识别，步骤 200。当视频的后续帧明显且基本上不同于其紧跟的在前帧的时候，表明出现了镜头变化或切换。例如，如果起初视频显示了很多帧的风景，然后变成了一张脸的特写肖像图像，则应当将该视频看作经历了一次“切换”。不过，如果视频起初显示风景并且推进到脸部肖像图像，则该视频图像应当没有切换，因为每个帧都应该是类似于其紧跟的在前帧的。切换的确定是通过将视频的一个帧与该视频的下一个后续帧进行比较来实现的。如果一个帧的某些特征与接下来的后续帧的类似特征相类似（例如，如果有两个帧的相同部分具有相同颜色的情况大量存在），则没有查找到切换。此外，可以使用其它的特征（例如，纹理、形状、边缘）或对象（例如，面孔、视频文本、车辆的存在等）来比较两个帧。不过，如果一个帧的某些特征没有出现在视频的后续帧中，则存在着切换。在美国专利 US6137544 中公开了这样的用于确定切换的视频分析，该美国专利的全部内容以引用的方式并入本文。该专利公开了使用离散余弦变换进行分解来得出帧的表示，以便能够实现有效比较。在对视频帧进行分析以识别出视频中的切换或找出视频中的切换的位置的时候，可以将视频保存在视频编辑器 104 的存储器 104b 内的缓冲区中，以进行随后的编辑。按照另一种可选方案，不需要将同一镜头中的帧保存在存储器中，而是可以给它们加上标记，以表明它们全都是同一镜头的部分。

一旦识别到了视频中的开头和结尾（将一个镜头定义为两次切换之间的视频部分），然后就对该镜头进行分析，步骤 202，以识别出检测到了哪种跳跃动作，如果有的话。视频的帧中的图像跳跃动作是，例如，在摄像机的用户在最初记录场景时抖动或移动摄像机的时候造

成的。跳跃动作包括不协调调焦、抖动摇摄和抖动的静止场景。不协调调焦是在摄像机推进和拉远对象但是却是以不均匀的速度进行的推进和拉远的时候出现的。抖动摇摄是在摄像机横越风景进行摇摄（例如）但是在摄像机摇动期间有抖动的时候出现的。抖动静止场景是在摄像机以一个对象为中心时摄像机抖动的时候出现的。视频场景中的这些不同类型的跳跃运动是通过比较场景中的连续帧来确定的。

在文献中提出了很多用于图像稳定的算法。例如，在 Carlos Morimoto 等人所著的《快速电子数字图像稳定 (Fast Electronic Digital Image Stabilization)》(IEEE 关于模式识别的国际会议 (IEEE International Conference on Pattern Recognition)，奥地利维也纳，1996 年，第 1 - 28 页) 中介绍的方法中，对一小组特征进行跟踪，并且估算连续帧之间的摄像机运动。稳定是通过结合来自参考帧的所有动作并且将当前帧掩蔽回到参考帧来实现的。在 Zhigang Zhu 等人所著的《在 2.5D 运动估测和内部运动过滤的基础上进行的摄像机稳定 (Camera Stabilization on 2.5D Motion Estimation and Internal Motion Filtering)》(IEEE 关于智能交通工具的国际会议 (IEEE International Conference on Intelligent Vehicles)，德国斯图加特，1998 年 10 月 28 日 - 30 日，第 329 - 334 页) 中公开了另一种图像稳定方法。按照这种方法，使用 2.5D 运动模型，首先通过基于锥形的匹配算法估测典型点的图像速度，该基于金字塔的匹配算法比较两个连续的图像并且针对图像帧中的小块得出图像的速度。然后根据图像速度，确定摄像机的运动类型（例如，移动摄像、水平跟踪、垂直跟踪）。然后使用一般的惯性模型，消除掉不需要的高频振动，以实现图像稳定。经过平稳的运动参数和原始的估测运动参数之间的差可用于补偿图像，以便获得稳定的图像序列。

如果一个镜头具有多于一种类型的跳动，则如果可以的话，对该镜头进行分段，使得每个段都具有一种类型的跳动，步骤 204。如果任一个段特别长，就是说，其长度大于预定的和用户选定的阈值长度（比如，30 秒钟），则可以根据需要对该段进一步分段，步骤 206。

然后对检测到跳跃运动的各个镜头进行分析，并且还有可能对那些没有检测到跳跃运动的镜头进行分析，以产生该段的关键帧。关键帧是镜头中代表该镜头中的视频的视频帧。这个关键帧可以是镜头中的第一个帧或者被认为更加具有该镜头的说明价值的另一个选定帧。

在美国专利 No. 6125229 中公开了这样的关键帧选取。关键帧给出了视频在某一时间点的瞬态图像。

对检测到跳跃运动的镜头的各段的开头和结尾进行识别，比如通过识别第一个和最后一个帧或者通过识别段开始时和段结束时在镜头中的时间，步骤 208。

然后依据针对特定段检测到的跳跃运动的类型为这些段加上标签，步骤 210。适用的标签包括对应于“抖动摇摄”的“SP”、对应于“不均匀推进”的“UZI”、对应于“不均匀拉远”的“UZO”和对应于“抖动静止镜头”的“SSS”。按照另外一种可选方案，数据库表中的清单可以包括连同段中检测到的跳跃运动的类型标志在一起的、该段的起始和结束时间。

然后将识别出带有跳跃运动的段的各个帧（或关键帧）在显示器 108 上显示给用户，同时将关于所检测到的跳跃运动的类型的识别结果显示在显示器 108 上，步骤 212。这一识别结果可以具有适当的字母/数字标签（比如，SP、UZI、UZO 和 SSS），或者具有颜色码。按照另外一种可选方案，所有段（有和没有检测到的跳跃运动）的关键帧都可以被显示给用户，从而用户可以选择（如果有的话）应当对该段进行哪种修正。

然后用户可以同意跳跃运动的判断结果、选择不同的跳跃运动选项或者选取不进行任何跳跃运动修正，步骤 214。然后假如用户发出了这样的指令，对段进行适当修正，步骤 216，以充分消除掉跳跃运动。然后将经过修正（如果要求这样）的段加到视频的前一部分上，步骤 218。按照另外一种可选方案，用户可以选择应当合成哪些切换并且按照什么样的顺序来利用输入装置 106 选择合乎要求显示的关键帧，并且利用关键帧作为相应段的代表来使用拖放操作进行非线性操作。这是通过使用可视化工具来显示关键帧而实现的。然后将经过重编的视频输出给显示器 108 和/或存储装置 110。附图 2 中的处理过程持续进行，直到输入视频结束或者直到用户命令该处理过程终止。按照另外一种可选方案，可以将经过修正的视频段保存在存储器 104b 或存储装置 110 中，以备之后使用，或者可以将其加到完全不同的视频上。

随着视频的前进，可以每次将每个镜头的各段的关键帧之一显示给用户，或者可以在同一时间将每个镜头的所有关键帧显示给用户，

以使用户对镜头有整体的观察。按照另外一种方案，可以将不止一个镜头的关键帧同时显示给用户。

5 虽然，在上面介绍的方法中，最初是在镜头变化处对视频进行分解，步骤 200，但是按照另外一种可选方案，可以以预定的时间间隔或在其它的点上对视频进行分解。按照另外一种实施方式，根本不需要在一开始就对视频进行分解，而是用所执行的跳跃动作检测（步骤 202）作为最初的步骤。

10 这样，虽然按照应用于本发明的优选实施方式的方式对本发明进行了展示和介绍并且指出了本发明的基本新颖特征，但是应当理解，在不超出本发明的思想的前提下，本领域的技术人员可以在形式和细节上对所介绍和图示的装置及其操作以及所介绍的方法进行各种省略和替换以及改变。例如，明确地将那些以基本相同的方式执行基本相同的功
15 能来实现相同的效果的元件和/或方法步骤的所有组合预定包含在本发明的范围之内。从一种所介绍的实施方式到另一种的元件替换也已经被充分预期和考虑。因此，我们期望，仅仅由本文所附的权利要求书的范围对本发明的范围加以限定。

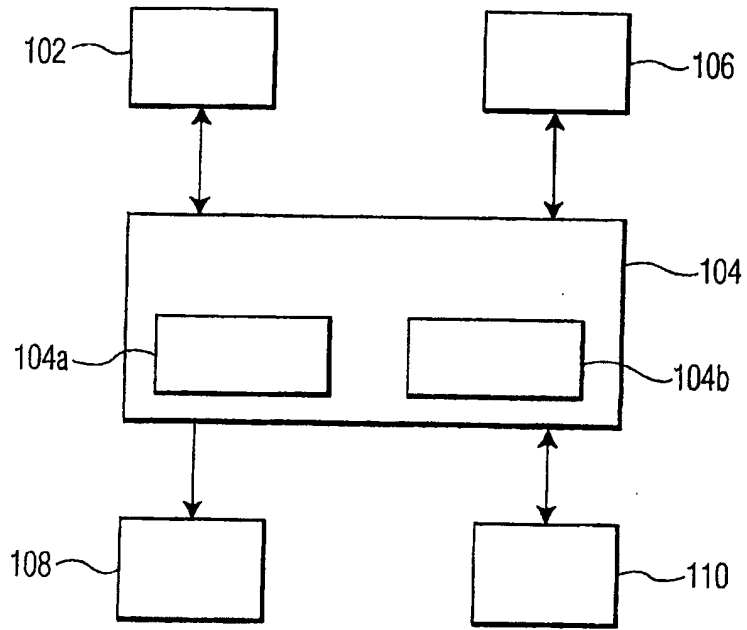


图 1

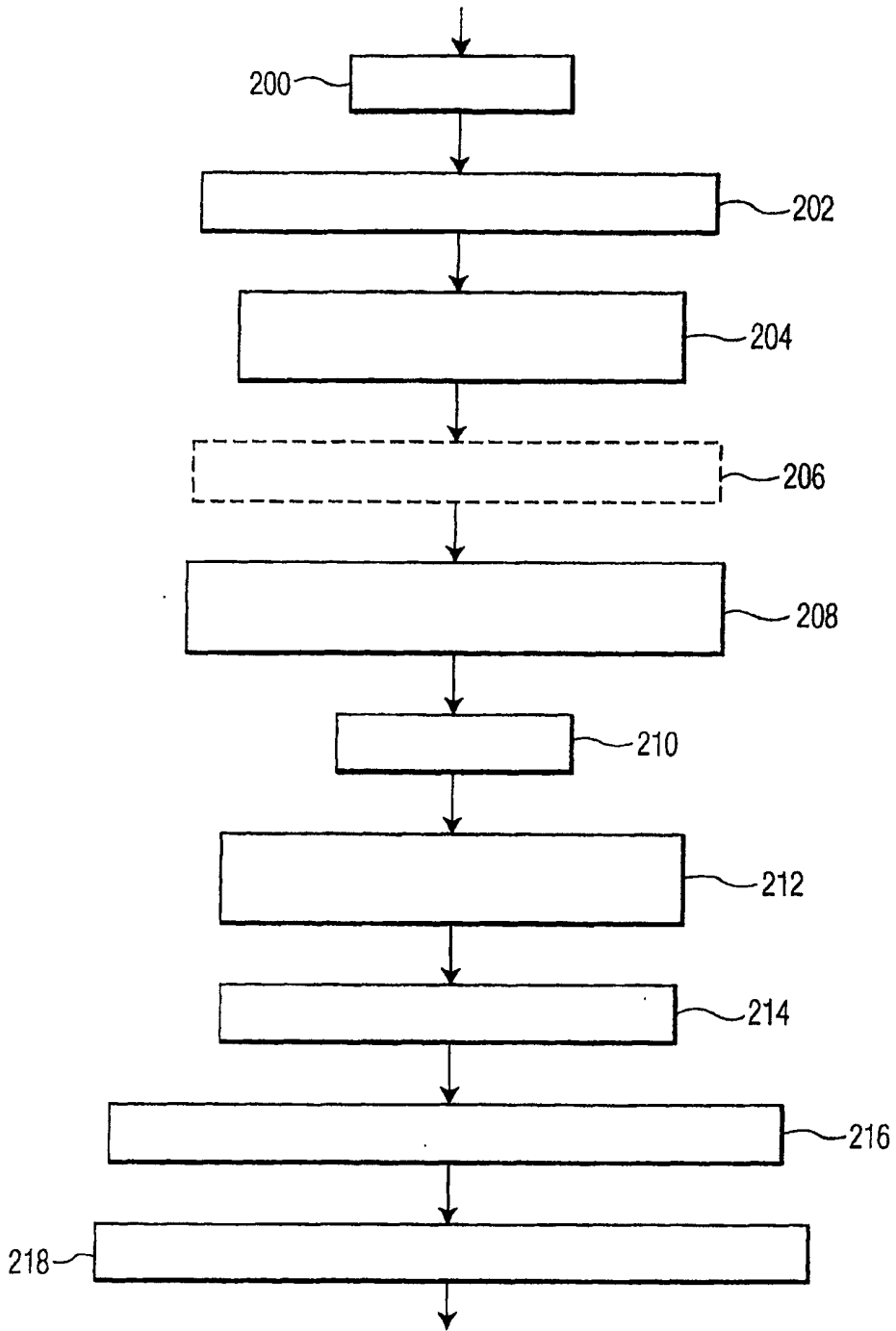


图 2