



(12) 发明专利

(10) 授权公告号 CN 101473653 B

(45) 授权公告日 2011. 11. 16

(21) 申请号 200780022364. 8

(22) 申请日 2007. 06. 11

(30) 优先权数据

06115337. 5 2006. 06. 13 EP

(85) PCT申请进入国家阶段日

2008. 12. 15

(86) PCT申请的申请数据

PCT/IB2007/052187 2007. 06. 11

(87) PCT申请的公布数据

W02007/144813 EN 2007. 12. 21

(73) 专利权人 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 W·F·J·胡根斯特拉滕

C·W·奎斯索特

(74) 专利代理机构 中国专利代理(香港)有限公司
72001

代理人 康正德 谭祐祥

(51) Int. Cl.

H04N 7/24(2006. 01)

(56) 对比文件

WO 2005/069640 A1, 2005. 07. 28, 全文.

WO 2005/069637 A1, 2005. 07. 28, 全文.

CN 1765133 A, 2006. 04. 26, 全文.

CN 1682533 A, 2005. 10. 12, 全文.

CN 1672211 A, 2005. 09. 21, 全文.

GAUCH J M ET AL. IDENTIFICATION OF NEW COMMERCIALS USING REPEATED VIDEO SEQUENCE DETECTION. 《IMAGE PROCESSING 2005. ICIP 2005. IEEE INTERNATIONAL CONFERENCE ON GENOVA ITALY identification of new commercials using repeated video sequence detection》. 2005, P1252-1253.

NATUREL X ET AL. A Fast Shot Matching Strategy for Detecting Duplicate Sequences in a Television Stream. 《PROCEEDINGS OF THE INTERNATIONAL WORKSHOP ON COMPUTER VISION MEETS DATABASES》. 2005, P21-27.

审查员 张伟

权利要求书 2 页 说明书 12 页 附图 3 页

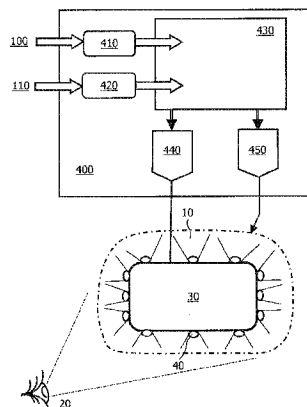
(54) 发明名称

用于识别并同步视频的指纹、器件、方法

(57) 摘要

提供了一种用于包括一系列场景 (200a, 200b, 200c) 的视频流 (110) 的指纹 (340), 其中该指纹 (340) 包括表示场景 (200a, 200b, 200c) 的持续时间的信息。所述表示场景 (200a, 200b, 200c) 的持续时间的信息以所述场景 (200a, 200b, 200c) 包括在所述视频流 (110) 中的顺序包括在所述指纹 (340) 中。而且, 每个持续时间被可选地定义为其中多个视频帧 (260, 270) 的计数。所述方法能够生成指纹, 该指纹易于更简单地和更有效地使第一和第二数据信号一起同步。

CN 101473653 B



1. 一种可被操作来生成指纹 (340) 的器件 (400), 所述器件 (400) 包括:
 - (a) 数据输入装置, 其用于接收视频流 (110);
 - (b) 数据处理器 (430), 其用于分析经由数据输入装置输送的视频流 (110) 以用于检测其中场景变化的发生, 并且用于根据所检测到的场景变化的发生来计算包括在视频流 (110) 中的场景的持续时间, 以及用于根据所述持续时间来生成所述指纹 (340)。
2. 如权利要求 1 所述的器件 (400), 其中所述数据处理器 (430) 可被操作以通过基于下列一个或多个情况的分析来检测视频流 (110) 中的场景变化:
 - (a) 在早期场景 (200a) 的至少最后一幅图像 (270) 相对于随后场景 (200b) 的至少第一幅图像 (310) 之间的视频材料照明度的突然改变;
 - (b) 在早期场景 (200a) 的所述至少最后一幅图像 (270) 相对于随后场景 (200b) 的所述至少第一幅图像 (310) 之间在空间图像特征信息方面的突然改变; 和
 - (c) 早期场景 (200a) 的所述至少最后一幅图像 (270) 的相关音频信息相对于随后场景 (200b) 的所述至少第一幅图像 (310) 的突然改变。
3. 一种根据视频流 (110) 生成指纹 (340) 的方法, 其中所述方法包括以下步骤:
 - (a) 接收所述视频流 (110); 和
 - (b) 检测所述视频流 (110) 中场景变化的发生;
 - (c) 根据所检测到的场景变化的发生来计算包括在所述视频流 (110) 中的场景的持续时间; 和
 - (d) 根据所述持续时间来生成所述指纹 (340)。
4. 如权利要求 3 所述的方法, 其中步骤 (b) 包括通过基于下列一个或多个情况的分析来检测视频流 (110) 中的场景变化:
 - (a) 在早期场景 (200a) 的至少最后一幅图像 (270) 相对于随后场景 (200b) 的至少第一幅图像 (310) 之间的视频材料照明度的突然改变;
 - (b) 在早期场景 (200a) 的所述至少最后一幅图像 (270) 相对于随后场景 (200b) 的所述至少第一幅图像 (310) 之间在空间图像特征信息方面的突然改变; 和
 - (c) 早期场景 (200a) 的所述至少最后一幅图像 (270) 的相关音频信息相对于随后场景 (200b) 的所述至少第一幅图像 (310) 的突然改变。
5. 一种在器件 (400) 中通过将场景变化检测应用于视频流 (110) 来同步辅助媒体 (100) 和所述视频流 (110) 的方法, 所述方法包括以下步骤:
 - (a) 在所述器件 (400) 的数据输入装置 (410, 420) 上接收辅助媒体 (100) 和视频流 (110);
 - (b) 在所述器件 (400) 的数据处理器 (430) 中, 分析经由所述数据输入装置 (410, 420) 输送的视频流 (110) 以用于检测其中场景变化的发生, 根据所检测到的场景变化的发生来计算包括在所述视频流 (110) 中的场景的持续时间, 并根据所述持续时间来生成标识符 (340); 和
 - (c) 在所述器件 (400) 中, 基于与对应于视频流 (110) 中场景变化的标识符 (340) 相关联的辅助媒体 (100), 使视频流 (110) 和辅助媒体 (110) 对用户 (20) 的呈现同步。
6. 如权利要求 5 所述的方法, 其中所述标识符与所述辅助媒体 (100) 之间的关联包括在所述辅助媒体 (100) 中。

7. 如权利要求 6 所述的方法,其中所述标识符(340)包括在所述辅助媒体(100)中。
8. 如权利要求 5 或 6 所述的方法,其中所述辅助媒体包括流光溢彩脚本(100),当所述方法应用在器件(400)中时可被操作来同步流光溢彩脚本(100)与视频流(110)。
9. 包括一系列场景(200a, 200b, 200c)的视频流(110)的指纹(340)的用途,其中所述指纹(340)包括表示所述场景(200a, 200b, 200c)的持续时间的信息,并且所述用途包括将从所述视频流的一部分获得的部分指纹与所述指纹的一部分(340a, 340b)进行比较。
10. 如权利要求 9 所述的指纹(340)的用途,其中通过将所述视频流的所述部分(200a, 200b)获得的部分指纹与所述指纹的所述部分(340a, 340b)进行比较来将所述指纹用于识别视频流(110)。
11. 如权利要求 9 或 10 所述的指纹(340)的用途,其中通过将所述视频流的所述部分获得的部分指纹与所述指纹的所述部分(340a, 340b)进行比较来将所述指纹用于确定所述视频流的所述部分的时间位置。

用于识别并同步视频的指纹、器件、方法

技术领域

[0001] 本发明涉及用于包括一系列场景的视频流的指纹,并且涉及其一般应用以及在可被操作以使得辅助媒体与视频流同步的器件中的应用;例如,本发明关注用于同步流光溢彩脚本与视频流的器件中的问题。而且,本发明还涉及使用指纹来识别所述视频流和使得辅助媒体和视频流同步的方法,例如涉及上述流光溢彩。此外,本发明涉及可在计算硬件上执行来操作以实现这种方法的软件。

背景技术

[0002] “流光溢彩”是对电视和类似图像显示设备的增强。参照图 1,流光溢彩涉及在工作中呈现给用户 20 的光边晕 10,其环绕着电视显示屏 30 或类似显示设备的至少部分周围区域;显示屏 30 可以基于阴极射线管技术、等离子体像素技术、液晶显示器 (LCD) 技术、发光二极管 (LED) 技术和 / 或有机发光二极管 (OLED) 技术 (仅作为一些示例提出)。流光溢彩开发了这样的特性:为用户 20 的眼睛提供进化,或者在用户视野的中心区域提供最大空间分辨率和在用户视野的周围区域提供更加发散的空间分辨率。因此,为了提供给用户 20 增强的视觉体验,期望为用户 20 呈现宽阔的显示图像,在其上的中心区域需要包括最多的空间的和颜色的信息。由于数据经济和显示器制造经济,流光溢彩易于为用户 20 提供宽阔的显示图像,在所述宽阔显示图像的最易于为用户眼睛所接受的中心区域呈现最多的空间和颜色信息。上述边晕 10 可以从各种周围的光源 (例如从光源 40) 产生,所述周围光源的颜色和光输出响应于在相关电视显示屏 30 或类似设备上呈现的图像而变化。

[0003] 在流光溢彩的更基本的实现方法中,用于驱动周围光源的信息可以从电视显示屏 30 上呈现的图像获得。这种方法使得从现存的视频材料呈现流光溢彩操作变得可能,例如从现存的录像带、DVD 和经通信网络 (例如经互联网或经卫星链路) 下载的视频材料。操作中涉及的从现存的视频材料中获得流光溢彩是所述边晕应当不仅正确地补偿在用户 20 的电视 30 上呈现的基于颜色的视频材料,而且随时基本准确地跟踪在电视 30 上呈现的图像颜色的一般改变。同时,对在电视显示屏 30 上的图像进行分析,以用于空间颜色内容分析和根据所述空间颜色内容分析结果获得的对周围光 40 的驱动。这种分析可在数字硬件或计算硬件中实现,所述计算硬件可被操作以执行导致所述计算硬件执行这种分析的软件。

[0004] 作为上面阐述的已知流光溢彩的进展,期望为用户提供带有流光溢彩脚本的数据流或数据文件。即,不是根据呈现在电视显示屏 30 上的视频图像来实时获得对周围光 40 的驱动,而是期望提供流光溢彩脚本以伴随现存视频材料以及将来的视频材料,例如作为对用户 30 的附加服务。应当进一步理解,易于将这种流光溢彩脚本经几个可能的可替代通道提供给用户,例如作为经互联网下载的文件、作为分离的 CD 或 DVD 或相似类型的数据载体、作为通过无线或卫星链路传输的数据流。这些可替代通道在当代被称为“多媒体”。

[0005] 参照图 2,由于在已知流光溢彩中的这种进展所引起的被本发明至少部分解决的技术问题是将所接收的流光溢彩脚本 100 同步到相应的视频材料 110 的方法,所述视频材料 110 已经被用户 20 的设备接收、或被存储在数据库 120 中并且潜在地可

被用户 20 访问。当视频材料 110 基本上缺乏任何同步标志 130 (或将实现流光溢彩脚本 100 与视频材料 110 时间同步从而使得所述脚本 100 和视频材料能够同步操作地被输出到电视 30 和流光溢彩光源 40 的类似物), 进一步加剧了该技术问题。

[0006] 一起同步两个或多个信号的方法是已知的。例如在公开的国际 PCT 专利申请 PCT/IB2004/051259 (W02005/011281) 中, 描述了一种生成用于同步至少两个信号的第一指纹和第二指纹的设备和方法。在该方法中, 在第一信号的片段和第二信号的片段的基础上生成指纹对; 换句话说, 不是对第一信号和第二信号的全部而仅仅是对其一部分来生成指纹。例如, 第一信号是音频信号而第二信号是视频信号, 其中当该音频信号和视频信号提供给用户消费时, 它们将准确地时间同步。在上述专利申请所描述的方法中, 所生成的关于音频和视频信号的指纹对存储在数据库中并且被传输或分发到同步设备。在使用所述同步设备同步音频和视频信号期间, 音频信号和视频信号的指纹都是由同步设备生成, 将这些指纹与数据库中的指纹相匹配。当发现匹配时, 所述指纹能够用于确定同步时间点, 这些时间点可用于将音频和视频信号同步到一起。而且, 可以在无需修改音频或视频信号情况下获得这种同步。

[0007] 在所公开的 PCT 专利申请中描述的方法存在一个问题: 对于音频以及视频, 必须准备好指纹对并将其存储在数据库中。而且, 在同步期间, 为了指纹匹配而实时提供对数据库的访问是不方便的或昂贵的。因此, 就消费型设备而言, 提供到这种具有存储在其中的预处理指纹对的数据库的连接不总是实用的或可行的。该方法的另外一个缺陷在于其在计算上没有吸引力, 因为所需的计算需要可观数量的资源。该方法的又一个缺陷在于它易受视频格式改变的影响。例如当将包括指纹的视频格式转换为高密度视频格式时, 为标准密度视频生成的指纹将典型地不合格。相似地, 在视频已经被裁剪为 16:9 的帧比率之后, 为 21:9 的帧比率生成的指纹将典型地不合格。

[0008] 因此, 由本发明至少部分解决的技术问题是提供同步两个数据信号的可替代方法, 特别是在上述流光溢彩技术的情况下。

发明内容

[0009] 本发明的目的在于提供一种用于从第一数据信号生成指纹的可替代方法, 和可操作来实现该可替代方法的器件; 所述指纹可选地易于被使用来将第一数据信号同步到第二数据信号, 例如在流光溢彩的情况下用于将流光溢彩脚本和视频内容进行时间同步。

[0010] 根据本发明的第一方面, 提供一种用于包括一系列场景的视频流的指纹, 其中所述指纹包括表示所述场景持续时间的信息。

[0011] 用于视频流的指纹可以考虑对视频流的编码, 因为它提供了基于压缩的信息来识别视频流。然而, 所述指纹自身典型地没有提供足够的用于从它来重构视频流的信息。视频流的指纹的概念同样是已知的。相似地, 如下面详细的描述, 在诸如视频领域, 场景的概念同样是已知的。

[0012] 指纹的基本思想是: 它充当输入串的紧凑表示, 并且可以使用它就像它是唯一可等同于所述串那样。所发明的指纹尤其具有识别视频流或与之同步的目的。这可以通过比较各个视频流的指纹而不用比较视频流来获得。可替代地, 这可以通过将从部分视频流得到的部分指纹与从全部视频流得到的指纹的一部分进行比较来获得。所述指纹进一步能够

确定所述视频流的一片段在完整视频流中的时间位置,所述片段是整个视频流的一部分。

[0013] 所述指纹与现存的指纹相比具有更低的计算复杂性。对于一些视频分发格式如 MPEG,为了优化分配解码器资源,所述解码器已经进行了场景变化检测。因此,在这种解码器中可以用很少的成本实现该指纹。特别地,所述指纹可以(而不是包括)由表示场景持续时间的信息构成。令人惊讶地,即使所述指纹仅仅由所述信息构成,如果它只涉及几个场景变化,这仍然会产生唯一指纹。这是由典型的视频流的场景持续时间的巨大可变性造成的,特别是如果用高达单帧的分辨率来进行测量更是如此。

[0014] 该指纹具有对于改变视频格式或转换视频格式编码的鲁棒性,因为这种操作典型地不会影响所述场景的持续时间。

[0015] 该指纹实现了一种用于识别任何包括一系列场景的视频流的相对高效的方法。它还能够识别仅仅在(通过例如跳入所播送的电影的中部)视频流开始之后的某一时间点接收的视频流。一个这样的方法操作如下。首先复位帧计数器。在检测到场景变化之后,读出计数器并随后复位。所读出的值在可选的传输之后被送入匹配设备,该设备确定记录集合的子集。在所述子集中的每个记录具有与所述读出值匹配的相关指纹。如果在指纹中出现所述读出值,则该指纹匹配该读出值。可替代地,如果在指纹中出现其他值而在所读出值与其他值之间的量差不超过阈值,则该指纹可以匹配所述值从而适应较小的误差。所述阈值可以预选确定或取决于所述值,以提供例如对在相对较长的场景中的丢帧的鲁棒性。

[0016] 当视频流在初始集合中没有相关记录时,重复前面的步骤,由此所述匹配设备继续先前创建的子集,直到所得到的子集包括识别视频流的单一记录,或直到所述子集为空。在远程匹配设备的情况下,最终结果可以返回到所述帧被计数的位置。可替代地,如果在一个元组中已经积累了多个帧计数,这所述匹配设备只能开始匹配。这具有以下优势:可以使得搜索记录更加高效,因为例如不需要保留中间子集。

[0017] 识别视频流本身具有几个应用,如链接到关于视频流的元数据,例如标题、演员、音乐(如果所述视频流是电影),或艺术家、唱片、卖点(如果所述视频流是用于歌曲的视频剪辑)。

[0018] 可以通过计数场景的帧来确定场景的持续时间。然而,如本领域所公知,世界上使用了不同的帧速率。因此,为了进一步增加所述指纹对转换视频的帧速率的鲁棒性,所述指纹可以考虑帧速率。这可以通过各种方式实现。

[0019] 在第一方式中,通过计数场景的帧并用所述帧周期乘以所得到的帧计数来确定持续时间,所述帧周期是在两个连续帧之间的固定时间量。以此方式,可以获得场景持续时间。在该第一方式中,所述指纹只需要包括表示持续时间的信息。该方式具有一些优势:它对于当前没有使用的其他帧速率是开放式的。所得到的持续时间可以以预选确定的单位(如秒)表示,并且其可在指纹中被编码为浮点数。有利地,所述持续时间可表示为更小预定单位(如毫秒或微秒)的整数倍数,并且在指纹中编码为二进制整数。

[0020] 在第二方式中,预选确定一个特定帧速率作为默认值。每个指纹包括在所述特定帧速率情况下的持续时间。如果通过计数具有不同于那一个特定帧速率的帧速率来确定场景持续时间,将所得到的帧计数转换为那一个特定帧速率。在实例中,那一个特定帧速率被预先确定为 300Hz。在该实例中,关于 50Hz 帧速率的帧计数被乘以六,同时关于 60Hz 帧速率的帧计数被乘以五。该方式具有下列优势:指纹本身不需要包括附加信息,因为那一个特

定帧速率隐含在指纹中。

[0021] 在第三方式中,除了帧数,指纹额外地包括表示帧速率的信息,以该帧速率对帧进行计数。在一个实例中,指纹的一个附加的单个位表示帧速率,该指纹的帧数适合于该帧速率。如果在该实例中,指纹中的那一个附加单个位被设置为一,则该指纹中的帧数适合于 60Hz 的帧速率,而如果那一个附加单个为被设置为零,则该指纹中的帧数适合于 50Hz 的帧速率。该实例的优势在于,仅仅需要单个位来编码帧速率。其他实例也是可能的,其中多个位用于表示多个预选确定的帧速率中的一个。该第三方式的优势在于,单个的指纹格式支持多个帧速率。

[0022] 在第四方式中,指纹仅仅包括帧数,并且使用启发式方法以确定所述多个帧速率适合的帧速率。在一个实例中,如果视频信号符号 NTSC 标准,则帧速率可被推断为 60Hz,而如果视频信号符号 PAL 标准,则帧速率可被推断为 50Hz。

[0023] 可以应用多个其他方式(例如通过将上述方式的措施组合)来确定基于帧数的场景持续时间。

[0024] 在根据权利要求 1 的指纹的一个应用中,指纹与流光溢彩脚本相关联。装配有流光溢彩“光扬声器”的娱乐设备(如电视机或多媒体中心和 PC)在其屏幕上示出了视频流。如本领域公知的,视频内容可以被电视机分析从而例如使得光扬声器在屏幕上模拟和扩展颜色。

[0025] 所述娱乐设备确定视频流中的场景变化、计算视频流的部分指纹、以及将所述部分指纹与相关于所述流光溢彩脚本的指纹进行匹配。只要发现匹配,所述当前同步的流光溢彩脚本从正被分析的视频内容中取出,以确定所述光扬声器的颜色。

[0026] 当流光溢彩脚本正在控制光扬声器时,所述娱乐设备继续确定视频流中的场景变化,以用于检测同步的损失。如果同步的损失被检测到,所述流光溢彩脚本将停止对光扬声器的控制并且所述娱乐设备回复到分析视频内容以用于控制光扬声器。在同步损失之后,所述娱乐设备仍然继续确定视频流中的场景变化,以用于试图恢复与流光溢彩脚本的同步。

[0027] 几个启发式方法可以应用于检测同步损失并用于恢复同步。

[0028] 在第一启发式方法中,一旦在被同步的指纹没有指定场景变化的瞬间检测到了视频流中的场景变化,则检测到同步损失。这具有的优点在于检测损失的快速响应时间。

[0029] 在第二启发式方法中,一旦在被同步的指纹没有指定任何场景变化的瞬间检测到了视频流中的两个连续的场景变化,则检测到同步损失。这具有的优点在于它可以阻止同步损失的错误检测。

[0030] 在第三启发式方法中,可以基于从视频流或用户操作娱乐设备而获得的另一个信号来检测同步的损失。举一个例子,在某些时刻期望将一块商业广告作为视频流的一部分进行广播。举另外一个例子,用户操纵遥控器切换到另一个频道可能立刻导致对同步损失的检测。这具有的优点在于,同步损失的检测不仅仅取决于检测场景变化。

[0031] 相似地,为了恢复同步,可以使用各种启发式方法。

[0032] 在用于恢复的第一启发式方法中,所述恢复基于同步的先前损失的内容。举一个例子,当娱乐设备存储损失内容时,如果同步损失是由于用户选择不同频道而损失的,则在该用户返回先前频道之后立刻就恢复所述同步。所述内容可以包括频道,连同匹配指纹与

发生损失处的帧数或时间点的关联。举另外一个例子,如果所述损失由商业广告中断造成,则可以在例如 30 秒(该实例中商业广告中断的典型长度)之后试图恢复。此外,所述恢复可以考虑在中断之后重复所述中断之前的一段视频流。这一段视频流典型地持续几秒与小于一分钟(如半分钟)之间。这可以通过校正用于商业广告中断的部分指纹来实现,例如通过从所述部分指纹中切去作为商业广告中断的一部分的场景变化。随后已校正的部分指纹对与流光溢彩脚步相关的全部指纹进行匹配。校正所述部分指纹也需要切去重复的片段。

[0033] 所述指纹也适合于使其他媒体类型与视频流,尤其是字幕同步。

[0034] 本发明的进一步的优点在于,实现了一种生成指纹的方法,该方法易于更简单和更有效地将第一和第二数据信号同步在一起。

[0035] 可选地,表示场景持续时间的信息以场景包括在所述视频流中的顺序包括在指纹中。

[0036] 这种指纹提供了在识别视频流的同时与所述视频流同步,因为只要所述指纹包括足够的、在记录的初始集合上唯一的持续时间,就可以确定在当前帧计数数量与视频流开始之间的偏移。

[0037] 可选地,在所述指纹中,每个持续时间被定义为其中多个视频帧的计数。

[0038] 一些 MPEG 解码器已经执行场景变化检测以用于优化分配解码器资源:仅仅需要有效地加入帧计数器。因此,可以在这种解码器中使用非常少的附加成本实现该指纹。

[0039] 可选地,在所述指纹中,通过对应于以下一个或多个情况的场景变化来定义持续时间:

[0040] (a) 在视频流中,在早期场景的至少最后一幅图像相对于随后场景的至少第一幅图像之间的视频材料照明度的突然改变;

[0041] (b) 在视频流中,在早期场景的所述至少最后一幅图像相对于随后场景的所述至少第一幅图像之间在空间图像特征信息方面的突然改变;和

[0042] (c) 在视频流中,早期场景的所述至少最后一幅图像的相关音频信息相对于随后场景的所述至少第一幅图像的突然改变。

[0043] 根据本发明的第二方面,提供了一种可被操作以生成根据本发明第一方面的指纹的器件,所述器件包括:

[0044] (a) 数据输入装置,用于接收视频流;

[0045] (b) 数据处理器,用于分析经输入装置输送的视频流以用于检测其中场景变化的发生,并且用于根据所检测的场景变化的发生来计算包括在视频流中的场景持续时间,以及用于根据所述持续时间生成指纹。

[0046] 可选地,在所述器件中,所述处理器被操作以通过基于以下情况中的一个或多个的分析来检测视频流中的场景变化:

[0047] (a) 在早期场景的至少最后一幅图像相对于随后场景的至少第一幅图像之间的视频材料照明度的突然改变;

[0048] (b) 在早期场景的所述至少最后一幅图像相对于随后场景的所述至少第一幅图像之间在空间图像特征信息方面的突然改变;和

[0049] (c) 早期场景的所述至少最后一幅图像的相关音频信息相对于随后场景的所述至少第一幅图像的突然改变。

[0050] 根据本发明的第三方面,提供了一种从视频流生成根据本发明第一方面的指纹的方法,其中所述方法包括以下步骤:

[0051] (a) 接收视频流;

[0052] (b) 在视频流中检测场景变化的发生;

[0053] (c) 根据所检测的场景变化的发生来计算包括在视频流中的场景持续时间;和

[0054] (d) 根据持续时间生成指纹。

[0055] 可选地,在所述方法中,步骤(b)包括通过基于以下情况中的一个或多个的分析来检测视频流中的场景变化:

[0056] (a) 在早期场景的至少最后一幅图像相对于随后场景的至少第一幅图像之间的视频材料照明度的突然改变;

[0057] (b) 在早期场景的所述至少最后一幅图像相对于随后场景的所述至少第一幅图像之间在空间图像特征信息方面的突然改变;和

[0058] (c) 早期场景的所述至少最后一幅图像的相关音频信息相对于随后场景的所述至少第一幅图像的突然改变。

[0059] 值得一提的是,也可以使用用于场景变化检测的其他步骤和技术。事实上,只要将相同的检测器既用于创建指纹又用于分析视频流以获得所述部分指纹,所精确检测的内容在某种程度上是不相关的。

[0060] 根据本发明的第四方面,提供了一种在器件中通过将场景变化检测应用于视频流来同步辅助媒体与视频流的方法,所述方法包括以下步骤:

[0061] (a) 在所述器件的数据输入装置上接收辅助媒体和视频流;

[0062] (b) 在所述器件的数据处理器中分析经所述输入装置输送的视频流以用于检测其中场景变化的发生;和

[0063] (c) 在所述器件中,基于包括了对应于视频流中场景变化的标识符的辅助媒体,使视频流和辅助媒体对用户的呈现同步,其中所述辅助媒体包括流光溢彩脚本,所述方法在应用于所述器件中时可被操作以同步所述流光溢彩脚本和视频流。

[0064] 根据本发明的第五方面,提供了可在根据本发明的第二方面的器件的计算硬件上运行的软件,以用于实现根据本发明第四方面的方法。

[0065] 根据本发明的第六方面,提供一种用于流光溢彩脚本的数据结构,所述流光溢彩脚本用作用于根据本发明第二方面的器件的辅助媒体,所述数据结构包括场景变化标识符,并且流光溢彩随着每个场景而变化。

[0066] 可选地,在所述数据结构中,流光溢彩脚本包括用于动态控制视频流的场景变化分析的一个或多个参数,所述流光溢彩脚本适合于补充所述视频流。

[0067] 短语“数据结构”将以与数据信号格式相类似的方式解释,所述数据信号格式是早期已经被示出以易于吸引专利权利保护的主体。

[0068] 应当理解,在不脱离如由所附权利要求定义的本发明的范围的情况下,本发明的特征易于被结合成任意组合。

附图说明

[0069] 现在将仅仅通过示例的方式并参照附图来描述本发明的实施例,其中:

[0070] 图 1 是流光溢彩系统的简单表示；

[0071] 图 2 是与图 1 的系统相关的数据同步的说明；

[0072] 图 3 是指纹和在易于呈现在如图 1 所描述的流光溢彩系统上的视频材料中形成场景的图像序列的说明；和

[0073] 图 4 是被配置以实现图 1 的系统的器件的说明。

[0074] 在附图中,使用下划线的数字表示所述下划线数字位于其上的项目或所述下划线数字邻近的项目。非下划线数字涉及这样的项目:其通过将所述非下划线数字连接到该项目的直线被识别。当数字是非下划线的并其伴随着相关箭头时,该非下划线数字用于识别箭头所指的一般项目。

具体实施方式

[0075] 参照上面所描述的图 1 和图 2,本发明关注同步两个数据信号的自动方法,例如前述的流光溢彩脚本 100 与其对应的视频材料 110 的同步。公知的是,通过水印的方式同步,和通过从信号中提取指纹方式同步。这种指纹按惯例被确定用于数据信号的各部分,例如国际 PCT 专利申请 PCT/IP2004/051259 (WO2005/011281) 所述,可以获得用于视频材料 110 的 2 秒播放部分的指纹。然而,就流光溢彩技术来说,这种先前已知的方法涉及计算并且该方法还是次优的。而且,通常不方便的是,为指纹匹配的目的必须访问外部数据库,因为这种访问阻止了对于消费型电子视频产品所期望的单机操作。

[0076] 本发明的意图是能够例如应付这样的情况:用户 20 购买 DVD,该 DVD 在其上记录了视频材料 110,其也方便地被称为视频内容。用户 20 在其流光溢彩系统 30、40 上播放所述 DVD,可选地利用从应用于视频材料 110 的分析中获得流光溢彩来发现适合的视频材料 110。用户 20 可以随后在使用互联网时发现对应的流光溢彩脚本 100 的最新版本可以用于伴随视频材料 110,所述视频材料包括图像和声音信息,并且接着可以从互联网下载所述脚本 100。可替代地,器件 400 可以证实脚本的可用性并且在没有任何用户交互的情况下将它更新,这提供了更多的方便,因为用户无需被寻找更新所打扰。当用户的流光溢彩系统 30、40 可以根据本发明操作时,该流光溢彩系统 30、40 能够使流光溢彩脚本 100 的最新版本与视频材料 110 同步并且为用户 20 提供了与呈现给用户 20 的视频材料 110 同步的所下载脚本 100 的增强的观察体验。现在将进一步描述按照本发明脚本 100 与视频材料 110 的同步。

[0077] 参照图 3,视频材料 110 一般是有区别的,因为它包括由 200a、200b、200c 等等表示的一系列场景。视频材料 110 包括视频图像序列,当用户 20 消费所述视频图像序列时,其易于以由箭头 230 表示的时间顺序而被呈现。例如,所述序列包括由 220 表示的视频图像,其包括在第一场景 200a 中。方便地,所述图像可选地被编码在视频材料 110 中作为包括 I 帧数据、P 帧数据和 B 帧数据的当前 MPEG-4 数据。当被编码为当前 MPEG-4 数据时,每个场景 200a、200b、200c 可以包括多个 I 帧;I 帧可被操作以定义对应图像的所有细节,而 B 帧和 P 帧定义与它们的最近的在前 I 帧有关的进行中的变化的发生。由于 MPEG-4 可以获得的数据压缩程度,在当前数据分发系统中通常使用 MPEG-4 编码。

[0078] 在第一场景 200a 中,从左到右示出了拍摄的运动,其中用于记录第一场景 200a 的图像的照相机镜头逐步从由 240 表示的人转到由 250 表示的盘旋在人 240 上面的飞行器;场景 200a 中的第一图像 260 完全不同于序列 200a 中的最后一幅图像 270,虽然从第一幅图

像 260 到最后一幅图像 270 是逐步过渡的。而且,在第二场景 200b 中,从序列 200b 的图像的右侧到其左侧来拍摄由 300 表示的树;序列 200b 的第一和最后一幅图像 310、320 分别包括至少部分相似的图像特征细节。此外,在第三场景 200c 中,场景 200c 的每个图像基本上彼此相同。

[0079] 在本发明的方法中,分析视频材料 110 (优选在其被呈现给用户 20 时的消费点上) 以确定在视频材料 110 中出现的场景。如前所述,在本发明的方法中,根据下面的一个或多个情况,视频材料 110 的场景变化潜在地可识别:

[0080] (a) 相对于图像 I_n 和 I_{n+1} 的各个像素指定的总体照明强度或颜色,从所述序列中的给定图像 I_n 到序列中随后的图像 I_{n+1} 的突然改变;例如,当图像 270 对应于开放晴朗的环境而随后的图像 310 对应于黑暗的森林时,在图像 270 与其随后图像 310 之间出现突然改变;

[0081] (b) 存在于给定图像 I_n 中的基本所有空间图像特征相对于其随后图像 I_{n+1} 的突然改变;

[0082] (c) 与给定图像 I_n 相关的音频频谱内容相对于其随后图像 I_{n+1} 的突然改变;

[0083] 相反地,根据下面一个或多个情况,通过所述方法可以识别给定场景中邻近的图像:

[0084] (d) 给定的图像 I_n 和其随后图像 I_{n+1} 具有相关的或对于图像 I_n 和 I_{n+1} 共同的空间图像特征;例如,场景 200b 中的图像 310 和其随后图像 330 具有基本相似的树特征;

[0085] (e) 给定的图像 I_n 和其随后图像 I_{n+1} 具有基本相似的总体照明强度;和

[0086] (f) 给定的图像 I_n 和其随后图像 I_{n+1} 具有基本相似的相关音频频谱内容。

[0087] 对于确定图像序列的两个连续图像是否包括视频材料 110,本发明的方法可选地被操作以使用阈值比较。因此,描述相似性的特征参数 K_i ,和因此还有所述序列中邻近图像之间的差异的度量,可选地被计算或由本发明的方法通过应用一般由公式 1 (Eq. 1) 描述的处理函数 F 来获得:

$$[0088] \quad K_i = F(\Delta A_i, \Delta S_i, \Delta I_i) \quad \text{Eq. 1}$$

[0089] 其中

[0090] ΔA_i = 与图像 I_n 和 I_{n+1} 相关的音频频谱特征差异的度量;

[0091] ΔS_i = 与图像 I_n 和 I_{n+1} 相关的空间特征信息差异的度量;和

[0092] ΔI_i = 与图像 I_n 和 I_{n+1} 相关的空间特征信息差异的度量。

[0093] 当将所述方法用于确定在视频材料 110 的图像序列中是否已经发生场景变化时,可选地可以这样应用阈值:如果特征参数 K_i 的值低于阈值 T (即 $K_i < T$), 图像 I_n 和 I_{n+1} 在场景内是相邻的,而当特征参数 K_i 超过阈值 T (即 $K_i > T$) 时,在图像 I_n 和 I_{n+1} 之间发生了场景变化。可选地,一个或多个度量 ΔA_i 、 ΔS_i 、 ΔI_i 可以用于在计算公式 1 的过程中与一个或多个阈值进行比较。可选地,在该方法中可以使用相对于应用多个阈值之一的其他识别形式。

[0094] 因此,可被操作以接收视频材料 110 和流光溢彩脚本 100 并将前述的本发明的方法应用于视频材料 110 的用户 20 的器件能够识别包括了视频材料 110 的图像序列中发生的场景变化;根据该场景变化中,获得了场景持续时间并由此获得了与视频材料 110 有关的指纹 340。而且,当操作视频材料 110 以将之呈现给用户 20 消费时,所述器件可被有益地操作以实时检测场景变化。可替代地,所述器件可以被指示来预先分析所接收到的视频材

料 110 并修改所述材料以包括表示发生场景变化的标识 130,或指纹 340。因此,前述的根据视频材料 110 生成的指纹 340 可以被预先生成或实时逐步生成。

[0095] 应当理解,包括在视频材料 110 的图像序列中的突发性影响(例如与突发闪电和雷声相关的雷击)易于造成器件潜在地不正确地检测场景变化。然而,如果将流光溢彩脚本 100 相似地编码从而使得雷击在其中被表示为脚本 100 中的场景变化,这就不是问题了。当在所述器件中解释脚本 100 时,这种雷击将造成流光溢彩源 140 模拟与出现在电视屏幕 30 上的雷击图像同步的照明效果。

[0096] 流光溢彩脚本 100 被有益地根据一系列场景编码,其中在脚本 100 中描述了包括在每个场景中的图像的数量,连同在所述场景期间以何种方式驱动光源 40 的控制参数。换句话说,流光溢彩脚本 100 优选被生成来匹配从视频材料 110 获得的前述指纹 340。因此,脚本 100 可选地提供一表示在一系列场景中正在被考虑的场景的索引 α ,和表示具有索引 α 的特定场景中多个图像的参数 N_α 。在给定的场景内(例如第一场景 200a),流光溢彩源 40 将接收变化驱动(例如在强度和/或颜色方面的变化),因为包括在给定的场景中的图像被呈现给用户 20 用于消费。可选地,脚本 100 可以包括一个或多个阈值参数(例如前述的阈值 T),以当应用本发明的方法时用于计算由公式 1 (Eq. 1) 描述的函数 F,以使得所述器件更可靠地检测如脚本 100 中所描述的场景变化。

[0097] 因此,例如可以可选地以如表 1 提供的方式构造流光溢彩脚本 100。

[0098] 表 1 :

[0099]

场景系列索引 α	包括具有索引 α 的场景的图像数量 N_α	用于流光溢彩强度的流光溢彩驱动参数 L 和用于流光溢彩颜色的 G	图像数量 Q, 驱动参数 L 和 G 用于所述图像来控制所述源 140	附加的可选编码参数 P_α
1	$N_1 = 200$	$L = 10, G = \text{灰色}$ $L = 20, G = \text{蓝色}$	$Q = 150$ $Q = 50$	500 550
2	$N_2 = 401$	$L = 15, G = \text{黄色}$ $L = 18, G = \text{橘黄色}$	$Q = 200$ $Q = 201$	450 480
3	$N_3 = 2000$	$L = 22, G = \text{蓝色}$	$Q = 2000$	500
...
Z	$N_z = 1002$	$L = 12, G = \text{红色}$	$Q = 1002$	500

[0100] 在表 1 中,流光溢彩脚本 100 被分为对应于以 $\alpha = 1$ 开始和以 $\alpha = z$ 结束的一系列场景的流光溢彩指令。用 N_α 表示包括给定场景的图像的数量,例如第一场景 $\alpha = 1$ 其中具有 200 个图像。在脚本 100 中经参数 L 控制用于每个场景的流光溢彩照明的幅度;例如第一场景 $\alpha = 1$ 具有对于所述第一场景的第一 150 个图像的流光溢彩照明幅度 10,和对于所述第一场景的第二 50 个图像的流光溢彩照明幅度 20,第一场景包括总共 200 个图像。相似地,用于第一场景的流光溢彩照明的颜色具有对于第一场景的第一 150 个图像的质量“灰色”,其用参数 G 表示,并且具有对于第一场景(即 $\alpha = 1$)的第二 50 个图像的质量“蓝色”;可选地,由每个光源 40 生成的颜色相对于所存在的其他光源可以在脚本 100 中单独指定。虽然在这里初步阐述了第一场景 $\alpha = 1$,应当理解脚本中随后的场景以相关的方式跟随。可选地,脚本 100 包括用于动态控制视频材料 110 的分析的参数 P_α ,例如动态地改变应用来检测在视频材料 110 中场景变化所出现的地方的阈值 T,以用于更可靠地使

之与流光溢彩脚本 100 同步。

[0101] 虽然表 1 说明了流光溢彩脚本 100 构造的方式,应当理解在由所附权利要求定义的本发明的范围内流光溢彩控制脚本数据的可替代的设置是可行的。

[0102] 在一个这样的实例中,数量 Q 被置为 1。也就是,流光溢彩脚本包括用于每个分离的视频帧的清楚设置。这具有的优点在于流光溢彩脚本可以具有在时间上带有高分辨率的完全控制。

[0103] 在另一个这样的实例中,流光溢彩脚本控制影响用于控制光扬声器的内容分析算法的参数。这种参数的一个实例是弛豫时间,其确定在颜色被相应地改变之前多少帧被分析。该参数通常被设置得十分长,因为太紧张的光效果将被避免,但是为了准确地跟踪闪电,它应该十分短,也许甚至为单个帧。根据本发明控制这种参数具有的优点在于闪电可以被准确地跟踪,因为所述参数被暂时设置为很小的值,同时在例如电影的其他片段期间避免太紧张的光效果。根据本发明控制参数与已知内容分析技术的这种结合具有附加的优点在于,流光溢彩脚本可以相对较小,因为只有很少的位就足以控制所述参数。

[0104] 在图 1 中,从用户 20 的视角示出了流光溢彩系统的典型的部分。在图 4 中,以示意性的形式示出了系统硬件的配置。参照图 4,用 400 表示用于同步视频材料 110 与流光溢彩脚本 100 的系统。所述同步系统 400 包括分别用于接收视频材料 110 和流光溢彩脚本 100 的数据缓冲器 410、420。如前所述,材料 110 和脚本 100 易于从相互不同的数据源被提供来作为在多媒体环境中的共同体验。所述同步系统 400 进一步包括在数字硬件中实现的同步处理器 430;可选地,所述处理器 430 可以使用计算硬件实现,该硬件可被操作以执行造成所述计算硬件可被操作来执行如先前阐述的同步功能的软件。而且,所述系统 400 进一步包括驱动器 440、450,其用于将视频数据分别输出到电视 30 和输出到一个或多个提供流光溢彩照明 10 的流光溢彩源 40。当存在许多基于包括在脚本 100 中的数据而被独立驱动的流光溢彩源 40 时,驱动器 450 包括相应的多个驱动器通道;换句话说,了流光溢彩照明 10 可以在运行中的任意给定的时刻具有空间变化的颜色。

[0105] 在运行中,处理器 430 可被操作以通过确定场景变化发生在哪里来同步视频材料 110,例如参照图 3 先前所述;由此,处理器 430 可被操作以确定视频材料 110 的相应指纹 340。随后,处理器 430 可以被操作以确定用于从缓冲器 410 接收的脚本 100 中相应场景变化的数据,并且由此确定合适的的数据输出到缓冲器 450 中,用于驱动一个或多个照明源 40 的颜色以及照明强度。如早期所述,通过基于以下一个或多个情况的分析可以发现场景变化:

[0106] (a) 在早期场景的至少最后一幅图像相对于随后场景的至少第一幅图像之间的视频材料照明度的突然改变;

[0107] (b) 在早期场景的所述至少最后一幅图像相对于随后场景的所述至少第一幅图像之间在空间图像特征信息方面的突然改变;和

[0108] (c) 早期场景的所述至少最后一幅图像相对于随后场景的所述至少第一幅图像在音频信息方面的突然改变。

[0109] 可选地,不仅基于紧密相邻的图像而且基于从多个先前图像和多个随后图像所采集的信息来检测视频材料 110 内的场景变化和前述指纹 340 的生成,从而改进了检测的可靠性。

[0110] 可选地,在流光溢彩脚本 100 和视频材料 110 同时呈现给用户 20 期间,在处理器 430 不可用的情况下,处理器 430 可被操作以切换到仅仅基于视频材料 110 来进行流光溢彩照明 10 的流光溢彩控制的普通当前模式,以充分地同步流光溢彩脚本 100 与视频材料 110;通过这种自动切换,至少部分地被消除了对已提供给用户 20 的增强视觉体验的破坏,即在处理器 430 被发现不能充分同步视频材料 110 与视频脚本 100 时避免了流光溢彩照明 10 的突然损失。

[0111] 同步系统 400 易于被包括作为电视接收机的集成部分。可替代地,系统 400 可以提供给用户 20 作为外接单元,例如作为外接的“数字黑匣子”,或如本领域公知的机顶盒或多媒体 PC。然而可替代地,同步系统 400 可以在软件中实现,所述软件可以经例如多媒体或其他通信信道下载到装配有计算硬件的电视上;这种实现方法的益处在于,由于将来有更高級的流光溢彩脚本 100 能为用户 20 所用,处理器 430 可以被有效地更新。

[0112] 可选地,视频材料 110 作为视频流被例如从数字广播服务商提供。而且,流光溢彩脚本 100 作为辅助媒体被提供。虽然在视频材料 110 中通过场景识别来描述流光溢彩脚本 100 与视频材料 110 的同步,应当理解,可以以相似的方式同步其他类型的脚本与视频材料 110。例如,可替代类型的脚本可以是“气味脚本”,其中同步系统 400 被提供有电子可控的气味生成器,其可以基于通过在视频材料 110 中出现的场景变化的检测而被同步的气味脚本有选择地激活。其他类型的辅助多媒体包括控制用户周围气流的气流脚本,但是其他类型也是可行的。

[0113] 流光溢彩系统是用于创建例如 TV 周围的光效果以生成身临其境的视觉体验。本领域所公知的是,如何从视频内容获得这个效果,但是主观上更好的效果可以通过指导或控制流光溢彩系统的特技专家来创建。专家可以使用用来显示视频并允许其定义特定光效果的工具。其结果可以是需要与原始视频内容密切相关的脚本。

[0114] 一种分配脚本的相对简单的方法是在将它与视频材料结合在一个媒体中(例如,如 DVD 的光盘)或在广播流中。由于流光溢彩当前不是标准特征,这种同步的方法当前不是广泛可用的。

[0115] 在如互联网中使用单独的来源,有利的是能够下载用于控制流光溢彩系统的脚本,而不必下载视频。这需要允许分别获得脚本和视频的同步方法,同时仍然能够同步脚本和视频。

[0116] 根据本发明的实施例,可以通过计数两个场景变化之间的时间从视频流中获得唯一指纹 340,其被称为“场景变化印记”。N 个场景变化印记的序列被定义为“序列指纹”340。通过在脚本中使用相同的序列指纹,同步脚本与视频是可能的。本发明还适合于需要与视频内容同步的其他辅助媒体。本发明克服了在一个媒体中具有脚本和视频的问题,并确保了二者的同步。本发明还允许可以在时间上独立生成的多个脚本或独立于屏幕上被观看视频的位置来生成脚本。

[0117] 本发明的一个特征是,在场景变化之间的时间可以用作低成本的可靠指纹方法,其对内容、内容恶化或视频/广播标准(NTSC(60Hz),PAL(50Hz)等等)相对不敏感,而对广泛的视频编辑(导致全新视频)相对敏感。

[0118] 为确保所述辅助媒体与可能来自完全不同来源的视频同步,可以使用一种唯一地识别所述视频的指纹。

[0119] 本发明部分地基于下列观点：视频材料无论来自何处，如果它没有被编辑（帧去除或重新排序），在场景变化之间的时间会帮助识别视频场景，特别是如果在许多场景变化上完成了匹配。可以根据视频帧速率和帧数量来获得两个场景变化之间的时间。

[0120] 根据本发明的指纹 340 的这个实施例具有相对较低的成本，因为只需要检测场景变化和确定视频帧速率的能力。这些能力本身分别在本领域中已知，并且可能被认为是标准功能，其已经出现在 TV 中并需要很少的硬件资源。

[0121] 本发明可以嵌入到脚本创作工具中，其将视频作为主要输入并生成视频相关输出，例如具有用于每个场景变化的场景变化印记的流光溢彩脚本。

[0122] 所得到的输出（例如脚本）可以下载到例如 TV 的设备。所述设备可以将第一 N 个场景变化印记作为唯一的启动键。分析屏幕上显示的视频并计数每个场景的帧并乘以帧速率的倒数，从而得到“场景变化印记”。一旦第一 N 个场景变化印记与脚本匹配，则检查第二 N 个场景变化印记，直到第 N 个场景变化印记。N 可以预选确定，并且其可以选择足够大以具有低的假阳性检测（如 0.01%）和选择为高检测概率（例如 99.99% 分值），并且足够小以防止太长的同步时间。

[0123] 一旦脚本被锁定（即与视频同步），其内容被分析或解释并用于控制流光溢彩系统，其结果作为视觉效果呈现在 TV 屏幕周围。

[0124] 在商业广告出现或视频被另外打断时，所述场景变化印记开始失配并且脚本可以被暂停。脚本中在所述中断之前和之后的搜索区域可以被检查以用于匹配场景变化印记。这个用于重新同步的策略基于电影的开始或继续比实际中断早一点以允许人们观看概要。

[0125] 当没有脚本运行或同步丢失时，流光溢彩系统可以被设置为缺省模式，其中该系统跟随视频内容。

[0126] 除了流光溢彩之外，任何其他需要与视频同步的应用可以使用本发明。示例包括：用于例如电影的可下载的字幕、附加的声道（如混音）、视频相关动画和评论。

[0127] 在不脱离如由所附权利要求定义的本发明的范围的情况下，对前面描述的本发明的实施例的修改是可能的。

[0128] 用于描述本发明和要求权利的表达（如“包括”、“结合”、“组成”、“具有”、“是”）应当以非排他性的方式被解释，即允许没有明确描述的项目、组件或元件的存在。单数的附图标记也可被解释为涉及复数。

[0129] 在所附权利要求中括号内包括的数字的意图是辅助理解权利要求，而不应该以任何方式解释为限制由这些权利要求所述的主体。

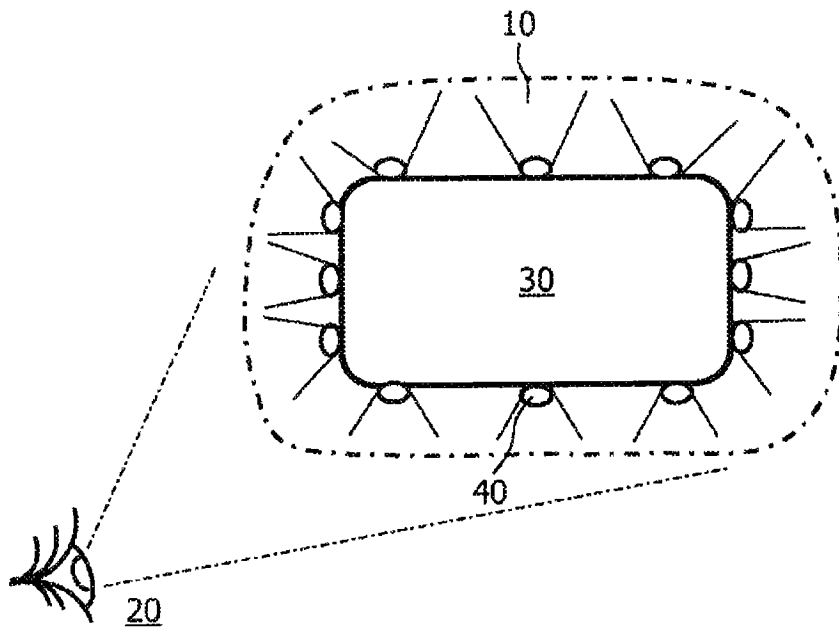


图 1

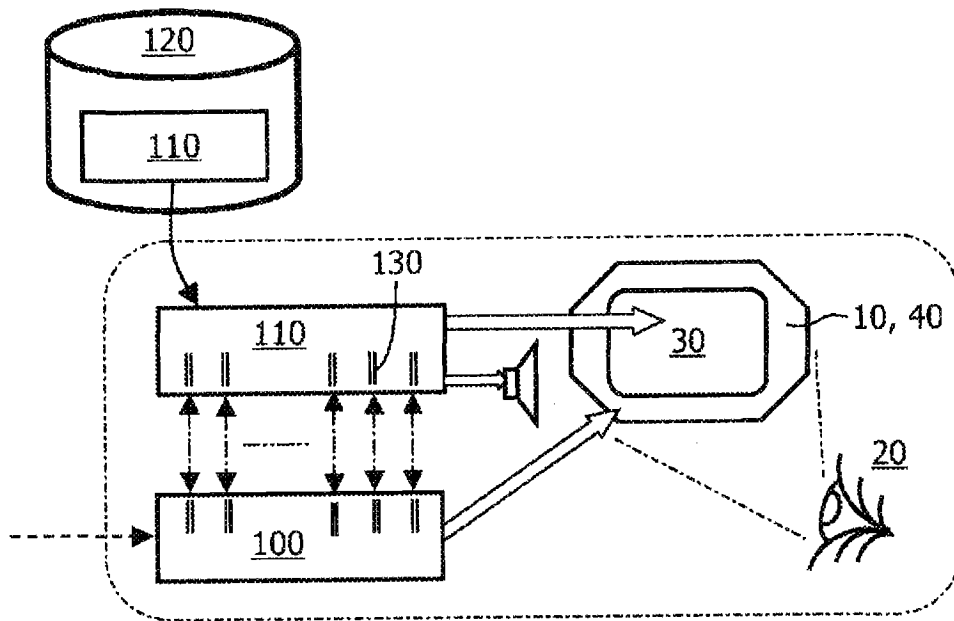


图 2

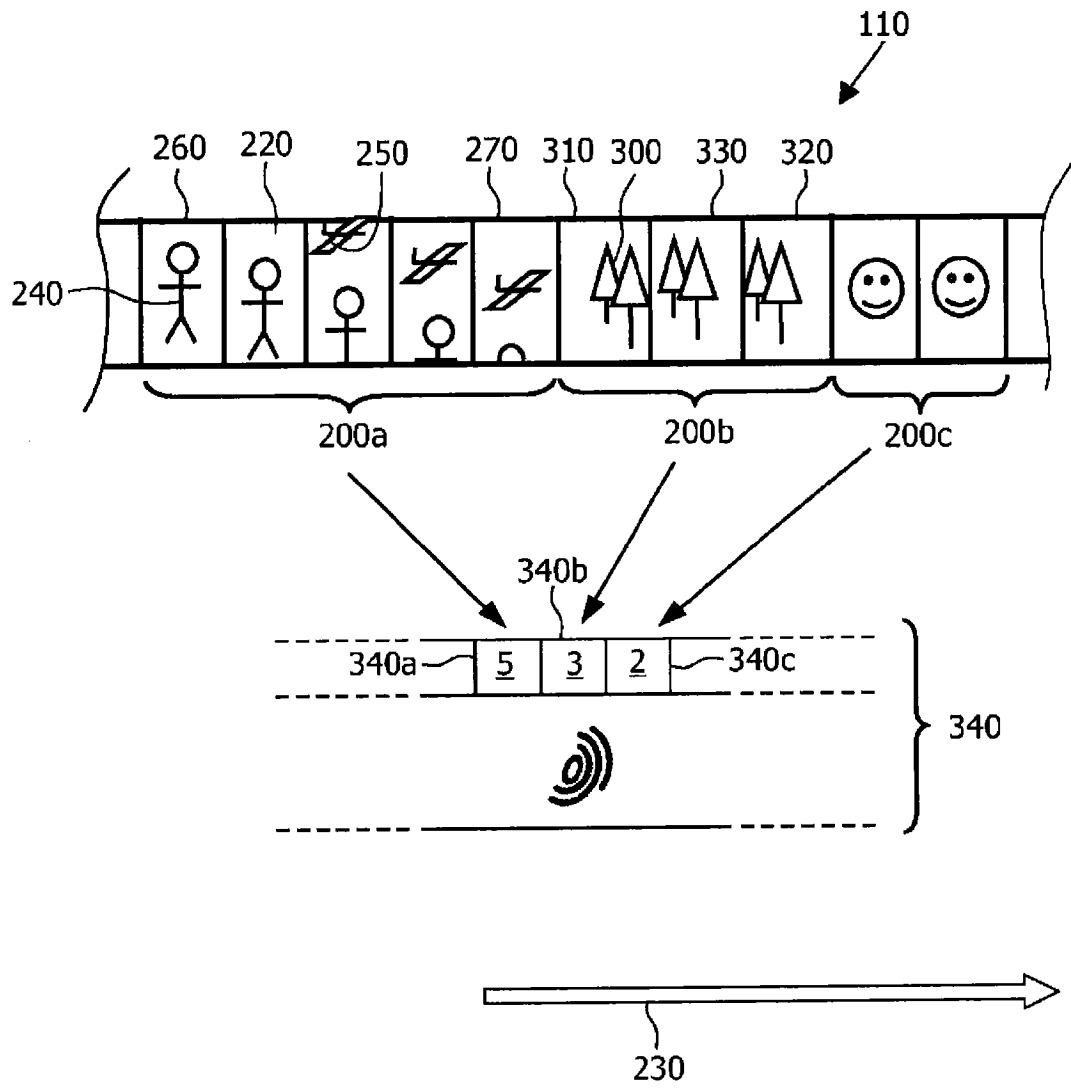


图 3

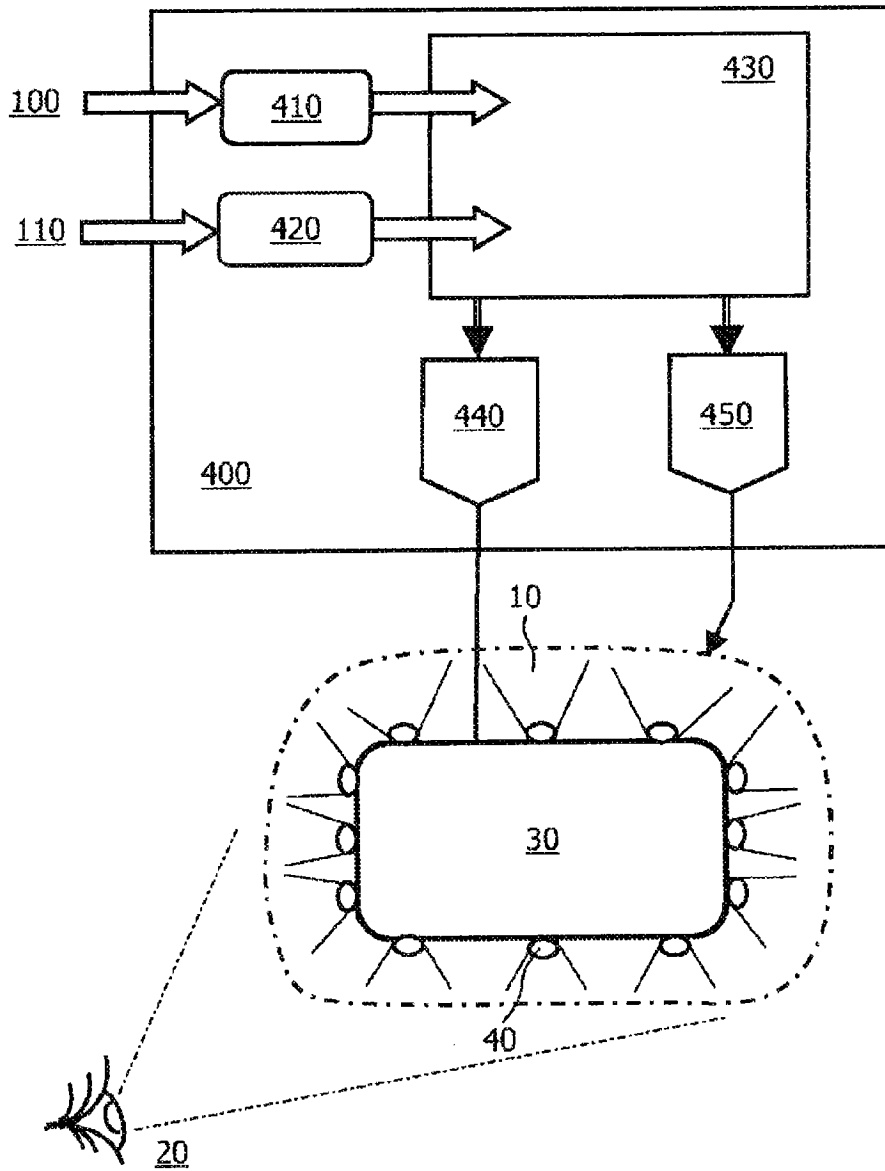


图 4