



(12) 发明专利

(10) 授权公告号 CN 102576411 B

(45) 授权公告日 2015. 01. 07

(21) 申请号 201080049516. 5

(56) 对比文件

(22) 申请日 2010. 08. 30

CN 1572117 A, 2005. 01. 26,

US 7313183 B2, 2007. 12. 25,

(30) 优先权数据

12/553069 2009. 09. 02 US

审查员 李东

(85) PCT国际申请进入国家阶段日

2012. 05. 02

(86) PCT国际申请的申请数据

PCT/US2010/047178 2010. 08. 30

(87) PCT国际申请的公布数据

W02011/028666 EN 2011. 03. 10

(73) 专利权人 索尼电脑娱乐公司

地址 日本东京都

(72) 发明人 H-J. 李

(74) 专利代理机构 中国专利代理(香港)有限公

司 72001

代理人 蒋骏 卢江

(51) Int. Cl.

G06K 9/36(2006. 01)

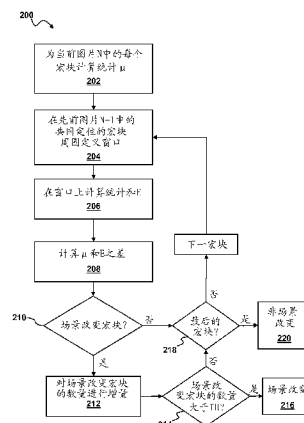
权利要求书2页 说明书8页 附图10页

(54) 发明名称

场景改变检测

(57) 摘要

公开了在编码数字图片中的场景改变检测。为当前图片中的给定段计算统计量  $\mu_M$ 。在先前图片中的共同定位的段周围定义一个或多个段的窗口。在该窗口中的段上计算统计和 E。计算统计和 E 与统计量  $\mu_M$  之差。使用 E 和  $\mu_M$  之差来确定该给定段是否是场景改变段。可从该场景改变段的数量确定当前图片是否是场景改变图片。可存储或传输指示该当前图片是否是场景改变图片的信息。



1. 一种用于在编码一个或多个数字图片中检测场景改变的方法,包括:
  - a) 为当前图片中的给定段计算统计量;
  - b) 在先前图片中的共同定位的段周围定义一个或多个段的窗口;
  - c) 在该窗口中的所述一个或多个段上计算统计和;
  - d) 为该给定段计算统计和与统计量之差;
  - e) 使用该给定段的统计和与统计量之差来确定该给定段是否是场景改变段;
  - f) 为当前图片确定场景改变段的数量;
  - g) 从该场景改变段的数量确定当前图片是否是场景改变图片;
  - h) 存储或传输指示该当前图片是否是场景改变图片的信息。
2. 根据权利要求 1 的方法,进一步包括:
  - i) 根据指示该当前图片是否是场景改变图片的信息来用编码器编码该当前图片。
3. 根据权利要求 2 的方法,其中, i) 包括:如果该信息指示该当前图片是场景改变图片,则将该当前图片编码为 I 帧,或者如果该信息指示该当前图片不是场景改变图片,则将该当前图片编码为 B 帧或 P 帧。
4. 根据权利要求 1 的方法,进一步包括,针对该当前图片中的多个段重复 a) 到 e)。
5. 根据权利要求 1 的方法,其中每个段是宏块。
6. 根据权利要求 1 的方法,其中 a) 包括为当前图片中的每个段计算平均像素值或像素值的方差。
7. 根据权利要求 1 的方法,其中该窗口包括邻近该共同定位的段的一个或多个段。
8. 根据权利要求 1 的方法,其中 c) 包括为该窗口中的每个段确定权重,以及为该窗口中的每个段确定统计值,以及其中统计和由每个段的权重与对应统计值的乘积之和所加权给出,其中该和在窗口中的所有段上获得。
9. 根据权利要求 8 的方法,其中为该窗口中的每个段确定该权重包括确定移动方向,以及根据该窗口中的对应段与该移动方向的接近程度来分配每个权重。
10. 根据权利要求 1 的方法,其中 e) 包括将统计和与统计量之差与阈值相比较。
11. 根据权利要求 1 的方法,其中 g) 包括将场景改变段的数量与阈值相比较。
12. 根据权利要求 1 的方法,其中 g) 包括确定场景改变段的数量是否是该当前图片中的段的大多数。
13. 一种用于在编码一个或多个数字图片中检测场景改变的系统,包括:
  - a) 用于为当前图片中的给定段计算统计量的装置;
  - b) 用于在先前图片中的共同定位的段周围定义一个或多个段的窗口的装置;
  - c) 用于在该窗口中的所述一个或多个段上计算统计和的装置;
  - d) 用于为该给定段计算统计和与统计量之差的装置;
  - e) 用于使用该给定段的统计和与统计量之差来确定该给定段是否是场景改变段的装置;
  - f) 用于为当前图片确定场景改变段的数量的装置;
  - g) 用于从该场景改变段的数量确定当前图片是否是场景改变图片的装置;以及
  - h) 用于存储或传输指示该当前图片是否是场景改变图片的信息的装置。
14. 根据权利要求 13 的系统,进一步包括用于根据指示该当前图片是否是场景改变图

片的信息来编码该当前图片的装置。

15. 根据权利要求 13 的系统,进一步包括用于针对该当前图片中的多个段重复 a) 到 e) 的装置。

16. 根据权利要求 13 的系统,其中 c) 包括用于为该窗口中的每个段确定权重,以及为该窗口中的每个段确定统计值的装置,以及其中统计和由每个段的权重与对应统计值的乘积之和加权给出,其中该和在该窗口中的所有段上获得。

17. 根据权利要求 16 的系统,其中用于为该窗口中的每个段确定该权重的装置包括用于确定移动方向的装置,和用于根据该窗口中的对应段与该移动方向的接近程度来分配每个权重的装置。

18. 根据权利要求 13 的系统,进一步包括用于如果该信息指示该当前图片是场景改变图片,则将该当前图片编码为 I 帧的装置,或者用于如果该信息指示该当前图片不是场景改变图片,则将该当前图片编码为 B 帧或 P 帧的装置。

## 场景改变检测

### 技术领域

[0001] 本发明的实施例涉及数字地编码视频图像,以及更特别地涉及视频图像的编码中的场景改变检测。

[0002] 有关申请的交叉引用

[0003] 本申请要求编号 12/553,069、标题为“场景改变检测”、2009 年 9 月 2 日提交的美国申请的优先权的权益,其整体公开通过引用结合于此。

[0004] 本申请涉及编号 12/533,070、2009 年 9 月 2 日提交并且标题为“用于视频编码的图片级速率控制”的共同转让、共同未决的美国专利申请(代理人案号 SCEA08073US00),其整体内容通过引用结合于此。

[0005] 本申请涉及编号 12/533,073、2009 年 9 月 2 日提交并且标题为“并行数字图片编码”的共同转让、共同未决的美国专利申请(代理人案号 SCEA08077US00),其整体内容通过引用结合于此。

[0006] 本申请涉及编号 12/533,075、2009 年 9 月 2 日提交并且标题为“利用阈值和提早终止以在视频编码器中实现快速运动估计”的共同转让、共同未决的美国专利申请(代理人案号 SCEA08078US00),其整体内容通过引用结合于此。

### 背景技术

[0007] 数字信号压缩被广泛地应用在许多多媒体应用和设备中。使用编码器/解码器(编解码器)的数字信号压缩允许了流媒体,诸如要在互联网上传输或在光盘上存储的音频或视频信号。已经出现了多个不同标准的数字视频压缩,包括 H.261、H.263;DV;MPEG-1、MPEG-2、MPEG-4、VC1;以及 AVC(H.264)。这些标准以及其它视频压缩技术寻求通过在图片中和在连续的图片间消除空间和时间冗余来有效地表示视频帧图片。通过使用这些压缩标准,视频内容可被携带在高度压缩的视频比特流中,并因此有效地存储在磁盘中或在网络上传输。

[0008] MPEG-4 AVC(高级视频编码),也称为 H.264,是一种提供了比其前任显著更大的压缩的视频压缩标准。该 H.264 标准被期望提供达到更早的 MPEG-2 标准的压缩的 2 倍。该 H.264 标准还被期望提供感官质量上的改善。从而,越来越多的视频内容正在以 AVC(H.264)编码流的形式被输送。两个竞争的 DVD 格式,HD-DVD 格式和蓝光盘格式,支持的 H.264/AVC High Profile 解码,以作为强制播放器特征。在 ISO/IEC MPEG&ITU-TVCEG(ISO/IEC JTC1/SC29/WG11 和 ITU-T SG16 Q.6)的联合视频小组(JVT),2005 年 1 月 18-21 日在中国香港的第 14 次会议上,由 Gray Sullivan、Thomas Wiegand 和 Ajay Luthra 所著的“Draft of Version 4 of H.264/AVC(ITU-T Recommendation H.264 and ISO/IEC 14496-10(MPEG-4 part 10)Advanced Video Coding)”中,AVC(H.264)编码被详细描述,其整体内容通过引用被结合于此以用于所有目的。

[0009] 现代视频编码器/解码器(编解码器),诸如 MPEG2、MPEG4 和 H.264 一般将视频帧划分为称为内帧、预测帧和双向预测帧的三个基本类型,其通常分别被称为 I 帧、P 帧和 B

帧。

[0010] I 帧是除了自身而不参考任何图片来编码的图片。I 帧被用于随机访问并且被用于解码其它 P 帧或 B 帧的参考。I 帧可由编码器生成以创建随机访问点(以允许解码器在给定图片位置从头开始正确地解码)。I 帧可在差分图像细节禁止生成有效 P 或 B 帧时被生成。由于 I 帧包含完整的图片,因此 I 帧通常需要比 P 帧或 B 帧更多的比特来编码。

[0011] 为了被解码,P 帧需要在先解码一些其它的一个或多个图片。P 帧通常比 I 帧需要更少的用于编码的比特。P 帧包含关于与解码顺序中的先前 I 帧有关的差异的编码信息。P 帧通常参考图片组(GoP)中的在先的 I 帧。P 帧可包含图像数据和运动向量位移两者,以及两者的组合。在一些标准编解码器(诸如 MPEG-2)中,P 帧在解码期间仅使用一个先前解码的图片作为参考,并且需要该图片在显示顺序中也在该 P 帧之前。在 H. 264 中,P 帧在解码期间可使用多个先前解码图片作为参考,并且可具有与用于其预测的一个或多个图片有关的任何的任意显示顺序关系。

[0012] 为了被解码,B 帧需要在先解码 I 帧或 P 帧两者之一。像 P 帧一样,B 帧可包含图像数据和运动矢量位移两者,和 / 或两者的组合。B 帧可包括一些预测模式,该预测模式通过对使用两个不同的先前解码的参考区域所获得的预测进行平均来形成运动区域(例如,帧的段,诸如宏块或更小范围)的预测。在一些编解码器(诸如 MPEG-2)中,B 帧从不被用于其它图片预测的参考。从而,可对这种 B 图片使用更低质量的编码(导致所使用的比特少于使用其它的方式),因为细节的损失将不会损害后继图片的预测质量。在其它编解码器中,诸如 H. 264,B 帧可被用作或不被用于其它图片解码的参考(任凭编码器自行处理)。一些编解码器(诸如 MPEG-2),在解码期间精确使用两个先前解码的图片作为参考,并且需要那些图片之一在显示顺序中在该 B 图片之前而另一个图片在该 B 图片之后。在其它编解码器中,诸如 H. 264, B 帧可在解码期间使用一个、两个或多于两个的先前解码的图片作为参考,并且可具有与用于其预测的一个或多个图片有关的任何的任意显示顺序关系。B 帧通常需要比 I 帧或 P 帧两者之一更少的用于编码的比特。

[0013] 如在此所使用的,术语 I 帧、B 帧和 P 帧可被应用到与 I 帧、B 帧和 P 帧具有相似属性的任何流数据单元,例如,如相对于流视频的上下文在以上所描述的。

[0014] 视频编码通常利用实际上在给定视频场景之内,该场景的视觉内容的特定元素倾向于保持相对静止。因此,通过按照该图片和用作参考的先前图片之间的差异来编码给定图片,可能减少对视频信号进行编码所需要的数据的压缩。然而,如果视频序列包含场景改变,则可能没有可用作参考的先前图片。由于场景的改变可对编码过程具有影响,因此编码程序能够检测场景改变是有用的。以前的场景改变检测算法已经基于了对视频帧内容的分析。

[0015] 本发明的实施例出自于该上下文内。

## 附图说明

[0016] 通过结合附图一起考虑随后的详细描述,本发明的技术可被容易的理解,其中:

[0017] 图 1A 是说明了两个视频帧之间的场景改变的示意图。

[0018] 图 1B 是说明了用于编码的视频图片的一种可能的划分的示意图。

[0019] 图 2A 是说明根据本发明的实施例的场景改变检测算法的示例的流程图。

[0020] 图 2B-2C 是说明了根据本发明的实施例的基于宏块的场景改变检测算法的框图。

[0021] 图 3A-3C 示出了用于加权过滤窗口的三个不同的开窗形状。

[0022] 图 4A-4B 示出了每个都由四个  $8 \times 8$  块所组成的宏块。

[0023] 图 5A-5B 示出了具有移动方向的宏块。

[0024] 图 6 是说明了根据本发明的实施例的用于在视频图片中实现场景改变检测的装置的框图。

[0025] 图 7 是说明了根据本发明的实施例的用于在视频图片中编码实现场景改变检测的装置的单元处理器实现的示例的框图。

[0026] 图 8 是依照本发明的实施例的包含用于实现图片级速率控制的计算机可读指令的计算机可读介质的框图。

### 具体实施方式

[0027] 尽管随后的详细描述包含出于说明目的的许多特定细节,但本领域任何普通技术人员将理解,对随后的细节的许多变形和变更在本发明的范围之内。由此,以下描述的本发明的示例性实施例在对要求保护的发明没有任何一般性损失,并且没有强加的限制的情况下被阐述。

[0028] 介绍

[0029] 可参照图 1A 来理解场景改变检测的问题的本质,图 1A 分别描绘了第一和第二视频帧 101、102。帧 101、102 是视频序列中的连续帧。第一帧 101 示出了针对特定背景 105 从特定角度获得的两个特征 103、104 的视图。第二帧 102 示出了相同的两个特征 103、104,并且在第一帧 101 和第二帧 102 之间,特征 103、104 的相对位置已经改变。此外,在第一帧 101 中,特征 103 之一是微笑,而在第二帧 102 中,该特征 103 是皱眉。

[0030] 从内容观点(现有技术所采取的方法)来看,第一帧 101 和第二帧 102 之间的差异可解释为场景改变。然而,从编码观点来看,由于该场景的元素(即,该示例中的特征 103、104 和背景 105)是基本上相同的,因此这不是场景改变。因此,第一帧 101 可被用作用于第二帧 102 的参考,只要它们之间的距离在基础的运动搜索范围之内。不幸的是,使用基于内容的场景改变检测的现有技术的视频编码方案将对此无法识别并将把第二帧 102 编码为独立帧(I 帧)。这将在帧 101、102 的编码中引起低效率。

[0031] 为克服这种问题,本发明的实施例可实现视频场景改变检测算法。在此提出的算法可在视频序列中有效地定位场景切换,使得随后的视频处理模块(例如,视频压缩器/编码器)能利用这种有用的信息。在视频编码应用领域中的典型应用是确定图片帧的类型。例如,具有场景剪切的图片帧将被标记为 I 图片帧,其中 MPEG 视频编码器能够独立地编码这种帧而不用参考过去的图片帧。

[0032] 在此描述的场景改变检测算法可在多个方面与其它场景改变算法相区别。特别是,该提出的算法可与视频编码在概念上结合,而不必与视频编码在实现上结合。就是说,该算法不需要成为视频编码程序的一部分,即使许多现有的场景改变算法被嵌入在运动估计模块中。这种算法可以与视频编码引擎分离并且可被看作要被编码的视频序列的预处理的形式。

[0033] 用该提出的算法,除了图片内容外,关于图片帧是否是场景改变的确定还基于编

码效率。例如，图片帧示出了与其邻近帧不同的具有非常相似的大背景对象的小前景对象。从内容观点来看，该图片帧可能被标记为场景改变帧，然而，从编码效率观点来看，该图片帧的大部分（例如，背景对象）是冗余的并且与在先帧中的该部分相似。这种冗余信息可由基础运动估计 / 补偿模块所提取并去除。

[0034] 提出的算法是基于宏块的方法，该方法与典型的基于 MPEG 的视频编码器中的编码粒度相匹配。提出的算法还利用了具有加权因子的窗口过滤以确定当前图片中的宏块与先前图片的共同定位和邻近宏块的差异的程度。通过基于窗口的加权过滤，可为最终的决策计算和收集统计的差异。

[0035] 为了编码和解码，可按合适大小的单元分解视频图片。例如，如图 1B 中所示，单个图片 110 可被分解成一个或多个段。如在此使用的，术语“段”是指图片内一个或多个像素的组。段的范围可以从图片内的单个像素，直到整个图片。段的非限制的示例包括片段 112、宏块 114、子宏块 116、块 118 和个别像素 120。如在图 1B 中说明的，每个片段 112 包含一行或多行宏块 114。在一行中的宏块的数量取决于宏块的大小和图片 110 的大小和分辨率。例如，如果每个宏块包含 16 乘 16 个像素，则每行中的宏块的数量可以通过将图片 110 的宽度（以像素为单位）除以 16 来确定。每个宏块 114 可被分解成多个子宏块 116。每个子宏块可被分解成多个块 118，以及每个块可包含多个像素 120。作为示例且在不限本发明的情况下，在普通视频编码方案中，每个  $16 \times 16$  像素宏块 114 可被分解成四个  $8 \times 8$  像素子宏块 116。每个子宏块可被分解成四个块 118，以及每个块可包含 4 像素乘 4 像素的排列，其包含 16 个像素 120。

[0036] 算法描述

[0037] 在图 2A 中示出了场景改变检测算法 200 的示例。如以上所提到的，视频流中的每个图片可被分解为多个段。在本发明的实施例中，场景改变的确定是基于相对于先前图片对当前图片逐段分析的。出于示例而非限制的目的，该分析被描述为以逐宏块为基础来完成。本领域技术人员将认识到，该算法可被推广到其它段大小。

[0038] 如在 202 处所指示的，可为当前图片 N 中的每个宏块计算统计量  $\mu_m$ ，以特征化每个宏块。有用的统计量的示例包括该宏块的平均像素值（例如色度或亮度）或这样的值的方差。每个宏块的统计量  $\mu_m$  的值可以以逐宏块为基础存储在存储器中。

[0039] 如图 2B-2C 中所描绘的，当前视频图片 N 可被逐宏块地划分，与在先图片 N-1 相同的方式。算法 200 使用在先帧基于分析先前图片 N-1 和当前图片 N 中宏块的分析来判断当前帧 N 是否是场景改变帧。特别地，如以下所讨论的，可为在先帧 N-1 中的每个宏块计算统计量  $\mu_m$  的值并将该值存储在计算机存储器中，以用于场景改变的确定。

[0040] 如在 204 处所指示的，可在先前图片 N-1 中的共同定位的宏块周围定义窗口。作为示例而不作为限制，对于当前图片 N 中的宏块 A，可在先前帧 N-1 中找到共同定位的宏块 a，并且邻近宏块 a 的所有八个宏块（从 b 到 i）也被定位以形成加权过滤窗口。如在此所使用的，如果宏块 A 和 a 在不同图片中但处于那些图片内的对应的位置，则宏块 A 和 a 可被称为是共同定位的。作为示例，宏块 A 可以是当前图片中的第 k 行宏块中的第 j 个宏块，而共同定位的宏块 a 可以是在先宏块 N-1 的第 k 行中的第 j 个宏块。注意的是，人们可选择不同的窗口形状，只要该形状对场景改变检测的目的有技术意义。例如，图 3A-3C 描绘了三种可能的替代的开窗形状，其可用于本发明的实施例中的过滤窗口。

[0041] 如在 206 处所指示的,一旦在先前图片中的共同定位的宏块周围选择过滤窗口,可基于该窗口中的每个宏块的统计量值来计算统计和 E。作为示例而不作为限制,该统计和 E 可以是在窗口上取得的统计量  $\mu_m$  的加权和。在一些实施例中,为窗口中的每个宏块分配加权因子以捕获每个宏块与中心宏块(例如 A 和 a)的相关程度。作为示例,如果宏块 A 和 a 在统计上显著相似,则窗口中的所有其它邻近宏块可被分配相等的加权因子。如果 A 和 a 在统计上不显著相似,则可得到关联的移动方向 M。基于检测的移动方向 M,位于沿移动方向 M 的窗口内的宏块相比于那些没有位于沿移动方向 M 的宏块可被分配更大的权重。有多种方式可得到移动方向。作为示例而不作为限制,如图 4A-4B 中所示,基于简单  $8 \times 8$  像素块(一个宏块通常包括 4 个  $8 \times 8$  块)的方向确定可被用于估计移动方向。如图 5A-5B 中所示,根据最大块差异可导致移动方向。

[0042] 可以以任何合适的方式得到移动方向。作为示例而不作为限制,该移动方向可以与先前图片 N-1 中的宏块 a 与当前图片 N 中的宏块 A 之间的在统计上最不相同块之间的直线平行。

[0043] 在图 5A-5B 中所描绘的示例中,位于沿移动方向 M 的宏块 f、c 和 d 相比于没有位于沿移动方向 M 的宏块 g、b、i、e 和 h 可被分配更多的权重。加权因子的和可被设置为常数 1 以用于标准化,即:

[0044]

$$\sum_{m=0}^i W_m = 1.0$$

[0045] 其中  $W_m$  是给定宏块 m 的权重因子,  $m = \{a, b, c, d, e, f, g, h, i\}$ , 并且  $0.0 \leq W_m \leq 1.0$ 。

[0046] 在分配加权因子之后,可计算总的加权和。例如,假设使用宏块的平均像素值(也可使用其它统计数据)并表示为  $\mu_m$ , 则总的加权和 E 可由以下给出:

[0047]

$$E = \sum_{m=0}^i (W_m \times \mu_m)$$

[0048] 如在 208 处所指示的,一旦已经计算出加权和 E,则可计算宏块 A 的平均值  $\mu_A$  与加权和 E 之差值。如果该差在预定义的阈值 (TH) 之上,则宏块 A 可被当作场景改变宏块,否则其是非场景改变宏块。注意的是,可基于图片的内容为每个图片更新阈值 TH。

[0049] 在已经作出关于当前图片 N 中的每个宏块是场景改变宏块还是非场景改变宏块的确定之后,可计算出当前图片 N 中的场景改变宏块的数量。然后可能基于场景改变宏块的数量来确定该图片帧是否具有场景改变。作为示例而不作为限制,该算法可以基于以下判断:场景改变宏块的数量是当前图片 N 中的宏块的大多数,还是大于该图片的预定义阈值。在图 2A 中所描绘的示例中,通过对场景改变宏块的计数进行增量来保持场景改变宏块的数量流水记数 (running tally)。例如,如果在 210 处确定了宏块为场景改变宏块,则该计数被增量,如在 212 处所指示的。在 214 处如果记数超出了阈值,则在 216 处该图片可被认为是场景改变图片,并且可避免另外的场景改变宏块计算。在 210 处如果发现宏块是非场景改变宏块,则在 218 处如果该宏块是最后的宏块,并且在 214 处该记数仍在阈值之下,则在 220 处可确定当前图片 N 为非场景改变图片。否则,为当前图片中的下一个宏块重复该过程。一旦已经为当前图片中的每个宏块重复了该过程,则可为后继图片 N+1 重复该



过程。要注意的是,为当前图片 N 中的每个宏块计算的  $\mu_m$  的值,可被存储以用于在计算图片 N 的统计和 E 中使用,该图片 N 是相对图片 N+1 的在先图片。

[0050] 一旦作出了关于当前图片是否是场景改变图片的确定,例如,使用软件编码器来可编码当前图片。算法 200 可存储或传输指示了当前图片是否包含场景改变的信息。编码器可在编码当前图片时使用该信息。例如,如果该信息指示了当前图片是场景改变图片,则编码器可将当前图片编码为 I 帧。否则,编码器可将该图片编码为 P 帧或 B 帧。

[0051] 以上描述的算法可被实现在合适的计算机装置上。图 6 说明了计算机装置 600 的示例的框图,该计算机装置 600 可被用于实现场景改变检测,用于如上所描述的图片编码。装置 600 一般可包括处理器模块 601 和存储器 602。处理器模块 601 可包括一个或多个处理器核。作为使用多个处理器模块的处理系统的示例,是单元处理器(Cell processor),单元处理器的示例被详细描述在例如单元宽带引擎体系机构(Cell Broadband Engine Architecture)中,其可在 [http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/\\$file/CBEA\\_01\\_pub.pdf](http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/$file/CBEA_01_pub.pdf) 在线获得,其通过引用结合于此。

[0052] 存储器 602 可以是以集成电路的形式,例如, RAM、DRAM、ROM 等等。该存储器也可以是由所有处理器模块 601 可访问的主存储器。在一些实施例中,处理器模块 601 可包括多个处理器核以及与每个核相关联的本地存储器。编码器程序 603 可以以可在处理器模块 601 上执行的处理器可读指令的形式存储在主存储器 602 中。该编码器程序 603 可被配置为将图片编码成压缩的信号数据和 / 或配置为解码压缩的信号数据。作为示例而非作为限制,可如在公开号为 20090010338 的共同转让共同未决的专利申请中所描述的来配置该编码器程序 603,该专利申请的内容通过引用将其整体结合于此。如以上所讨论的,编码器程序 603 可被配置为接收关于给定的图片是否是场景改变图片的信息,并且部分地基于该当前图片是否是场景改变图片来将该图片编码为 I 帧、P 帧或 B 帧。

[0053] 场景改变检测程序 604 也可被存储在存储器 602 中。场景改变检测程序可包括指令,当由处理器模块 601 执行该指令时,例如,根据图 2A 的算法 200,实现如以上所描述的场景改变检测。该编码器程序 603 和场景改变检测程序 604 可以以任何合适的处理器可读语言来编写,处理器可读语言例如是 C、C++、JAVA、汇编、MATLAB、FORTRAN 以及多个其它语言。

[0054] 输入数据 607 可被存储在处理器 602 中。这种输入数据可包括流数据的缓冲部分,例如编码的视频图片或其部分。在执行编码器程序 603 和 / 或场景改变检测程序 604 期间,程序代码和 / 或数据 607 的部分可被加载到存储器 602 或用于由多个处理器核所并行处理的处理器核的本地存储设备中。作为示例而不作为限制,输入数据 607 可包括编码前或处于编码中间阶段的视频图片或视频图片的段。这些各种段可被存储在一个或多个缓冲器中。特别是,这些段可被存储在实现在存储器 602 中的输出图片缓冲器中。

[0055] 装置 600 还可包括公知的支持功能 610,诸如输入 / 输出(I/O)元件 611、电源(P/S)612、时钟(CLK)613 和高速缓存器 614。装置 600 可以可选地包括大容量存储设备 615,诸如磁盘驱动器、CD-ROM 驱动器、录音带驱动器等等,以便存储程序和 / 或数据。装置 600 还可以可选地包括显示单元 616 和用户接口单元 618 以促进在装置 600 和用户之间的交互。显示单元 616 可以是以显示文本、数字、图形符号或图像的阴极射线管(CRT)或平板屏幕的形式。用户接口 618 可包括键盘、鼠标、操纵杆、光笔或可连同图形用户接口(GUI)一

起使用的其它设备。装置 600 还可包括网络接口 620 以使设备能够与网络上的其它设备通信,网络例如是因特网。这些组件可以在硬件、软件或固件或这些中的两个或更多个的一些组合中实现。

[0056] 有多种附加方式来用装置 600 中的多个处理器进行流水线并行处理。例如,可能“展开(unroll)”处理循环,例如在两个或更多个处理器核上复制代码并使每个处理器核实现该代码以处理不同的数据块来“展开”处理循环。这种实现可避免与设置循环相关联的等待时间。

[0057] 能够在三个或更多个处理器上实现并行处理的处理器系统的尤其一个示例是单元处理器。存在可被分类为单元处理器的多种不同的处理器体系结构。作为示例且在不限制的情况下,图 7 说明了一种类型的单元处理器 700。该单元处理器 700 包括主存储器 702、单功率处理器元件(PPE) 704 以及 8 个协同处理器元件(SPE) 706。可替代地,该单元处理器 700 可被配置具有任何数量的 SPE。相对于图 7,存储器 702、PPE 704 和多个 SPE 706 可互相通信,并且在环形元件互连总线 710 上与 I/O 设备通信。存储器 702 可包含输入数据 703、编码器程序 709 以及场景改变检测程序 711,该输入数据 703 具有和以上所描述的输入数据 607 一样的特征,该编码器程序 709 具有和以上所描述的编码器程序 603 一样的特征,该场景改变检测程序 711 具有和以上所描述的场景改变检测程序 604 一样的特征。例如,如上所述,至少一个 SPE 706 可在其本地存储设备(LS)中包括场景改变检测指令 705 和/或要被并行处理的缓冲数据的一部分。PPE 可在其 L1 高速缓存器中包括编码指令 707,该编码指令 707 具有和以上所描述的编码器程序 603 一样的特征。指令 705 和数据 707 也可被存储在存储器 702 中,用于在需要时由 SPE 和 PPE 访问。

[0058] 作为示例,PPE 704 可以是具有关联的高速缓存器的 64 位 PowerPC 处理器单元(PPC)。该 PPE 704 可包括可选的向量多媒体扩展单元。每个 SPE 706 包括协同处理器单元(SPU)和本地存储设备(LS)。在一些实现方式中,本地存储设备可具有例如用于代码和数据的大约 256 千字节的存储能力。该多个 SPU 是没有 PPU 复杂的计算单元,因为该多个 SPU 通常不执行任何系统管理功能。该多个 SPU 可具有单指令多数据(SIMD)能力,并且通常处理数据和发起任何需要的数据传输(服从由 PPE 所设置的访问属性)以便执行它们被分配的任务。该多个 SPE 允许系统 700 实现应用,该应用需要更高的计算单元密度并可有效地使用提供的指令集。由 PPE 704 所管理的系统中大量的 SPU 706 允许在广泛的应用上的成本效益处理。作为示例,单元处理器 700 可由被称为单元宽带引擎体系结构(CBEA)的体系结构所特征化。在遵从 CBEA 的体系结构中,可将多个 Pep 组合成 PPE 组,以及可将多个 SPE 组合成 SPE 组。出于示例的目的,单元处理器 700 被描绘为仅具有带有单个 SPE 和单个 PPE 的单个 SPE 组和单个 PPE 组。可替代地,单元处理器可包括功率处理器单元的多个组(PPE 组)和协同处理器单元的多个组(SPE 组)。遵从 CBEA 的处理器被详细描述在例如单元宽带引擎体系机构(Cell Broadband Engine Architecture)中,其可在 [http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/\\$file/CBEA\\_01\\_pub.pdf](http://www-306.ibm.com/chips/techlib/techlib.nsf/techdocs/1AEEE1270EA2776387257060006E61BA/$file/CBEA_01_pub.pdf) 在线获得,其通过引用结合于此。

[0059] 本发明的实施例提供了用于并行场景改变检测和流数据编码的系统和方法,该流数据诸如是流视频。这样的实施例可被应用到大多数视频编码器,尤其是 H. 264/AVC 解码器,以及更特别地,可应用到具有作为模块的视频解码器的产品。这种产品的示例包括但不

限于视频游戏控制台、DVD 播放器、软件(PC)视频解码器 / 播放器、视频开通(video on)手机等等。在替代的实施例中,这种系统和方法可被应用到除了视频外的流数据的解码。这样的实施例的示例包括用于解码流音频数据、图形渲染流、静止图片和 XML 文档的系统和方法。本发明的实施例可能对于记录视频游戏、流送游戏图像到便携式游戏设备、以及上载并与通过网络连接的多个游戏系统共享游戏结果的各种应用是所期望的。

[0060] 如以上所记录的,本发明的实施例可通过修改现有视频编码标准来实现,以允许压缩和编码一系列视频图片中的给定图片之前进行场景改变检测。尽管以上已经根据编码视频图像中的场景改变检测来描述了多个示例,人们可替代地利用本发明的实施例来编码静止图片,例如,以 JPEG 的静止图片。

[0061] 根据另一个实施例,如用于连同以上所描述的图像编码一起执行场景改变检测的指令可被存储在计算机可读存储介质中。作为示例而不作为限制,图 8 说明了计算机可读存储介质 800 的示例。该存储介质包含可由计算机处理设备检索、翻译的格式存储的计算机可读指令。作为示例而不作为限制,该计算机可读存储介质 800 可以是计算机可读存储器,诸如,随机存取存储器(RAM)或只读存储器(ROM)、用于固定的磁盘驱动器(例如,硬盘驱动器)的计算机可读存储磁盘、或可移除磁盘驱动器。另外,该计算机可读存储介质 800 可以是闪存设备、计算机可读磁带、CD-ROM、DVD-ROM、蓝光盘、HD-DVD、UMD 或其它光学存储介质。

[0062] 该存储介质 800 包含配置为在处理器执行时实现场景改变检测的场景改变检测指令 801。该场景改变检测指令 801 可包括一个或多个统计计算指令 802,该统计计算指令 802 实现了为当前图片中的每个宏块来计算统计量  $\mu_m$ 。另外,该场景改变检测指令可包括一个或多个窗口定义指令 803,该窗口定义指令 803 被配置为在先前图像中的共同定位的宏块周围定义窗口。此外,场景改变检测指令 801 可包括一个或多个统计求和指令 804,该统计求和指令 804 在窗口中的宏块之上计算统计和 E。另外,该场景改变检测指令 801 可包括一个或多个差计算指令 805,该差计算指令 805 被配置成为当前图片中的给定宏块计算 E 和  $\mu_m$  之差。该场景改变检测指令 801 还可包括场景改变宏块确定指令 806,该场景改变宏块确定指令 806 使用 E 和  $\mu_m$  之差来确定该给定宏块是否是场景改变宏块。该场景改变检测指令 801 还可包括场景改变宏块计数指令 807,该场景改变宏块计算指令 807 使用当前图片的场景改变宏块的数量来确定该当前图片是否是场景改变图片。

[0063] 存储介质 800 可以可选地包括一个或多个图片编码指令 810,该图片编码指令 810 被配置为部分地基于当前图片是否被确定为场景改变图片来将当前图片编码为 I 帧、IDR 帧、P 帧或 B 帧。

[0064] 虽然以上是本发明的优选实施例的完整描述,但使用各种替代、修改和等同方式是可能的。因此,本发明的范围不应当参照以上描述来被确定,而应当改为参照附加的权利要求连同它们的等同方式的全部范围一起来确定。在此描述的无论是否优选的任何特征,都可与在此描述的无论是否优选的任何其它特征相组合。在随后的权利要求中,除了以其它方式明确阐明的地方,不定冠词“A”或“An”指的是在该冠词之后的一个或多个项中的一些。附加的权利要求不被解释为包括“装置加功能”的限定,除非这种限定在给定的权利要求中使用短语“用于……的装置”来明确叙述。

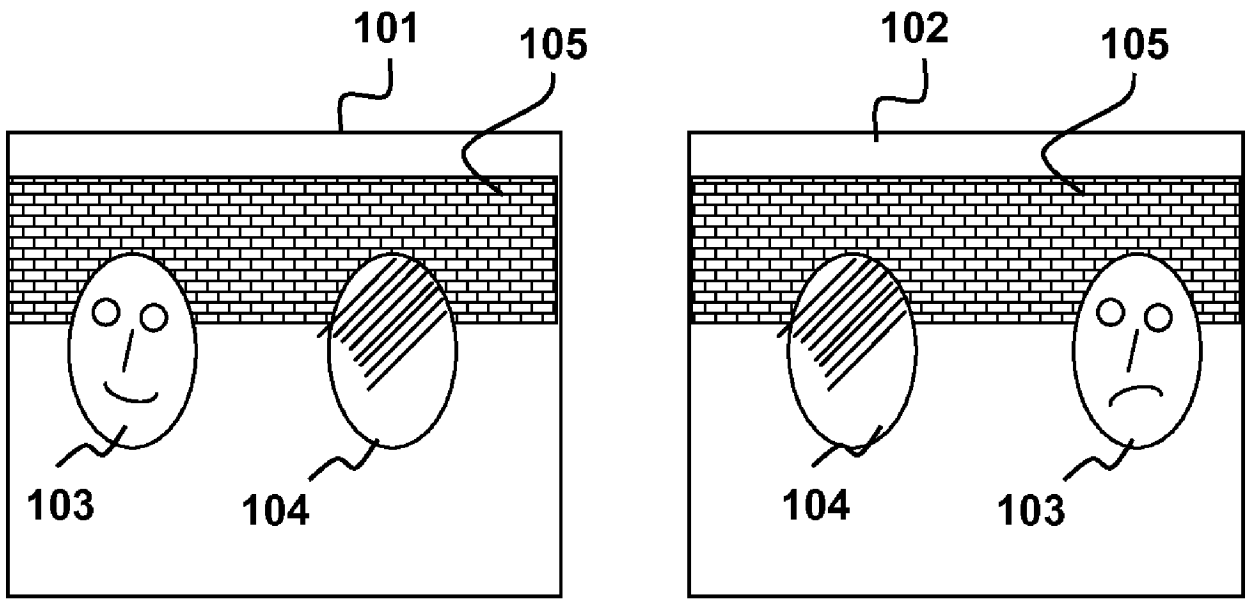


图 1A

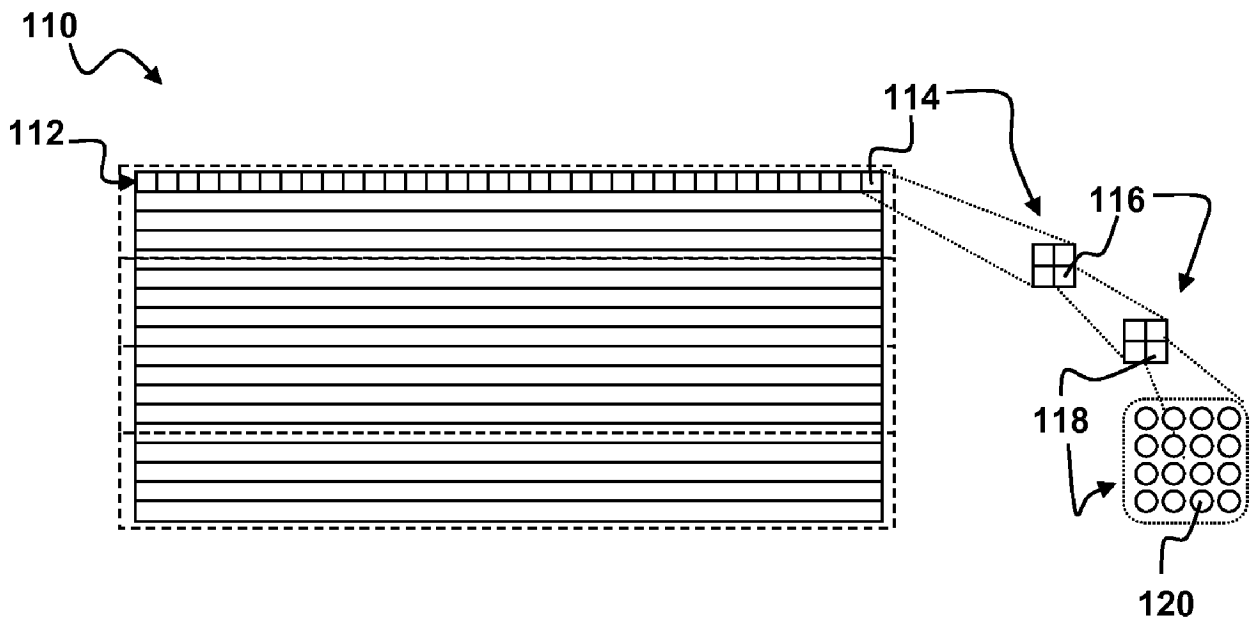


图 1B

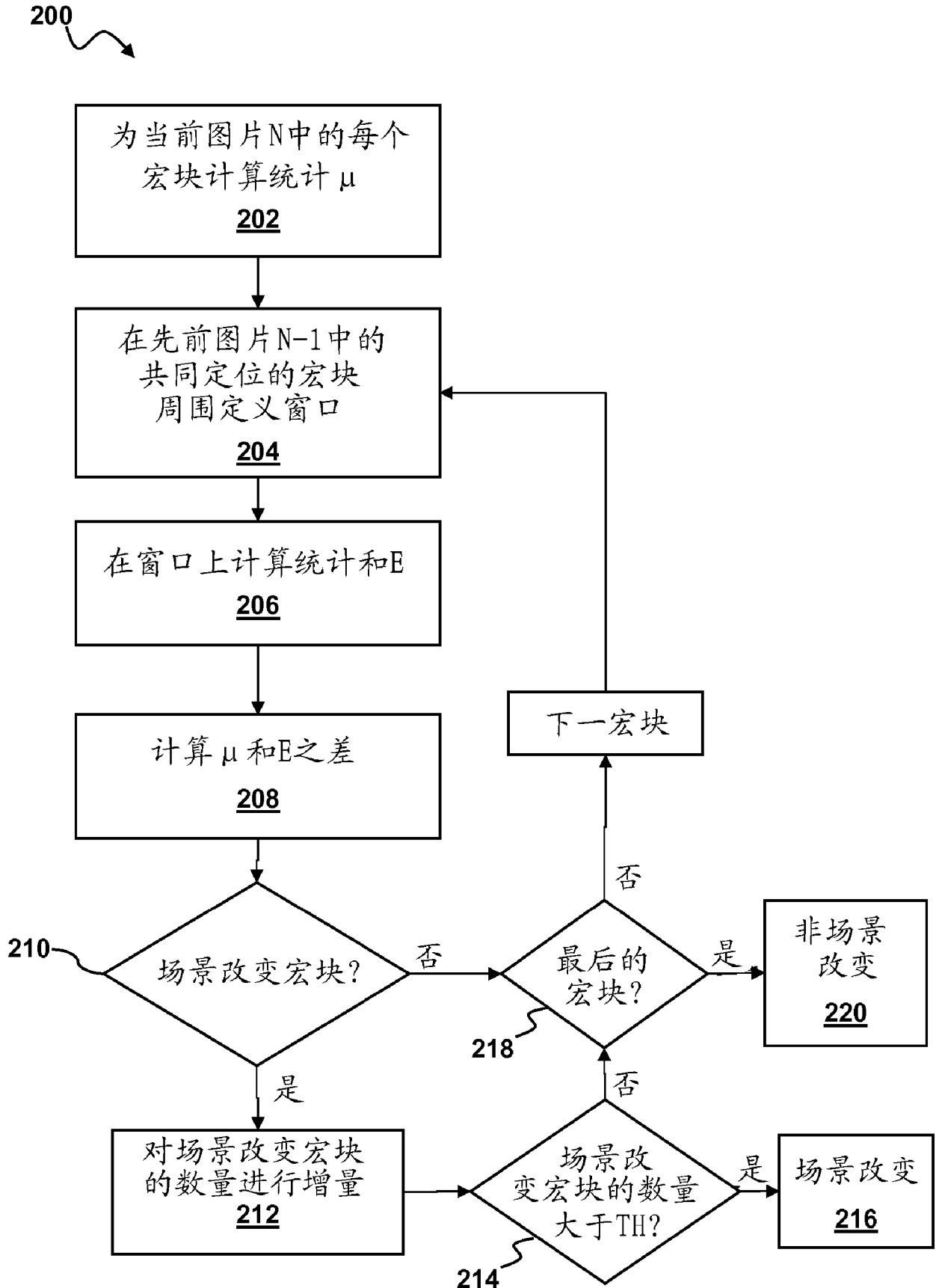


图 2A

# 图片帧N-1

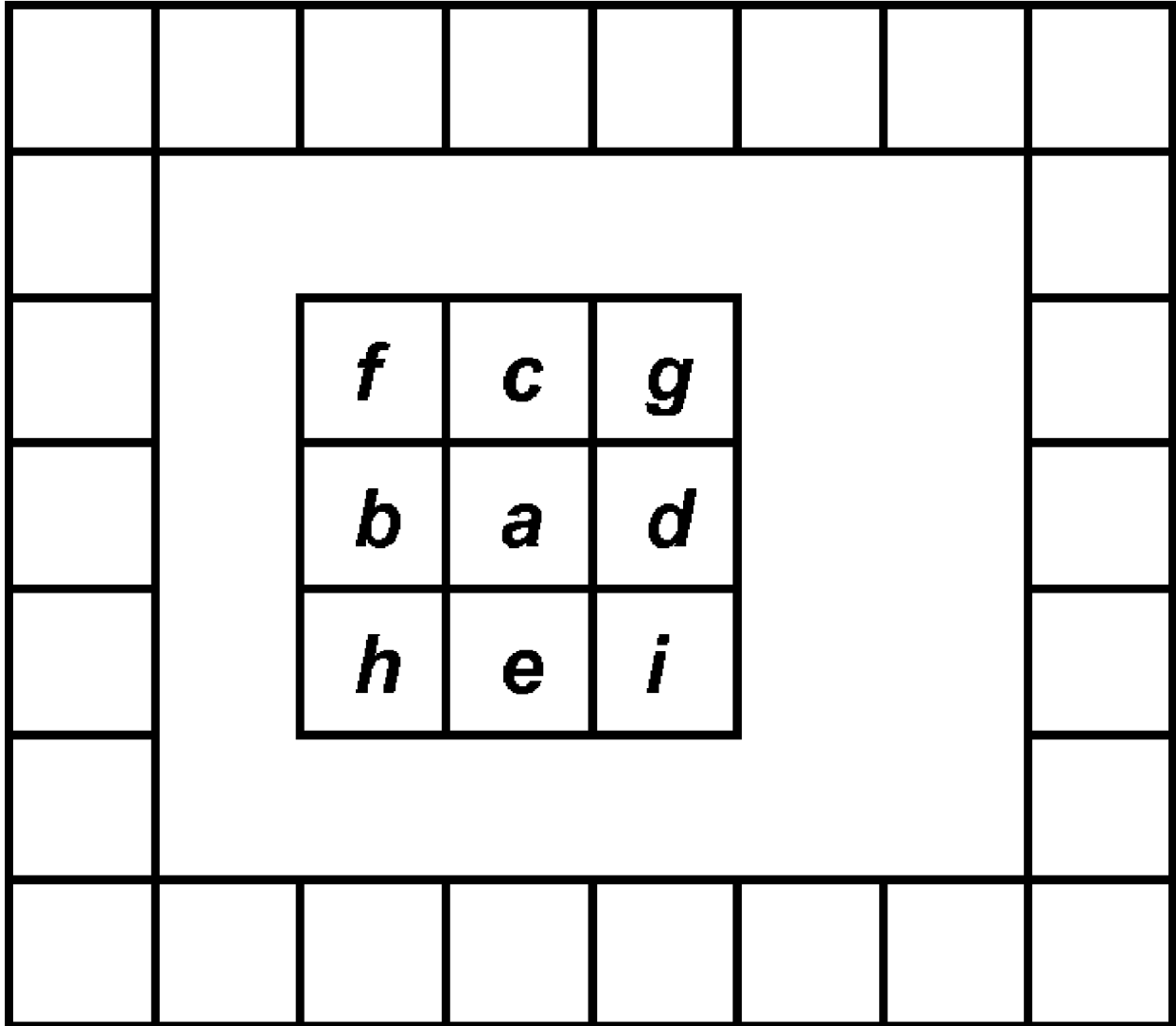


图 2B

# 图片帧N

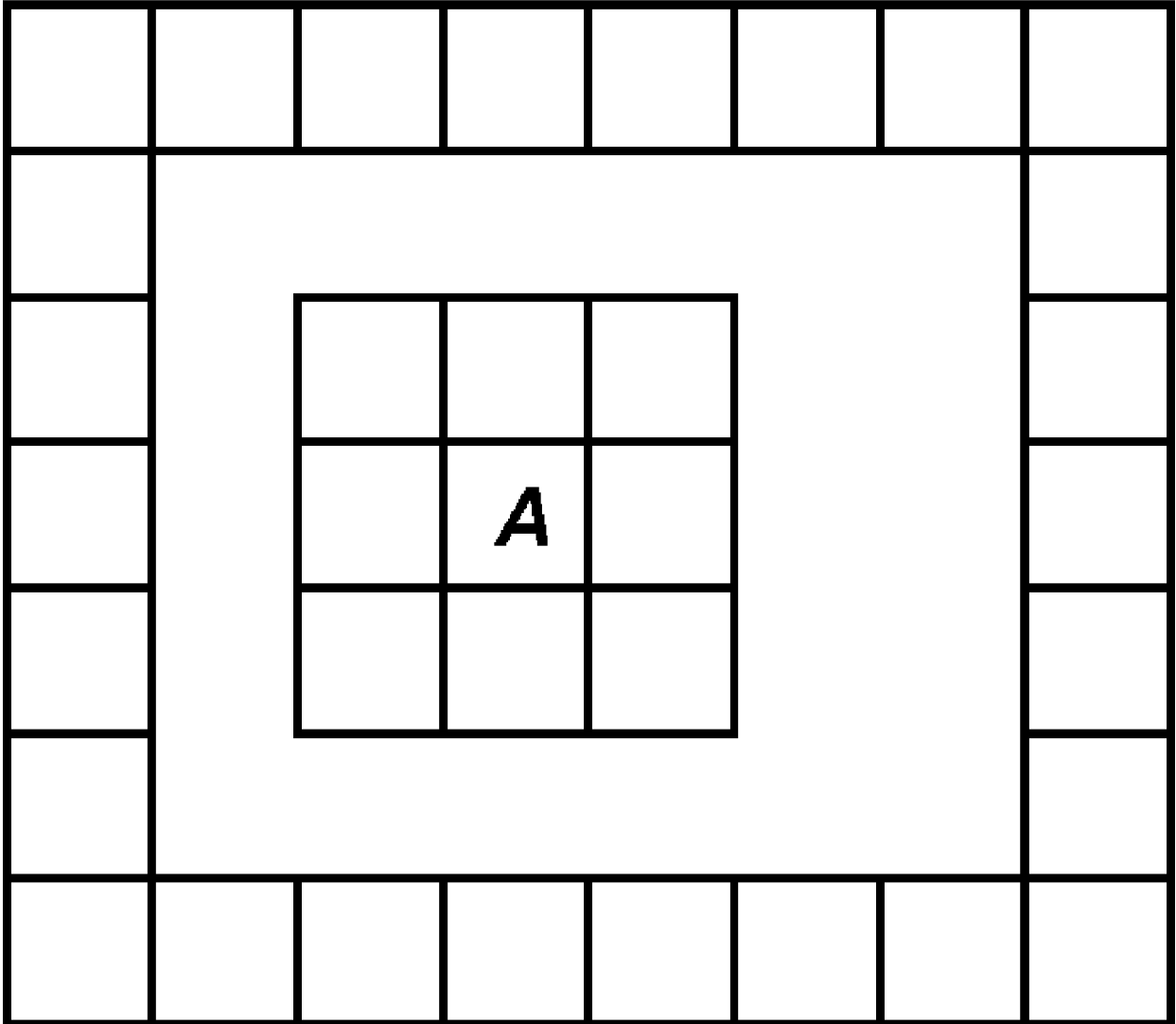


图 2C

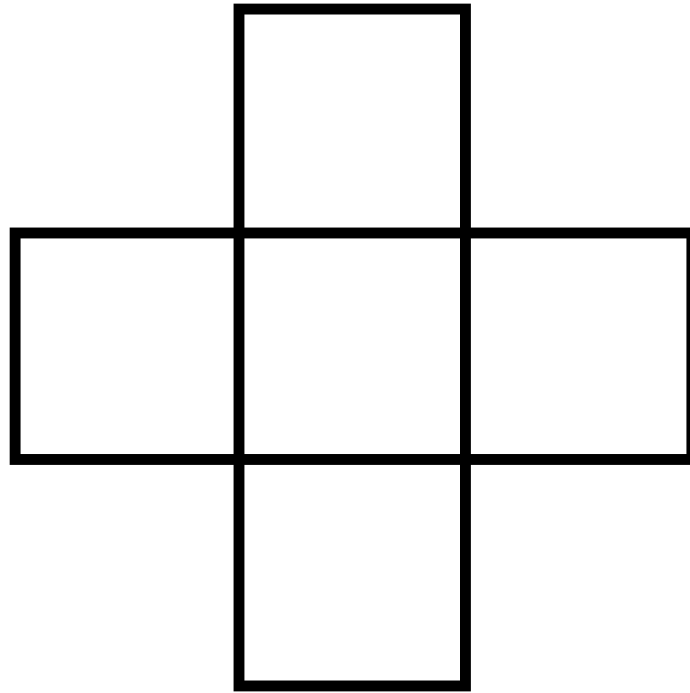


图 3A

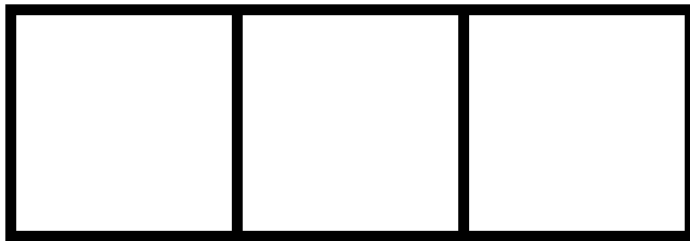


图 3B

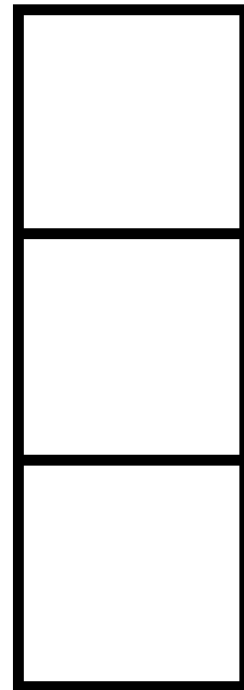


图 3C



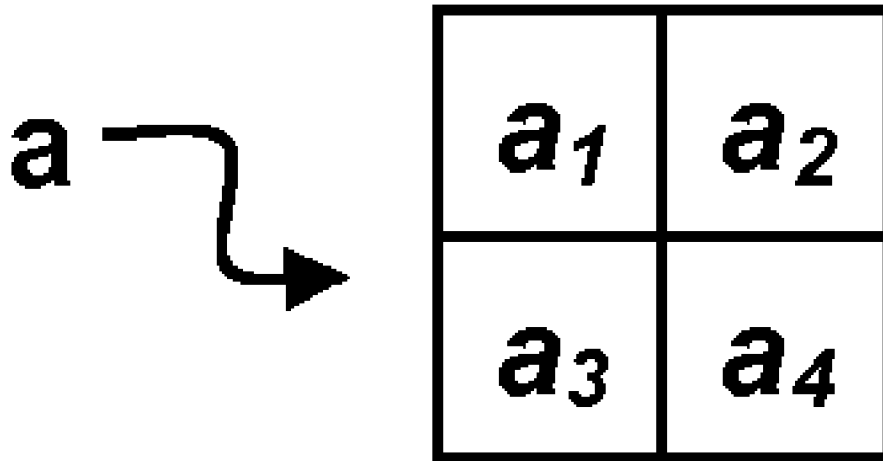


图 4A

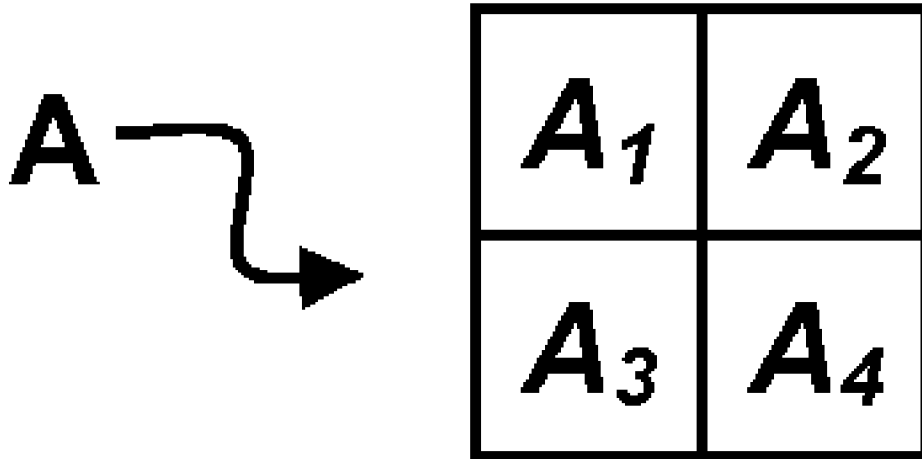


图 4B

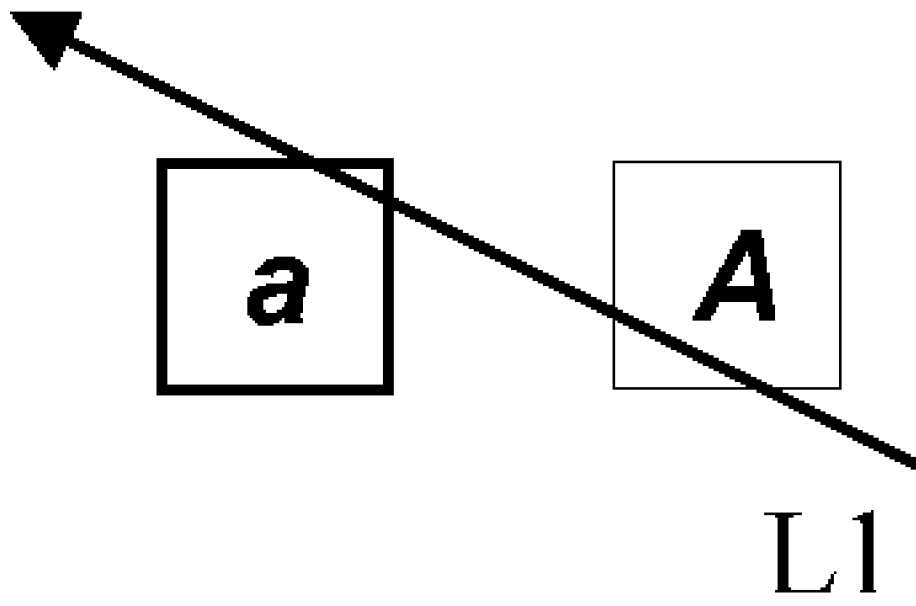


图 5A

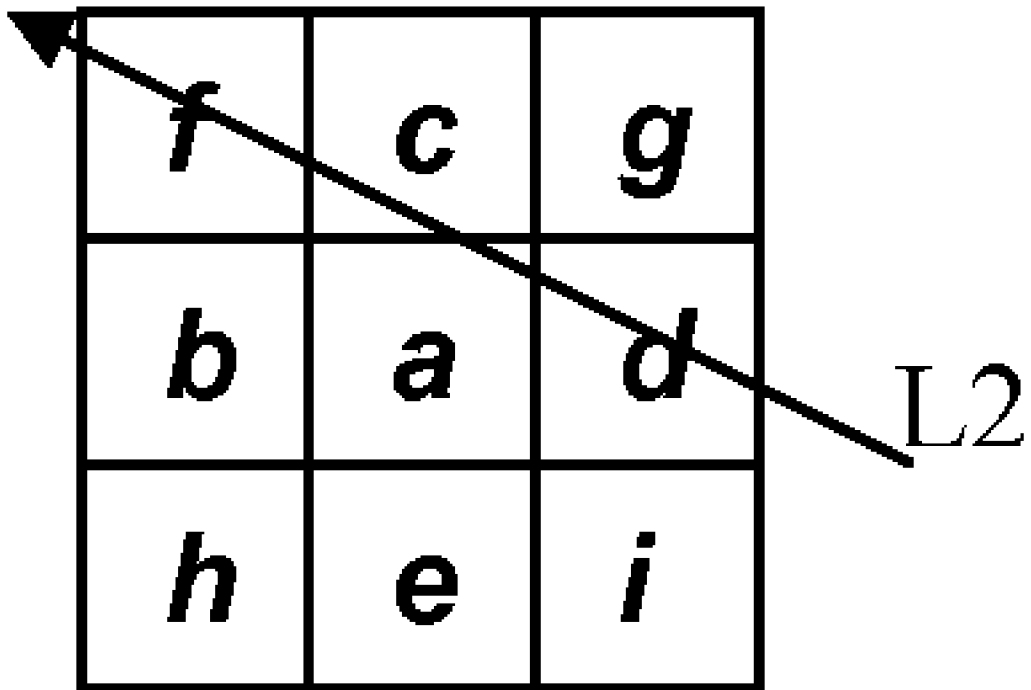


图 5B

**600**

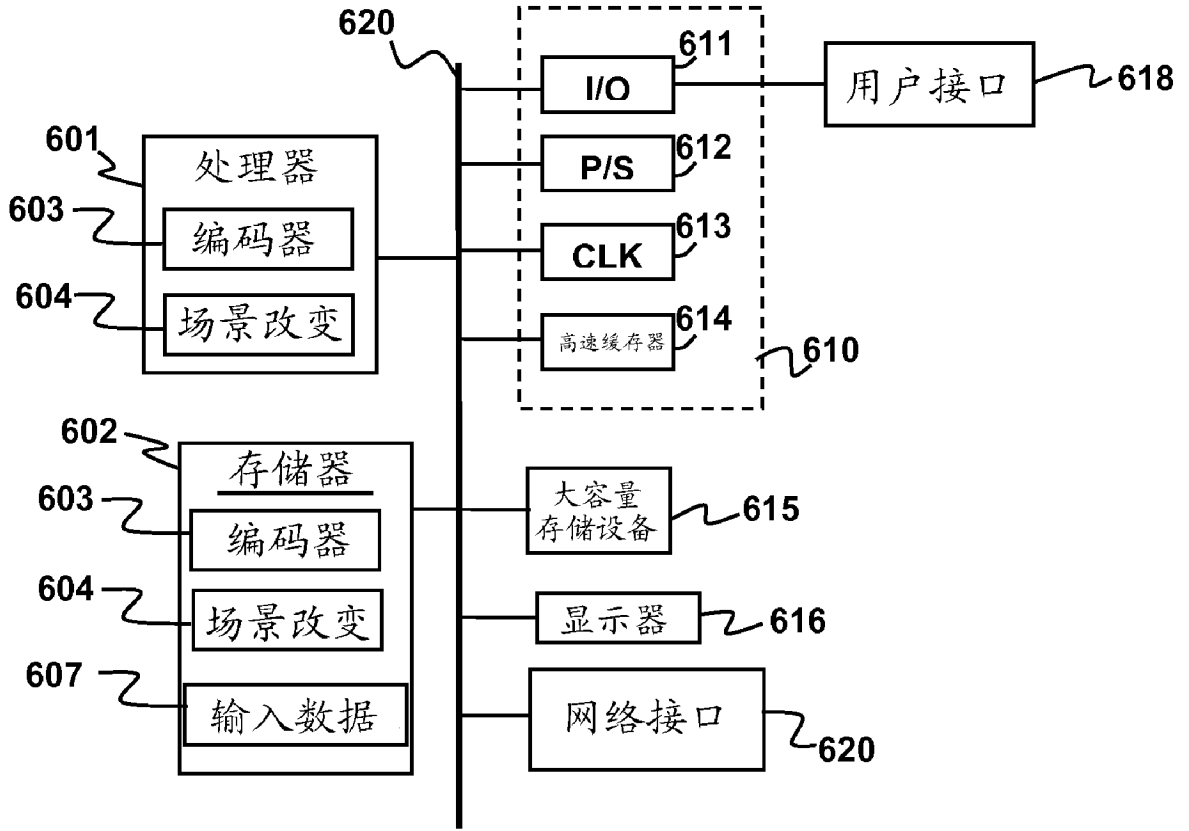


图 6

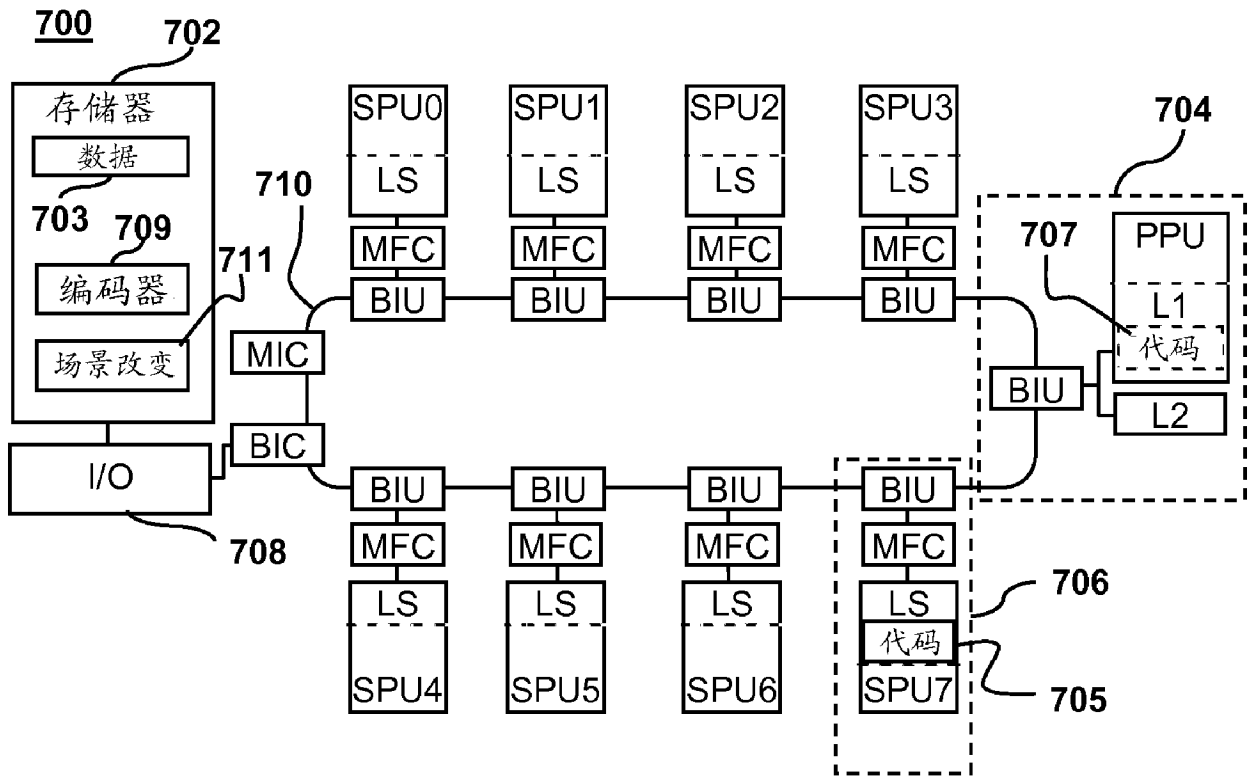


图 7

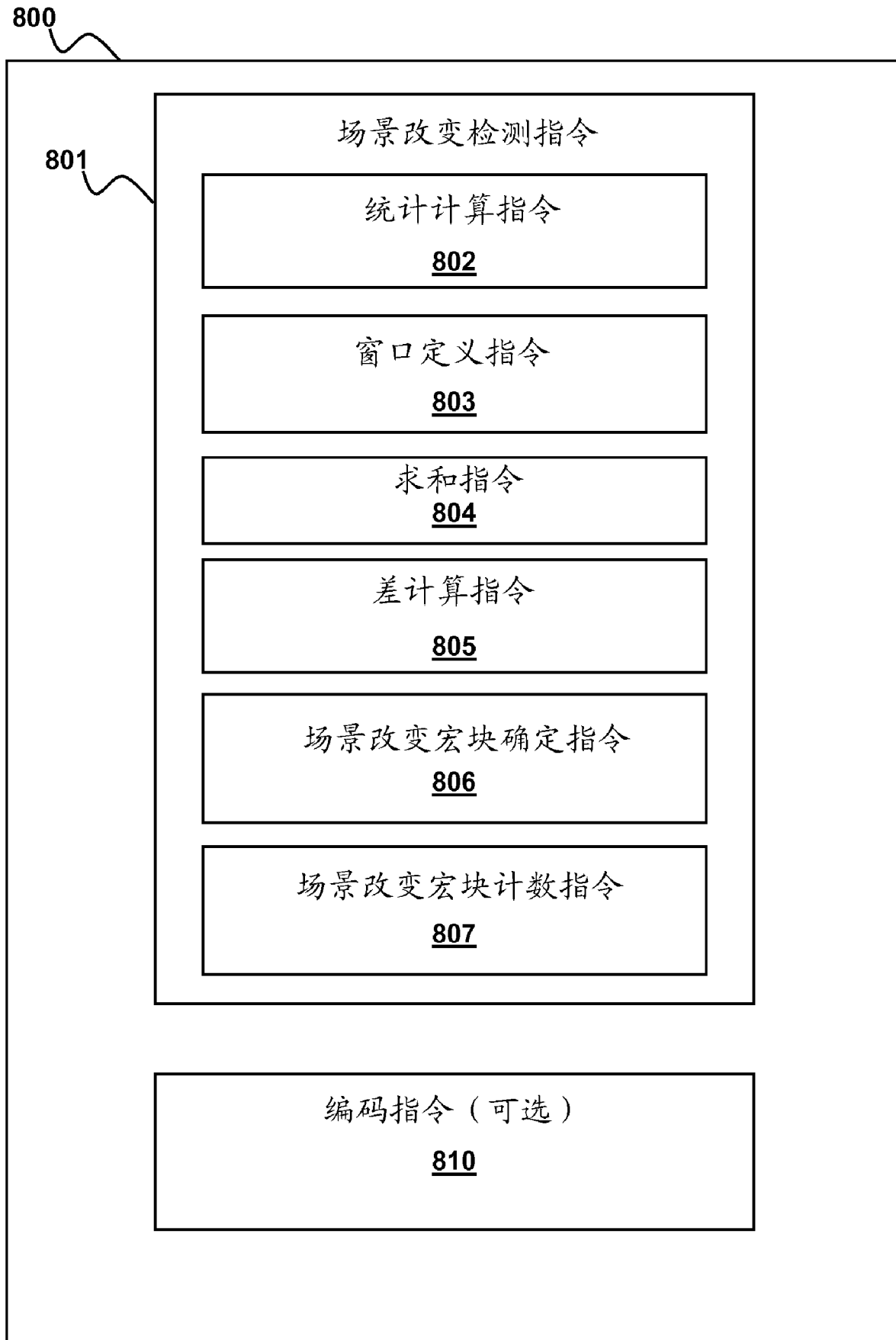


图 8