



[12] 发 明 专 利 说 明 书

[21] ZL 专利号 97118915.3

[43] 授权公告日 2003 年 1 月 22 日

[11] 授权公告号 CN 1099801C

[22] 申请日 1997.9.26 [21] 申请号 97118915.3
[30] 优先权
[32] 1996. 9.27 [33] US [31] 08/722,468
[71] 专利权人 通用仪器公司
地址 美国宾夕法尼亚州
[72] 发明人 陈学敏 刘承德
审查员 金源

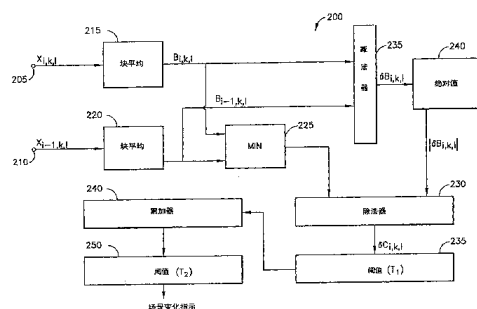
[74] 专利代理机构 永新专利商标代理有限公司
代理人 蹇 炜

权利要求书 3 页 说明书 9 页 附图 2 页

[54] 发明名称 数字视频信号的场景变化检测器

[57] 摘要

检测图象序列的前一视频图象和当前视频图象间场景变化的方法。确定前一和当前视频图象块的平均亮度直。用平均亮度值的差值确定增量视觉检测值。若增量视觉检测值超过块对比阈值水平表示有场景变化。若当前和前一图象块平均亮度值的最小值超过暗场景阈值，则用平均亮度值差值的绝对值和当前和前一图象块的平均亮度值的最小值之比确定增量视觉检测值。否则，用差值的绝对值和暗场景阈值之比确定增量视觉检测值。



1. 一种检测前一视频图象和当前视频图象之间场景变化的方法，包括下列步骤：

确定所述前一和当前视频图象的块对的平均亮度值；

用所述平均亮度值之间的差值来确定增量视觉检测值；

确定所述当前图象块的平均亮度值和前一图象块的所述平均亮度值中的最小值，

如果所述最小值超过暗场景阈值，则用（a）所述差值的绝对值，和（b）所述最小值之比来确定所述增量视觉检测值；

如果所述最小值不超过暗场景阈值，用（a）所述差值的绝对值，和（b）所述暗场景阈值之比来确定所述增量视觉检测值；及

如果所述增量视觉检测值超过块对比阈值水平，则表示有场景变化。

2. 如权利要求 1 的方法，其中所述块对比阈值水平为韦伯分数常数的十五到二十五倍。

3. 如权利要求 1 的方法，其中所述块对的所述块在所述前一和当前图象中分别位于同一相对位置。

4. 如权利要求 1 的方法，其中：

所述暗场景阈值为最大灰度级的 10%。

5. 如前面权利要求之一的方法，其中：

确定所述前一和当前视频图象的多个块对的平均亮度值之间的所述差值；以及

用所述差值来确定每个所述块对的所述增量视觉检测值；其中：

如果对于所述当前和前一视频图象中块对的阈值比例，所述增量视觉检测值超过块对比阈值水平，则表示有场景变化。

6. 如权利要求 5 的方法，其中所述阈值比例为 80%到 90%。

7. 如权利要求 1 的方法，还包括下列步骤：

确定所述块对的所述块之间的运动的相对量；以及

根据所述运动相对量来调节所述块的大小。

8. 如权利要求 7 的方法，其中当所述运动相对量增加时，增加所述块的大小。

9. 如权利要求 7 的方法, 其中所述确定运动相对量的步骤还包括下列步骤:
确定水平运动矢量的绝对值与垂直运动矢量的绝对值的和数;
其中所述水平与垂直运动矢量分别表示所述当前图象块的视频图象相对于所述前一图象块的视频图象的水平和垂直运动; 以及
确定所述和数是否超过运动阈值。
10. 如权利要求 9 的方法, 其中:
根据所述当前图象的图象类型来调节所述运动阈值。
11. 一种检测前一视频图象和当前视频图象之间场景变化的装置, 包括:
用于确定所述前一和当前视频图象块对的平均亮度值的设备;
用来用所述平均亮度值之间的差值来确定增量视觉检测值的设备;
用于确定所述当前和前一图象块的所述平均亮度值的最小值的设备, 其中:
如果所述最小值超过暗场景阈值, 则用 (a) 所述差值的绝对值, 和 (b) 所述最小值之比来确定所述增量视觉检测值;
如果所述最小值不超过暗场景阈值, 则用 (a) 所述差值的绝对值, 和 (b) 所述暗场景阈值之比来确定所述增量视觉检测值, 其中:
如果所述增量视觉检测值超过块对比阈值水平, 则表示有场景变化。
12. 如权利要求 11 的装置, 其中所述块对比阈值水平为韦伯分数常数的十五到二十五倍。
13. 如权利要求 10 或 11 的装置, 还包括:
用于确定所述前一和当前视频图象的多个块对的平均亮度值之间的所述差值的设备; 以及
用所述差值来确定每个所述块对的所述增量视觉检测值的设备; 其中:
如果对于所述当前和前一视频图象中块对的阈值比例, 所述增量视觉检测值超过块对比阈值水平, 则表示有场景变化。
14. 如权利要求 13 的装置, 其中所述阈值比例在 80%到 90%范围之内。
15. 如权利要求 11 的装置, 还包括:
用于确定所述块对的所述块之间运动相对量的设备; 以及
用于根据所述运动相对量来调节所述块的大小的设备。
16. 如权利要求 15 的装置, 还包括:
用于当所述运动相对量增加时, 增加所述块的大小的设备。

-
17. 如权利要求 15 的装置, 其中 用于所述确定运动相对量的设备还包括:
用于确定水平运动矢量的绝对值与垂直运动矢量的绝对值的和数的设备;
其中所述水平和垂直运动矢量分别表示所述当前图象块的视频图象相对于所述前一图象块的视频图象的水平和垂直运动; 以及
用于确定所述和数是否超过运动阈值的设备。
18. 如权利要求 17 的装置, 还包括:
用于根据所述当前图象的图象类型来调节所述运动阈值的设备。

数字视频信号的场景变化检测器

技术领域

本发明涉及一种用于检测数字图象帧（例如图象）序列中场景变化的装置和方法。特别是利用图象帧中不同块的亮度相对于前一帧中相应块的亮度的变化来提供可靠的场景变化指示。

背景技术

近来，用于给用户家庭提供电视及其它音频、视频和 / 或数据服务的数字视频传输格式越来越流行。这类节目可通过直播卫星链路，或通过有线电视网传输到用户家庭。除提供新的和扩大的节目源外，数字视频可比传统的模拟电视广播提供更高的图象质量。

为在可利用的带宽内传输数字视频信号，需使用数据压缩技术。特别是，空间压缩技术利用一帧中相邻像素或像素块之间的相关性，而时间技术则利用连续图象帧中像素或像素块之间的相关性。此外，运动补偿技术可提供更大的时间压缩。

然而，为提供最佳的视频帧序列的压缩，需要有检测该序列中场景变化的能力。场景变化通常可定义为帧的组成部分的任何显著变化。例如，在电影的视频序列中，当摄影机的角度改变时，或当在特写和全景之间进行切换时，会发生场景变化。此外，场景变化常常是通过连续视频帧之间亮度级的明显改变表现出来的。例如，当在暗室中打开亮的灯时，会发生场景变化。

一旦检测到场景变化，就可以相应地修改视频序列的编码。例如，当检测到场景变化时，由于新的场景的第一帧和前一场景最后一帧之间时间相关的可能性降低，可暂时中止运动补偿。另外，可根据场景变化信息选择图象的具体类型（如 I，P，或 B 图象）。I，P，或 B 图象是根据如下面将更详细讨论的 MPEG-2 标准来定义的。

各种现有的场景检测系统试图提供可靠的场景变化检测。例如，一种系统使

用在当前帧和前一帧之间相应像素值之差的绝对值的和数，并将此和数与预定的常数阈值比较以确定是否有场景变化。然而，如果在连续帧中发生快速运动，该系统可能不能给出可靠的结果。另外，当存在不同水平的运动（例如中速，快速，高速）时，这种系统也并非始终都是可靠的。

另一种系统确定当前帧和前一帧之间相应像素值之差的绝对值。然后确定当前帧和下一帧之间相应像素值之差的绝对值。最后得到上述两差值的差值和数，并将其与预定的常数阈值比较以确定是否有场景变化。然而，这一系统当场景从没有运动转变为快速运动时，可能不能提供可靠的结果。

因此，需要提供一种用于数字视频的场景检测系统，当存在不同水平的运动时，以及当场景从没有运动转变为快速运动时，该系统能够可靠地检测场景变化。而且，该系统应与现有的包括运动图象专家组（MPEG）标准的数字视频编码标准兼容，该标准的详细情况可见这里作为参考资料编入的文件：ISO/IEC JTC1/SC29/WG11 N0702，标题为“信息技术—运动图象及有关音频的遗传编码，建议 H.262，” March 25, 1994。

MPEG 系统包括三种类型的视频图象；它们分别是帧内编码图象（I—图象），预测编码图象（P—图象）及双向预测编码图象（B—图象）。而且，帧或场结构视频序列都是适和的。I—图象完整地描述单个视频图象，而不参考任何其它图象。为了提高隐错能力，在 I—图象中可包括运动矢量。由于在底层（base layer）中 P—图象和 B—图象都是从 I—图象预测的，因此 I—图象中的错误有对所显示的图象产生更大影响的可能性。P 图象是根据前一 I 或 P 图象预测的。参考是从前面的 I 或 P 图象到将来的 P—图象，称为前向预测。B—图象是从最接近的后面的 I 或 P 图象及最接近的后面的 I 或 P 图象预测的。

一个好的场景检测系统还提供速率控制的自动控制和视频压缩编码器的自适应量化处理，该编码器使用各种标准，包括 MPEG-1, MPEG-2, ISO/IEC H.261（视频会议）及 ISO/IEC H.263。此外，该系统还应与各种彩色电视广播标准兼容，如在北美使用的国家电视标准委员会（NTSC）标准，在欧洲使用的逐行倒相（PAL）标准，并且该系统还应和帧和场模式的视频兼容。本发明提供具有上述和其它优点的系统。

发明内容

根据本发明，提供一种检测视频图象序列中前一图象和当前图象之间场景变化的方法和装置。

根据本发明的检测前一视频图象和当前视频图象之间场景变化的方法，包括下列步骤：确定所述前一和当前视频图象的块对的平均亮度值；用所述平均亮度值之间的差值来确定增量视觉检测值；确定所述当前图象块的平均亮度值和前一图象块的所述平均亮度值中的最小值，如果所述最小值超过暗场景阈值，则用（a）所述差值的绝对值，和（b）所述最小值之比来确定所述增量视觉检测值；如果所述最小值不超过暗场景阈值，用（a）所述差值的绝对值，和（b）所述暗场景阈值之比来确定所述增量视觉检测值；及如果所述增量视觉检测值超过块对比阈值水平，则表示有场景变化。

根据本发明的检测前一视频图象和当前视频图象之间场景变化的装置，包括：用于确定所述前一和当前视频图象块对的平均亮度值的设备；用来用所述平均亮度值之间的差值来确定增量视觉检测值的设备；用于确定所述当前和前一图象块的所述平均亮度值的最小值的设备，其中：如果所述最小值超过暗场景阈值，则用（a）所述差值的绝对值，和（b）所述最小值之比来确定所述增量视觉检测值；如果所述最小值不超过暗场景阈值，则用（a）所述差值的绝对值，和（b）所述暗场景阈值之比来确定所述增量视觉检测值，其中：如果所述增量视觉检测值超过块对比阈值水平，则表示有场景变化。

特别是，确定当前和前一图象块的平均亮度值的最小值，其中，如果这一最小值超过暗场景阈值，则用（a）平均亮度值之间差值的绝对值，和（b）当前和前一图象块的平均亮度值的最小值之比来确定增量视觉检测值。否则，用（a）差值的绝对值，和（b）暗场景阈值之比来确定增量视觉检测值。暗场景阈值可约为最大灰度级的 10%。

此外，可确定前一和当前视频图象多个块对的平均亮度值之间的差值。最好用图象中的每个块对来确定总的图象场景变化。用差值来确定每个块对的增量视觉检测值，其中，如果增量视觉检测值超过块对比阈值水平当前和前一视频图象的块对的阈值比例，则表示有场景变化。这一阈值比例可约为 80%到 90%。

另外，通过确定块对的块之间的运动相对量，并根据该运动相对量来调节块的大小，该方法可自适应地优化。特别是，当运动相对量增加时，增加块的大小。此外，运动相对量可通过确定水平运动矢量的绝对值和垂直运动矢量的绝对值的

和数来获得, 其中, 水平和垂直运动矢量分别表示当前图象块的视频图象相对于前一图象块的视频图象的水平和垂直运动。然后确定该和数是否超过了运动阈值。运动阈值可根据当前图象的图象类型(例如, 当前图象是否是 I, P 或 B 图象)来调节。

附图说明

图 1 示出根据本发明两相继视频帧的块之间的比较。

图 2 是根据本发明的场景变化检测器的方框图。

具体实施方式

本发明提供一种检测数字视频帧序列中场景变化的方法和装置。

场景的辉度级由包括该场景的各象素的平均亮度和各象素亮度值的动态范围决定。此外, 人眼对辉度的视觉检测通常被认为是图象亮度的自然对数的函数。在场景变化的帧和 / 或场中, 对辉度的视觉检测随前一帧或前一相应场而明显改变。

而且, 由于人对亮度的对比比对绝对亮度值本身更敏感, 因此两场景之间的增量视觉检测 δC 是场景变化的一个好的标志。 δC 定义为在两帧和 / 或场中, 具有同一相对位置的区域(例如块)的平均辉度之差值。

特别是, 根据韦伯定理, 如果一物体的亮度 b_0 与周围区域的亮度 b_s 之间的差别正好是可察觉的, 则可定义下面被称为韦伯分数的比值:

$$\frac{|b_s - b_0|}{b_0} = C = \text{常数}$$

在高的背景亮度级, 例如大于 0.5mL (毫流明) 时, 韦伯分数基本保持为常数。已发现该常数 C 的值为 0.02, 这意味着在 0 到 1 的刻度上, 对于人可察觉的级之间的对比, 至少需要五十个不同的亮度级。

令 $b_0 = b$, 则有 $b_s = b + \delta b$, 其中 δb 是最小可察觉的亮度变化。这样,

$$\frac{\delta b}{b} \approx d(\log_e b) = \delta C \quad (\text{常数})$$

这显示 $\frac{|b_s - b_0|}{b_0}$ 是与辉度增量视觉检测值成比例的。

图 1 示出根据本发明两相继视频帧的块之间的比较。当前帧，即帧(i)，在 100 上示出，它包括块 110。前一帧，即帧(i-1)，在 150 上示出，它包括块 160，该块在帧 150 中位于与帧 100 中的块 110 同一相对位置上。

例如，对于 NTSC 格式，帧 100 和 150 可各包括三十片(slice)，每一片有四十四个宏块。这样，整个 NTSC 帧包括 1,320 个宏块。而且，宏块一般包括一个 16x16 的像素块，在 MPEG-2 标准中，该块例如是由四个 8x8 的像素块构成的。从而，一个 NTSC 帧可能包括在宽度上为 44x16=704 个像素，在高度上为 30x16=480 个像素，总数为 337,920 个像素。此外，本发明与 PAL 格式是兼容的，该格式包括在 36 片中的 1,584 个宏块，每片有 44 个模块，每个宏块有 16x16 个像素。

块 110 和 160 用坐标组(k,l)表示，其中 k 是块的水平索引，l 是垂直索引。此外，每个块 110 和 160 的大小可为例如在高度上为 16 个像素而宽度上为 32 个像素。在这种情况下，k 的范围从 1 到 704/32=22，l 的范围从 1 到 480/16=30。下面各项定义为：

h	帧高度（像素）
w	帧宽度（像素）
m	块高度（像素）
n	块宽度（像素）
i	帧索引
k	水平块索引(k=1, ..., h/m)
l	垂直块索引(l=1, ..., w/n)
$X_{i,k,l}$	第 i 个帧，第 k 个水平块，第 l 个垂直块的像素密度

这样，我们有了由像素组定义的两个相继帧和 / 或两个顶(top)（或底(bottom)）场。特别是，第(i)帧，即帧 100 由像素组 $X_{i,k,l}$ 来定义，第(i-1)帧，即帧 150 由像素组 $X_{i-1,k,l}$ 来定义。为有效地分辨场景变化，将每帧划分为一组 k x l 的单独的块，每块有 m x n 个像素。

注意块的大小可设计为根据当前运动信息自适应地改变。特别是运动越快，则块大小 m x n 应越大。根据运动量对于每一帧调节块大小的一种方法是执行下列步骤。首先，对于每个 16x16 的宏块，计算索引 v[x][y]，其中 x=1,2, ..., [w/16]，y=1,2, ..., [h/16]。如果全像素前向运动矢量，即 vector[x][y][z]满足下列不等式：

$$|\text{vector}[x][y][0]| + |\text{vector}[x][y][1]| > T_3,$$

则表示两块之间有快速运动。 $\text{vector}[x][y][0]$ 和 $\text{vector}[x][y][1]$ 分别为当前帧块（例如块(x,y)）相对于前一帧块的水平和垂直运动矢量。这样，如果不等式成立，则令索引 $v[x][y]=1$ ；否则令 $v[x][y]=0$ 。

注意运动矢量 $\text{vector}[x][y][z]$ 是从具有同样图象类型的最近的可利用的图象获得的。例如，如果当前图象类型是P-图象，则运动矢量 $\text{vector}[x][y][z]$ 是以前的预测P-图象的运动矢量。由于对于每个图象的场景变化检测发生在图象的运动估值之前，因此这是正确的。

阈值 T_3 是根据视频帧序列中出现的不同图象类型来选择的。例如，如果比特流中没有B-图象，例如序列I,P,P,...，则 $T_3=16$ 是合适的选择。如果出现一个B-图象，例如序列P,B,P,B,...，则在当前图象是B-图象时， $T_3=16$ 是合适的选择，在当前图象是P-图象时， $T_3=32$ 是合适的选择，如此等等。

其次，相应地调节块大小。可使用初始（缺省）的 16×16 的块大小。然后，可根据 $v[x][y]$ 调节块大小。例如，如果 $v[x][y]=1$ ，则可增加块大小，例如增至 16×32 或 32×32 。类似地，如果 $v[x][y]=0$ ，则可减小块大小，例如减至 8×16 。但注意块大小的增加不能使得块超过宏块的右和/或底边界。通常，当运动越快时，块大小应越大。而且，在某些应用中，最大允许的块大小可能被限制在 32×48 个像素。

接下来，确定在帧（或顶场）中每块的平均亮度。对于第(i)帧，块(k,l)，平均亮度为：

$$B_{i,k,l} = \frac{1}{mn} \sum_{c1=0}^{m-1} \sum_{c2=0}^{n-1} X_{i,kh+c1,lw+c2}$$

$k=1, \dots, h/m$, $l=1, \dots, w/n$ 。c1和c2为虚设的计数索引。之后，第(i)和第(i-1)帧（或顶场）之间的块亮度增量 $\delta B_{i,k,l}$ 由下式确定：

$$\delta B_{i,k,l} = B_{i,k,l} - B_{i-1,k,l}$$

此外，对于第(i)帧，块(k,l)，可定义相对块增量对比 $\delta C_{i,k,l}$ 如下：

$$\delta C_{i,k,l} = \begin{cases} \frac{|\delta B_{i,k,l}|}{\min\{B_{i,k,l}, B_{i-1,k,l}\}}, & \text{对于 } \min\{B_{i,k,l}, B_{i-1,k,l}\} > T_0; \\ \frac{|\delta B_{i,k,l}|}{T_0}, & \text{对于 } \min\{B_{i,k,l}, B_{i-1,k,l}\} \leq T_0. \end{cases}$$

T_0 是表示场景被认为是暗场景的阈值。通常, 可使用 $T_0=25.5$, 该值是最大灰度级 255 的 10%。

对于场景变化, 将块亮度对比的相对变化的有效阈值 T_1 设置为: $T_1=15-25\text{ C}$ 。现在, 考虑索引阵列, 对于 $k=1,\dots,h/m$, $l=1,\dots,w/n$, $\text{index}[k][l]$ 定义为:

$$\text{index}[k][l] = \begin{cases} 1, & \text{如果 } \delta C_{i,k,l} > T_1 \\ 0, & \text{否则} \end{cases}$$

那么, 如果一帧中块的约 80%-90% 具有比有效阈值大的相对块增量对比, 即

$$\frac{\sum_{k=1}^{\lfloor h/m \rfloor} \sum_{l=1}^{\lfloor w/n \rfloor} \text{index}[k][l]}{\lfloor h/m \rfloor \cdot \lfloor w/n \rfloor} > T_2, \quad \text{其中 } T_2=80\%-90\%$$

则根据本发明, 这表示有场景变化。80%-90% 的范围是根据大范围的测试确定的, 但实际的最佳值可随具体的场景而改变。注意数学表示式 $\lfloor x \rfloor$ 代表将非整数 x 舍入成下一个最小的整数。

图 2 是根据本发明的场景变化检测器的方框图。检测器, 在 200 上示出, 包括输入端 205 和 210。在输入端 205, 接收来自当前帧的像素数据 $X_{i,k,l}$, 并将其提供给块平均函数 215 以为第 i 帧中的每一块产生平均亮度值 $B_{i,k,l}$ 。类似地, 在输入端 210, 接收来自前一帧的像素数据 $X_{i-1,k,l}$, 并将其提供给块平均函数 220 以为第 $(i-1)$ 帧中的每一块产生平均亮度值 $B_{i-1,k,l}$ 。

最小化单元(MIN)225 确定 $\min\{B_{i,k,l}, B_{i-1,k,l}\}$ 并将该项输出到除法器(divisor)函数 230 中。同时, 减法器 235 确定 $\delta B_{i,k,l} = B_{i,k,l} - B_{i-1,k,l}$ 。绝对值函数 240 确定 $|\delta B_{i,k,l}|$, 并将该项提供给除法器 230。除法器 230 根据是否 $\min\{B_{i,k,l}, B_{i-1,k,l}\} > T_0$ 为第 (i) 帧, 块 (k,l) 确定相对块增量对比 $\delta C_{i,k,l}$ 。然后将 $\delta C_{i,k,l}$ 提供给阈值函数 235, 该函数 235 确定块 (k,l) 是否表示有场景变化 (例如是否 $\delta C_{i,k,l} > T_1$)。如果是, 可相应地设定 $\text{index}[k][l]$ 。累加器 240 累加每块的场景变化结果, 并对整个帧或其一部分的结果求和。最后, 阈值函数 250 接收来自累加器 240 的相加的结果, 并用阈值 T_2 确定是否表示整个帧有场景变化。

用不同的视频序列对本发明的场景变化检测系统进行了大范围的测试。特别是对在测试模式编辑委员会 “Test Model 5”, ISO/IEC JTC1/SC29/WG11

MPEG93/457, April 1993 中描述的“Football”, “Mobile Calendar”, “Flower Garden”及“Table tennis”视频序列以及在关于 MPEG-4 视频 VM 编辑的特设组, “MPEG-4 视频确认模型版本 3.0”, ISO/IEC JTC1/SC29/WG11 N1277, Tampere, Finland, July 1996 中描述的“Destruct”, “Street Organ”, “Silent”及“Fun Fair”视频序列进行了分析。

本发明场景检测系统的样本测试结果示于下面的表 1 中。阈值选择为 $T_1=0.3$ 和 $T_2=0.85$, 块大小为 $m=16, n=32$ 。具体的视频序列在第一栏中给出。所涉及的帧序列示于第二栏中。例如, [0:50]表示分析了 0 到 50 帧的场景变化。第三栏表示是否检测到场景变化, 且如果是, 是在哪一帧。例如, 第三帧中的场景变化意味着在第二和第三帧之间检测到场景变化。第四栏, 如果可使用, 提供有关视频序列特性的额外信息。

表 1

序列	帧	场景变化	备注
Football	[0:50]	无	快速运动
Mobile Calendar	[0:44]	无	
Street Organ	[0:50]	无	
Silent	[0:50]	无	没有运动到运动
Flower Garden	[0:30]	无	摄像机移动
Fun Fair	[0:50]	无	快速运动
Table tennis	[90:100]	97	
Destruct	[0:40]	25,26	亮灯
任意两个序列的结合		有	每次均检测到场景变化

此外, 对于在上面视频序列中检测到场景变化的帧, 使用 MPEG-2 WG-11 方案检查编码效率。通过确定由达到所给图象质量的编码方法产生的比特数来确定编码效率。特别地, 对于一恒定的质量水平, 较少的编码比特表示较高的编码效率。如果有根据本发明检测的场景变化的帧被编码为 P-图象类型, 则多于 46% 的宏块被编码为 I- 图象。于是, 发现本发明的场景检测系统如所预期的那样工作, 因为它成功地将用预测编码技术难于有效地编码的帧定位。由于编码器的速率控制工具给 I-图象分配更多的比特, 这样的帧可被方便地编码为 I-图象, 从而也改善了图象质量。

通常, 在典型的编码器中, 如使用 Test Model 5 标准的 MPEG-2 编码器, 对于

P-图象或 B-图象的每一宏块有一个选择宏块编码类型的过程。这一过程将决定是应将宏块编码为帧内编码块(I-编码)还是非帧内编码块(P-或 B-编码),这要基于哪种编码类型可提供更好的质量或使用更少的数据比特。对于 P-图象或 B-图象,如果预测是有效的,则图象中只有小比例的宏块将为 I-编码(例如少于百分之五)。这是必要的,因为 I-编码块由于没有时间压缩而要使用相对来说大量的数据比特。

如果图象中 I-编码的宏块比例大于例如百分之三十或四十,图象质量就会不好。在这种情况下,在场景变化处,如所与预期的一样,预测编码对于该图象是无效的。这样,当发生场景变化时,通常需要新场景的第一帧不被编码为 P-图象。

虽然结合各个特定的实施例来描述了本发明,但本领域技术人员懂得对其可做多种修改和变化而不脱离如在权利要求中提出的本发明的实质和范围。例如,在此提出的各阈值水平可根据所分析的具体场景或视频序列来调节。即,有些类型的视频序列,如动作电影,其特点可能是有更频繁和显著的场景变化活动。而且,特定的照明情况可能与具体的视频序列有关,例如恐怖电影,其中相对来说低的照明水平可能贯穿于整个序列中。在这种情况下,可相应调节场景变化检测阈值。

此外,可能需要仅对视频图象的一部分进行分析以确定场景变化,或者可用不同的阈值分析不同的部分。例如,在具有在图象底部的相对来说暗的地面及图象顶部的相对来说亮的天空的风景场景的视频序列中,对于图象的底部,可使用更灵敏的场景变化阈值。类似地,在图象的不同区域可使用不同大小的块。例如,当向图象中部的运动比向图象边缘的运动更普遍时,可在图象的中部使用更大的块。

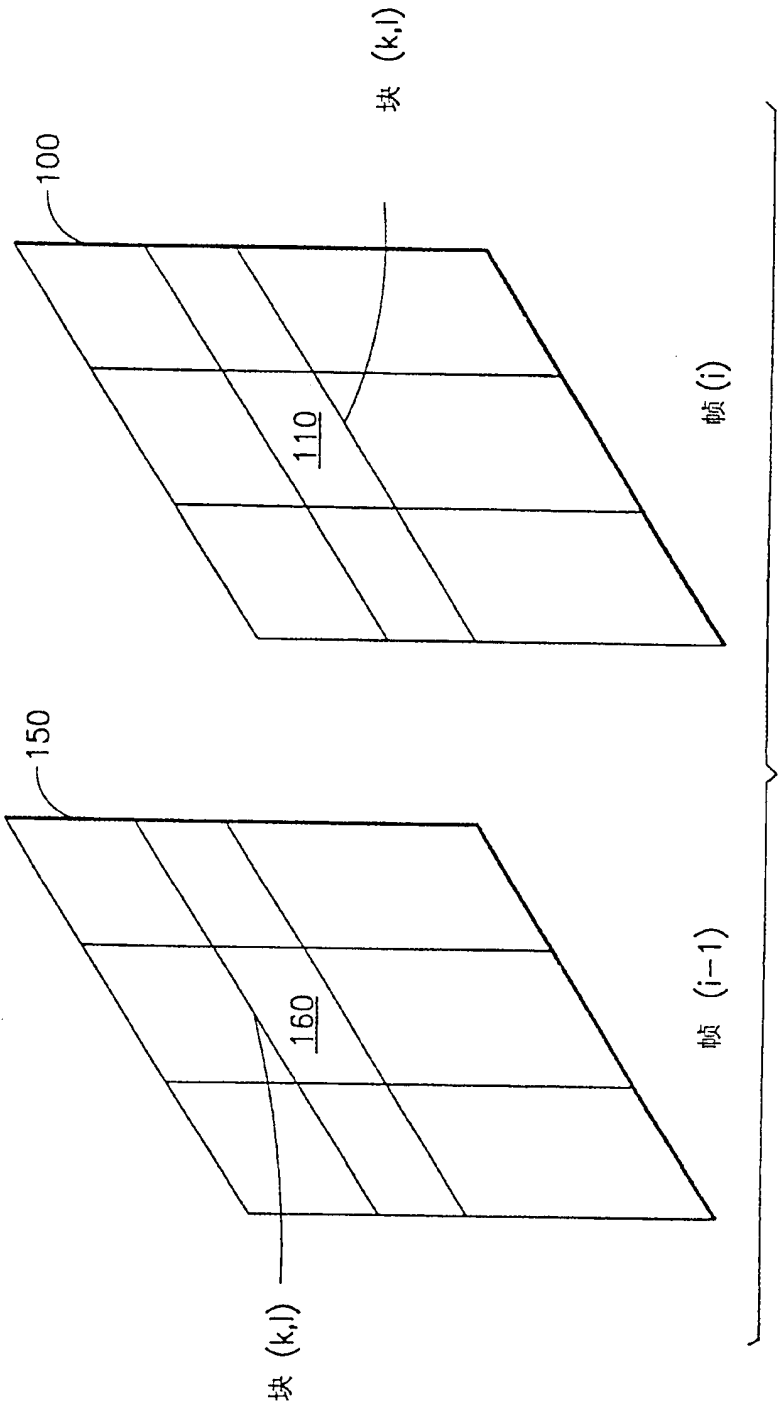


图1

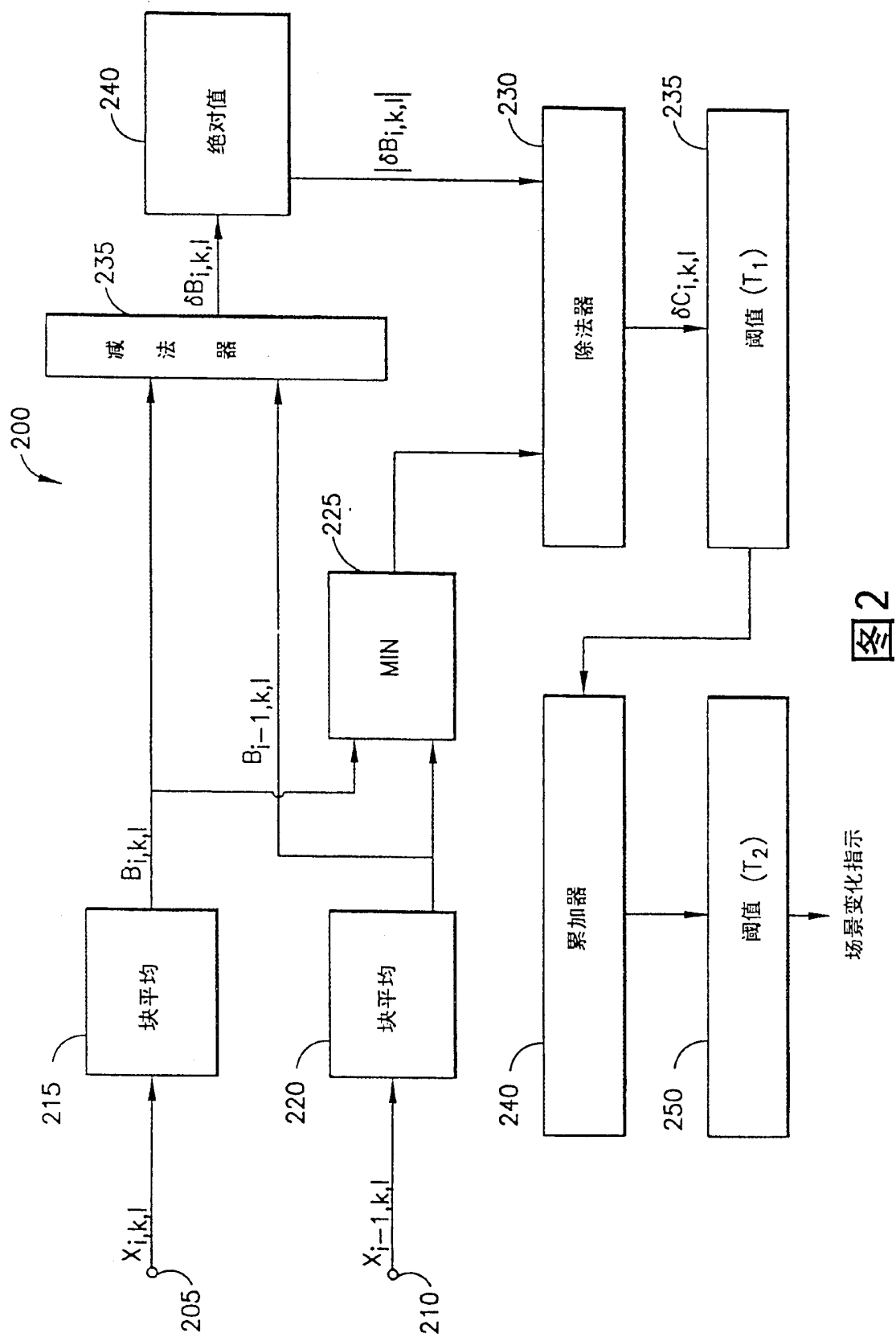


图2