



(12) 发明专利

(10) 授权公告号 CN 101090491 B

(45) 授权公告日 2016. 05. 18

(21) 申请号 200610064777. X

US 2003/0156646 A1, 2003. 08. 21,

(22) 申请日 2006. 12. 15

US 2003/0072374 A1, 2003. 04. 17,

(30) 优先权数据

审查员 吴黄飞

11/455, 009 2006. 06. 16 US

(73) 专利权人 香港科技大学

地址 中国香港九龙清水湾

(72) 发明人 区子廉 黄海明

(74) 专利代理机构 中国专利代理(香港)有限公司

司 72001

代理人 柯广华 王忠忠

(51) Int. Cl.

H04N 19/176(2014. 01)

H04N 19/51(2014. 01)

(56) 对比文件

US 2003/0072374 A1, 2003. 04. 17,

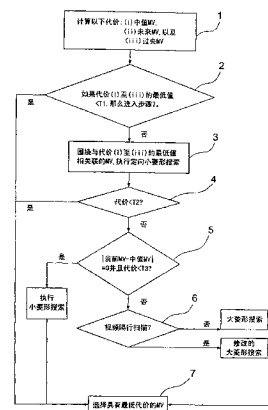
权利要求书1页 说明书15页 附图4页

(54) 发明名称

用于视频压缩的增强的基于块的运动估计算法

(57) 摘要

提出了一种方法、系统和软件,用于获得与第二图像(“参考图像”)的块相类似的第一图像的块。对第一图像的块顺序地进行处理,为每个块得到在第二图像中的多个候选位置,并且为每个候选位置评估代价函数。在第二图像中的每个候选位置是由来自第一图像的块的各自运动矢量来替换的。在本发明的第一方面中,所述代价函数是第一图像的未来块(即还没有被处理的第一图像的块)的预测运动矢量的函数。在本发明第二方面中,所述运动矢量由位置值给出,所述位置值并非全部是全像素空间、半像素空间、或1/4像素空间的。



1. 一种为第一图像中的一系列像素块中的每一个选择第二图像中的各自的相似像素块的方法,所述方法包括:

为所述第一图像的所述一系列像素块中的当前块,在所述第二图像中的多个候选位置之间进行选择,所述多个候选位置包括由两个轴方向上的两个位置值所定义的候选位置的组,所述位置值中的至少一个位置值对于所述两个轴方向的至少一个轴方向来说是不均匀的。

2. 如权利要求1所述的方法,其中,所述候选位置的组中的每一个通过相应位置值与各自的运动矢量相关联,所述运动矢量不同于预测运动矢量。

3. 如权利要求1所述的方法,其中,所述候选位置的组中的每一个通过相应位置值与各自的运动矢量相关联,所述运动矢量不同于(0,0)运动矢量。

4. 如权利要求1的方法,其中,所述位置值能够写作 DA_m/n ,其中 $m = \dots, 2, -1, 0, 1, 2, \dots$, n 是1或2的幂, D 是常量,而对于 m 的至少一个的值而言 A_m 是小于1的值,并且对于 m 的所有值而言 $0.75 \leq A_m < 1$ 。

5. 如权利要求4所述的方法,其中,所述候选位置的组中的每一个通过相应位置值与各自的运动矢量相关联,所述运动矢量不同于预测运动矢量。

6. 如权利要求4所述的方法,其中,所述候选位置的组中的每一个通过相应位置值与各自的运动矢量相关联,所述运动矢量不同于(0,0)运动矢量。

7. 如权利要求4所述的方法,其中,对于 m 的所有值而言, A_m 的值至少为0.85。

8. 一种为第一图像中的一系列像素块中的每一个选择第二图像中的各自的相似像素块的系统,所述系统包括处理器,所述处理器配置为,为所述第一图像的所述一系列像素块中的当前块,在所述第二图像中的多个候选位置之间进行选择,所述多个候选位置包括由两个轴方向上的两个位置值所定义的候选位置的组,所述位置值中的至少一个位置值对于所述两个轴方向的至少一个轴方向来说是不均匀的。

用于视频压缩的增强的基于块的运动估计算法

[0001] 相关申请

[0002] 本申请请求了于2005年6月17日申请的美国临时专利申请60/691,181的优先权,在此将其整体引入作为参考。

技术领域

[0003] 本发明通常涉及用于数字信号压缩、编码和表示的方法和系统,并且更加确切地说,本发明涉及使用多帧运动估计(ME)的方法和系统。本发明进一步涉及一种计算机程序产品,诸如记录媒体,承载可以由计算设备读取的程序指令,以使得所述计算设备执行根据本发明的一种方法。

背景技术

[0004] 由于由现代多媒体应用使用的原始数字视频数据(或者图像序列)的巨大尺寸,必须对这种数据进行压缩以便可以传输和存储这些数据。存在许多重要视频压缩标准,包括ISO/IEC MPEG-1、MPEG-2、MPEG-4标准和ITU-T H.261、H.263、H.264标准。ISO/IEC MPEG-1/2/4标准广泛地运用于娱乐业以发行电影、包括视频压缩光盘或者VCD(MPEG-1)的数字视频广播、数字视频光盘或者数字多用途光盘或者DVD(MPEG-2)、可记录DVD(MPEG-2)、数字视频广播、数字视频广播或者DVB(MPEG-2)、视频点播或者VOD(MPEG-2)、在US中的高清晰度电视或者HDTV(MPEG-2)等等。MPEG-4标准比MPEG-2更加先进,可以在较低的比特率下实现高质量视频,这使得其非常适合于因特网、数字无线网络(例如3G网络)、多媒体信息服务(来自3GPP的MMS标准)等等上的视频信息流。MPEG-4被下一代高清晰度DVD(HD-DVD)标准和MMS标准所接受。ITU-T H.261/3/4标准设计用于低延迟电视电话和视频会议系统。早期的H.261标准设计为在 $p \times 64 \text{ kbit/s}$ 下工作, $p=1,2,\dots,31$ 。后期的H.263标准非常成功,被广泛地用于现代电视会议系统,并且用于宽带网络和无线网络中的视频信息流,其中无线网络包括在2.5G和3G网络以及其他网络中的多媒体信息服务(MMS)。最新标准,H.264(也称作MPEG-4版本10,或者MPEG-4AVC)是当前最新技术水平的视频压缩标准。它如此强大以至MPEG决定与Joint Video Team(JVT)的框架中的ITU-T联合开发。新的标准在ITU-T中称为的H.264,并且被称作MPEG-4高级视频编码(MPEG-4AVC),或者MPEG-4版本10。H.264用于HD-DVD标准、直接视频广播(DVB)标准并且可能用于MMS标准。基于H.264,当前在中国正在开发称作视听标准(AVS)的相关标准。AVS1.0设计用于高清晰度电视(HDTV)。AVS-M设计用于移动应用。H.264具有超过MPEG-1/2/4及H.261/3标准的目标和主观的视频质量。除使用整数 4×4 离散余弦变换(DCT)代替传统的 8×8 DCT以外,H.264[1]的基本编码算法类似于H.263或者MPEG-4,并且还有额外的特点,包括I帧的帧间预测模式、用于运动估计/补偿的多种块大小以及多种参考坐标系、用于运动估计的四分之一像素精度、回路内去块效应(in-loop deblocking)滤波器,内容自适应二进制算术编码(context adaptive binary arithmetic coding),等等。

[0005] 运动估计是大多数视频压缩标准(诸如MPEG-1/2/4和H.261/3/4)的核心部分,其

充分利用时间冗余度,因此其性能直接影响视频编码系统的压缩效率、主观视频质量以及编码速度。

[0006] 在块匹配运动估计(BMME)中,在ME中,对当前块和参考块之间的失真的最通用的测量是绝对差值的和(SAD),对于一个 $N \times N$ 块而言,定义为

$$[0007] \quad SAD(mvx, mvy) = \sum_{m=0, n=0}^{N-1} |F_t(x+m, y+n) - F_{t-1}(x+m+mvx, y+n+mvy)|$$

[0008] 其中 F_t 是当前帧, F_{t-1} 是标准帧, (mvx, mvy) 表示当前运动矢量(MV)。对于宽度= X ,高度= Y ,并且块大小= $N \times N$ 的帧而言,在搜索范围 $\pm W$ 中需要对SAD进行评估以查找最优运动矢量的搜索点总数等于:

$$[0009] \quad \left(\frac{X}{N}\right)\left(\frac{Y}{N}\right)(2W+1)^2,$$

[0010] 对于 $X=352, Y=288, N=16$ 并且 $W=32$ 的情况下,其等于1673100。这是在视频编码器中消耗巨大计算能力的巨大数值。已经提出了许多快速算法[2]-[9]来减少在ME中搜索点的数目,例如三步搜索(TSS)[11],2D对数搜索[12],新三步搜索(NTSS)[3],MVFAST[7],以及PMVFAST[2]。MVFAST和PMVFAST显著地优于前三个算法,因为它们使用中值运动矢量预测器作为搜索中心执行中心偏离ME,由此通过平滑运动矢量场降低了MV编码的位的数目。

[0011] PMVFAST算法(其是对MVFAST及其它快速算法的重要改进,并且因而被MPEG标准[10]所接受)最初考虑一组MV预测器,包括中值、零点、左边、顶部、右上方的和先前的共位(co-located)MV预测器。图1举例说明了当前块、左块、顶部块、右上方块、右右上方块、以及右块(其是"未来块",即,在当前块之后被处理的块)的位置。它计算每个预测的SAD代价。在后来的发展中,对PMVFAST进行修改以计算RD(速率失真)代价[13]来替代使用以下代价函数的SAD代价:

$$[0012] \quad J(m, \lambda_{\text{motion}}) = SAD(s, c(m)) + \lambda_{\text{motion}}(R(m-p)) \quad (1)$$

[0013] 其中 s 是原始视频信号, c 是参考视频信号, m 是当前MV, p 是当前块的中值MV预测器, λ_{motion} 是Lagrange乘法器, $R(m-p)$ 表示用于编码运动信息的位。在PMVFAST中的下一步骤是选择具有最小代价的MV预测器,并且根据从MV预测器获得的最小代价的值来执行数量大菱形搜索(diamond search)或者小菱形搜索。

[0014] 在定义当前视频编码标准中独立但是重要问题是使用亚像素预定矢量,包括半像素、1/4像素或者可能甚至1/8像素运动矢量,其提供对运动的更加精确的描述,并且可以提供整像素运动估计的大约1dB的PSNR增益。采用半像素精度,运动矢量可以采用等间距位置值,诸如0.0,0.5,1.0,1.5,2.0等等。采用1/4像素精度,运动矢量可以采用诸如0.00,0.25,0.50,0.75,1.00,1.25,1.50,1.75,2.00等等之类的位置值。采用1/8像素精度,运动矢量可以采用诸如0.000,0.125,0.250,0.375,0.500,0.625,0.750,0.875,1.000,1.125,1.250,1.375,1.500,1.625,1.750,1.875,2.000等等之类的位置值。

[0015] 众所周知,运动矢量分布趋向于中心偏移,这意味着运动矢量趋向于非常地接近于(0,0)。在图6(a)中示出了这种情况,其示出为(0,0)MV使用完全搜索(FS)算法的在Foreman序列中的运动矢量分布。此外,如图6(b)所示,运动矢量分布还向中值预测器(中值MV)偏移,它是在图1中示出的运动矢量左块、顶部块和右上方块的中值。此外,如图6(c)所示,运动矢量还向在当前帧中的相邻运动矢量(leftMV, topMV, topRightMV)和先前帧中所

设置的运动矢量(preMV)偏移,如图6(d)所示。这些可以都是用于当前矢量的运动矢量的可考虑的预测器,并且它们可以被用于PMVFAST。

发明内容

[0016] 本发明目的是提供用于运动估计的新的并且有益的技术,其适用于数字信号压缩、编码和表达的方法和系统。

[0017] 特别是,本发明设法提供新的并且有益的有效运动估计技术,其可以例如应用在MPEG-1、MPEG-2、MPEG-4、H.261、H.263、H.264或AVS或者其它相关视频编码标准中。

[0018] 本发明第一方面基于这样的实现:PMVFAST算法的运动估计尽管与在前技术相比确实具有优点,但是其并不是最佳的。在原理上,对于视频中的每个帧,存在整体上使得整个帧的RD代价最小化的运动矢量场 $\{m_{i,j}, i=0..M-1, j=0..N-1\}$:

$$[0019] \quad total_RD_Cost = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [SAD(s_{i,j}, c(m_{i,j})) + \lambda_{i,j}(R(m_{i,j} - p_{i,j}))] \quad (2)$$

[0020] 其中 (i, j) 表示在包含 $M \times N$ 个块的帧中的第 (i, j) 个块。对于固定 Q_p (其是量化参数), $\lambda_{i,j} = \lambda = \text{恒量}$,而且

$$[0021] \quad p_{i,j} = \text{median}(m_{i,j-1}, m_{i-1,j}, m_{i-1,j+1}) \quad (3)$$

[0022] 然而,考虑到整个帧的全部RD代价同时需要指数级的计算复杂性,这是不实际的。因而,PMVFAST和其它已知算法每次仅仅考虑仅一个块的RD代价,而不是一帧中全部块。

[0023] 特别是,PMVFAST或PMVFAST都没有考虑当导出相对于当前块的运动矢量时,这导致下一个块的在中值MV预测器中的变化。这能够影响整个运动矢量场的平滑。

[0024] 一般地说,本发明的第一方面通过改善PMVFAST的代价定义和运动预测器候选的选择而提出了一种新的ME算法。特别是,对于第一图像的每个当前块(当前块可以是 $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 4 \times 8, 8 \times 4, 4 \times 4$ 或其它矩形长度,甚至非矩形),根据一个代价函数来选择第二图像(参考图像)的类似块,所述代价函数包括两项:(i)当前块与相似块的不相似测量(例如SAD, SAE)的项,以及(ii)作为至少是对第一图像的未来块的运动矢量的预测的函数的项。

[0025] 特别是,所提出的算法通过包含当前中值MV预测器并且还包含未来(即至今未处理)的编码块的估计中值MV预测器,使得改善运动场平滑性成为可能。

[0026] 本发明的许多变化是可能的。特别是,所述块可以具有任何大小和任何形状。

[0027] 可以有多个第二图像(即多个参考值)并且所述搜索可以包括在所有第二图像中的候选位置。

[0028] 此外,可以为一起构成所述第一图像中的较大区域的、并且采用由编码数字所定义的编码顺序进行编码的多个子块执行所述新的代价函数。这些子块无须具有相同大小或形状。

[0029] 本发明具有另一方面,其可以与本发明第一方面组合或独立使用。一般地说,本发明第二方面提出当对第一图像(当前块可以是 $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 4 \times 8, 8 \times 4, 4 \times 4$,或其它矩形大小,甚至非矩形)的当前块进行编码时,使用所选出的、具有从与已知技术中所使用的值不同的一组值中选出的位置值(即两个轴线方向中的各自组件)的运动矢量进行编码。

[0030] 考虑一个可能的运动矢量预测器： $(0,0)$ 。而整像素的常规技术允许运动矢量采用诸如 $-2.0, -1.0, 0, 1.0, 2.0$ 等的位置值，本发明第二方面提出修改接近于所述预测器的一组可能位置值。对于最接近0的位置值1.0，我们可以使用另一位置值诸如0.85，从而使得可允许的位置值将会包括 $-2.0, -0.85, 0, 0.85, 2.0$ 等等。其优点在于，在统计上，运动矢量趋向接近于0。并且，因而通过选择更加接近于0的位置，我们将更加接近真实运动矢量，并且因而可以给出可以导致较高压缩效率的更好的运动补偿。

[0031] 因而，在本发明一个特定表达中，可以选择为至少一个轴方向选择一组可能位置值，从而使得它们无法全部写作 Lm ，其中 $m = -\dots, 2, -1, 0, 1, 2, \dots$ ，并且 L 是常量（例如1个像素间隔， $1/2$ 像素间隔，或 $1/4$ 像素间隔）；即，位置值是不均匀的。特别是，可以选择为至少一个轴方向选择一组可能位置值，从而使得它们无法全部写作 m/n ，其中 $m = -\dots, 2, -1, 0, 1, 2, \dots$ 并且 n 是1或2的幂。

[0032] 注意，本发明的第二方面并不局限于从一组非均匀的空间位置值中选择位置值；与传统的位置值组相比，也不局限于仅仅选择最接近于零的两个位置值。作为示例，在本发明的第二方面的另一个示例里，位置值2.0可变成1.9，从而可允许的位置值将包括 $-1.9, -0.85, 0, 0.85, 1.9$ 等。

[0033] 因而，在本发明第二方面的可替换的特定表达中，（为至少一个所述方向）选择一组可能位置值，以包括可以写作为 LA_m/n 的一个或多个位置值，其中 $m = -\dots, 2, -1, 0, 1, 2, \dots$ ， n 是1或2的幂， L 是常量（例如1像素间隔， $1/2$ 像素间隔，或 $1/4$ 像素间隔），并且 A_m 是小于1但是至少是0.75的值（对于不同的 m 值选择性地不同），更加优选地至少是0.80，并且最优选地至少是0.85。

[0034] 我们已经发现 A_m 的最佳值取决于视频。

[0035] 本发明第二方面的特定实施例的一个优点在于，它们所产生的运动矢量可以采用与常规算法相同的格式码进行编码，除了该位置值的常规码应当分别解释为该实施例所使用的可能位置值之外。例如，如果由特定实施例所使用的位置值是 $-1.9, -0.85, 0, 0.85, 1.9$ 等，则位置值1.0的常规码应该解释为0.85，并且位置值2.0的常规码应该解释为1.9，等。

[0036] 根据本发明的第二方面的方法可以包括以下步骤：定义搜索区域，在所述搜索区域内定义多个候选位置，所述多个候选位置包括由本发明第二方面的新的位置值所定义的一组多个位置。这些位置值是来自关键位置（例如， $(0,0)$ 运动矢量位置，预测运动矢量，等）处的候选位置的相应位移值。对于每个候选运动矢量而言，我们计算代价函数，所述代价函数是第一图像中的当前块与第二图像中的所述候选运动矢量处的块之间的相似度测量（例如SAD, SAE）的函数。可选地，其还可以是以下运动矢量的函数：所述候选运动矢量、当前预测运动矢量，以及可选地，如本发明第一方面，一个或多个未来预测运动矢量。例如，可选地，如本发明第一方面，可以给定该代价函数。

附图说明

[0037] 现在将仅仅参考以下附图描述本发明实施例作为示例，其中：

[0038] 图1示出了当前块、左块、顶部块、右上方块、右右上方块、和右块；

[0039] 图2示出了Diff的典型分布，其是 $|m_{i,j+1} - p_{i,j+1}|$ 与 $|m_{i,j} - p_{i,j+1}|$ 之间的差异；

[0040] 图3a示出了如本发明第一实施例所使用的巨大菱形搜索的搜索方式；

- [0041] 图3b示出了如本发明第一实施例中所使用的修改的巨大菱形搜索的搜索方式；
- [0042] 图4示出了如本发明第一实施例中所使用的小菱形搜索的搜索方式；
- [0043] 图5a和5b比较PMVFAST和本发明第一实施例的MV域的平滑性；
- [0044] 图6示出了对当前帧中的(a)(0,0)MV、(b)中值MV、(c)相邻MV和先前帧中所配置的MV、(d)在先前帧中的右下方的MV(PreBottomRightMV)使用完全搜索(FS)算法得到的在Foreman序列中的运动矢量分布；
- [0045] 图7是本发明第一实施例的流程图。

具体实施方式

[0046] 本发明第一实施例采用了PMVFAST算法的许多特征,但是是通过考虑几个邻近块而不是仅仅一个块来对PMVFAST(及其它现有算法)加以改进的。等式(2)和(3)示出,当前块MV的选择直接影响邻近块的RD代价,所述邻近块包括右块(或第(i,j+1)块),左下方块(或第(i+1,j-1)块),以及下方块(或第(i+1,j)块)。这是因为,当前MV将会影响这些邻近块的所预测MV,并且因此进而影响那些块的最佳运动矢量。这些是“未来”块,因为当处理当前块时,还没有对它们执行运动估计。我们无法与当前块同时地计算这些未来块的最佳运动矢量,因为我们将需要按照等式(2)同时计算整个帧中所有块的最佳运动矢量,这将会非常复杂。

[0047] 作为替代,为了评定在右块或第(i,j+1)块上的当前块的当前MV的选择的蕴含式,我们可以为等式(1)的当前块的RD代价函数增加一项:

$$[0048] \quad R(|m_{i,j+1}-p_{i,j+1}|) \quad (4)$$

[0049] 其中, $m_{i,j+1}$ 是基于当前块的当前MV的右块的最佳运动矢量,并且 $p_{i,j+1}$ 是基于当前块的当前MV的右块的中值MV预测器,即

$$[0050] \quad p_{i,j+1} = \text{median}(m_{i,j}, m_{i-1,j+1}, m_{i-1,j+2}) \quad (5)$$

[0051] 然而,等式(4)中的 $m_{i,j+1}$ 是未知的,因为右块(未来块)的运动估计还没有执行。然而,我们注意到, $|m_{i,j+1}-p_{i,j+1}|$ 可以通过 $|m_{i,j}-p_{i,j+1}|$ 很好地近似。让Diff为 $|m_{i,j+1}-p_{i,j+1}| - |m_{i,j}-p_{i,j+1}|$ 。我们对许多视频测试顺序进行实验并且研究Diff的分布。图2显示了Foreman序列的概率密度函数(概率分布函数),如图2所示,典型的结果严重地偏向零。这意味着这两个量在大部分情况中是基本相同的(在大约70%情况下相同,而在大约23%的情况下仅仅相差1)。这意味着, $|m_{i,j}-p_{i,j+1}|$ 是 $|m_{i,j+1}-p_{i,j+1}|$ 的良好近似。因此, $R(|m_{i,j+1}-p_{i,j+1}|)$ 可以由 $W \cdot R(|m_{i,j}-p_{i,j+1}|)$ 来近似,其中 $W > 0$ 。同样,我们可以对左下方块和下方块添加附加项到代价中。

[0052] 对于第(i,j)块,让medianMV表示由等式3给出中值MV预测器。让FmedianMV表示由等式(5)给出的未来中值MV预测器(用于右块的中值MV预测器)。从而,FmedianMV是MV候选的函数。在此,第一实施例被称为对于第(i,j)块的“增强预测运动矢量场自适应搜索技术”(E-PMVFAST)。该实施例的步骤如下。

[0053] 实施例的各步骤如下,示于图7中。

[0054] 对任何候选MV,如下定义代价。

$$[0055] \quad \text{cost}(\text{MV}) = \text{SAD} + \lambda * [w * R(\text{MV} - \text{medianMV}) + (1-w) * R(\text{MV} - \text{FmedianMV})]$$

[0056] (6)

[0057] 1. 计算三个运动矢量预测器的代价:(i)中值MV预测器(“medianMV”),(ii)右块的估计运动矢量(“futureMV”),其定义为如下:

[0058] $futureMV \equiv \text{median}(\text{TopMV}, \text{TopRightMV}, \text{TopRightRightMV})$ 和(iii)来自过去块(“pastMV”)的MV预测器,其是先前的共位MV(“PreMV”)和远离于medianMV的先前右下方的MV(“PreBottomRightMV”)中的一个,即

[0059] $pastMV \equiv \arg \max_{MV \in \{\text{PreMV}, \text{PreBottomRightMV}\}} \{\text{abs}(MV - \text{medianMV})\}$

[0060] 注意,项(ii)可以由用于另一邻近未来块(诸如左下方的、底部,和/或右下方的块)的估计运动矢量来补充或代替。

[0061] 还应当注意,在项(iii)中,先前右下方的MV可以由用于另一邻近块的先前MV预测器补充或替换。

[0062] 注意,项(ii)和(iii)形代价发明的独立方面。

[0063] 如果以上任意一个MV预测器不可用(例如在帧的边界),则跳过该预测器。

[0064] 2. 如果运动矢量预测器的最小代价小于阈值T1,则停止搜索并且转到步骤7。否则,选择具有最小代价的运动矢量作为currentMV(当前MV)并且转到下一步骤。注意,3个运动矢量预测器的代价可以采用预定顺序来计算(例如medianMV,接着是futureMV,接着是pastMV),并且在任何时刻,如果任何运动矢量预测器的代价小于特定阈值,则搜索可能停止并且转到步骤7。

[0065] 3. 围绕currentMV执行定向小菱形搜索的一次迭代。下面解释定向小菱形搜索的概念。

[0066] 4. 如果最小代价小于阈值T2,则停止搜索并且转到步骤7。否则,选择具有最小代价的运动矢量作为currentMV并且转到下一步骤。

[0067] 5. 如果(currentMV=medianMV)并且当前最小代价小于阈值T3,则执行小菱形搜索并且转到步骤7。

[0068] 6. 如果视频不是隔行扫描,则执行大菱形搜索,如图3(a)所示;否则,执行如图3(b)所示的修改的大菱形搜索。在这些步骤中的每个步骤中,对菱形的每一标志点评估代价函数。

[0069] 7. 选择具有最小代价的MV。

[0070] 在我们的实验中,发现w的值大约0.8是有效。

[0071] 现在解释定向小菱形搜索的步骤。假定,centerMV是当前搜索中心,并且MV1,MV2,MV3和MV4是四个围绕搜索点,如图4所示。为每个MVi计算 $R(MVi - \text{medianMV})$ 。如果 $R(MVi - \text{medianMV}) < R(\text{centerMV} - \text{medianMV})$,则计算MVi的SAD和代价。否则,忽略该MVi。选择具有最低代价的MV作为currentMV。注意,定向方块搜索的概念被认为是新的,并且构代价发明一个独立方面,其无须与使用futureMV的概念相结合来执行。

[0072] 大菱形搜索和修改的大菱形搜索的步骤是相同的,但是所述搜索是对于图3(a)和3(b)中分别示出的点的所有集合而完成的。

[0073] 我们现在考虑在本发明范围内的实施例的多个可能的变体。

[0074] 首先,注意,对于不同的块,在代价函数中的加权系数w可以是不同的。此外,可选择地,对于不同的MV候选所述w可以是不同的。特别地是,w的定义可以取决于诸如MV候选是否接近于medianMV和/或futureMV、或MV候选的X轴分量或Y轴分量是否与所述FmedianMV的

X轴分量或Y轴分量相同之类的情况。

[0075] 此外,代价函数可以不限于等式(6)的形式。其可以是包括失真测量项(例如SAD、失真平方和(SSD)、平均偏差失真(MAD)、MSD等)和考虑了对当前块和某些邻近块(例如右块、下方块、左下块等)的运动矢量进行编码所必需的位的项的任何函数。

[0076] 此外,在步骤1中,futureMV的定义不局限于以上步骤1给出的形式。对于futureMV的两个可能的替换定义是:

[0077] $futureMV \equiv \text{median}(\text{leftMV}, \text{TopRightMV}, \text{TopRightRightMV})$

[0078] $futureMV \equiv \text{median}(\text{medianMV}, \text{TopRightMV}, \text{TopRightRightMV})$

[0079] 此外,在以上表达的步骤1中,pastMV被选为在先前帧中的可能MV(preMV和preBottomRightMV)的列表集合中距离medianMV最远的一个。然而,要考虑的MV列表可以包含两个以上可能的MV(例如preMV,preLeftMV,preRightMV,preTopMV,preTopLeftMV,preTopRightMV,preBottomMB,preBottomLeftMV,preBottomRightMV,等)。此外,来自一个以上先前编码帧的MV可以包括在所述列表中(例如如果当前帧是帧N,则所述列表可以包含帧N-1,N-2,N-3,...)。如果当前帧是B帧,则先前编码帧的列表可以包括未来P帧。

[0080] 此外,在步骤1中,选择pastMV为距离参考MV(在步骤1中的medianMV)最远的可能的MV。其它参考MV也是可以的,包括leftMV、或TopMV、或TopRightMV、或某组合。从可能的MV的列表中进行选择的其它方法也是可能的。

[0081] 在步骤2中,所述3个运动矢量预测器的代价是在按照某种预定义顺序中得到的。可能的预定义顺序包括:

[0082] a)medianMV,随后是futureMV,随后是pastMV

[0083] b)medianMV,随后是pastMV,随后是futureMV

[0084] c)futureMV,随后是medianMV,随后是pastMV

[0085] d)futureMV,随后是pastMV,随后是medianMV

[0086] e)pastMV,随后是medianMV,随后是futureMV

[0087] f)pastMV,随后是futureMV,随后是medianMV

[0088] 此外,虽然如以上表达,在步骤3中执行了定向小菱形搜索的一次迭代,可以应用一个以上的迭代。

[0089] 模拟结果

[0090] 我们现在给出实施例E-PMVFAST的模拟结果。将所述实施例嵌入到H.264参考软件JM9.3[13]中,并且使用各种QP、视频序列、分辨率和搜索范围对其进行模拟。表1(a-c)和2(a-c)示出一些典型的模拟结果。PSNR(峰值信号与噪声的比例)变化和BR(比特率)变化是相对于完全搜索(FS)的PSNR和比特率的变化。模拟结果示出,所提出的E-PMVFAST的比特率和PSNR趋向于与完全搜索和PMVFAST类似,但是在大范围的视频序列和比特率上,E-PMVFAST趋向于比PMVFAST快大约40%。E-PMVFAST的一个重要特征在于,其运动矢量场趋向于非常地平滑,从而使得运动矢量可以比其它快速移动估计算法更加准确地表示对象的移动。

[0091] 在图5(a)以及5(b)中,左侧的图像示出(如短线)由PMVFAST算法获得的运动矢量场,而右侧的成像示出通过所述实施例获得的相同成像的运动矢量。E-PMVFAST的运动矢量场显著地比PMVFAST的运动矢量场更加平滑,特别是在所圈出的区域中。对于将感知转换编

码(perceptual trans-coding)、速率控制、多种块大小运动估计、多种标准帧运动估计等中的视频运动内容进行分类而言,平滑的运动场非常有用。

[0092]

Foreman CIF								
FS			PMVFAST			E-PMVFAST		
QP	PSNR	BR	PSNR 改变	BR 改变	速度 提高	PSNR 改变	BR 改变	速度 提高
24	38.61	10457168	-0.01	1.12%	269.90	-0.01	0.73%	345.62
26	37.25	7498832	-0.02	1.13%	259.61	-0.02	1.17%	337.49
28	35.99	5397584	-0.02	1.53%	250.27	-0.03	1.21%	329.79
30	34.64	3804232	-0.02	1.27%	241.91	-0.04	1.20%	326.41
32	33.34	2730064	-0.02	1.19%	235.26	-0.03	1.22%	321.96
34	32.25	2044176	-0.05	1.57%	230.48	-0.05	1.38%	325.01

[0093]

36	31.70	1529848	-0.07	1.20%	226.76	-0.09	1.28%	325.47
38	29.9	1184744	-0.05	1.75%	223.29	-0.06	0.99%	328.10
40	28.89	960056	-0.09	1.85%	220.98	-0.11	1.09%	334.25
42	27.79	794312	-0.14	1.71%	218.94	-0.16	1.74%	341.62
44	26.7	672224	-0.17	1.58%	217.42	-0.22	0.89%	351.22
平均			-0.06	1.45%	235.89	-0.07	1.17%	333.36

[0094] 表格1a-foreman CIF序列的模拟结果

[0095]

Coastguard CIF								
FS			PMVFAST			E-PMVFAST		
QP	PSNR	BR	PSNR 改变	BR 改 变	速度 提高	PSNR 改变	BR 改变	速度 提高
24	37.32	19901336	0	0.02%	245.39	0	0.02%	328.81
26	35.77	15169632	-0.01	-0.03%	236.12	0	0.01%	319.07
28	34.35	11403728	-0.01	-0.07%	232.24	-0.01	-0.03%	308.17
30	32.8	8097840	0	0.07%	229.89	0	0.04%	301.18
32	31.34	5616880	0.01	0.23%	228.38	0	-0.04%	296.55
34	30.08	3913296	-0.01	-0.08%	227.20	-0.01	-0.23%	294.13
36	28.75	2546328	-0.02	0.03%	225.88	-0.01	-0.04%	292.18
38	27.55	1684056	-0.03	0.48%	225.13	-0.03	0.35%	293.95
40	26.55	1172448	-0.04	-0.01%	224.78	-0.04	0.15%	299.45
42	25.52	800328	-0.06	0.65%	225.22	-0.06	1.00%	308.85
44	24.54	572224	-0.07	-0.10%	225.47	-0.06	0.39%	321.38
平均			-0.02	0.11%	229.61	-0.02	0.15%	305.79

[0096] 表格1b—Coastguard CIF序列的模拟结果

[0097]

Hall CIF								
FS			PMVFAST			E-PMVFAST		
QP	PSNR	BR	PSNR 改变	BR 改 变	速度 提高	PSNR 改变	BR 改变	速度 提高
24	39.63	6954760	-0.01	0.26%	424.55	-0.01	0.18%	700.53
26	38.6	4539912	-0.02	0.17%	397.68	-0.02	0.43%	704.18
28	37.58	2993752	-0.04	0.07%	370.17	-0.04	0.06%	686.71
30	36.36	2021712	-0.04	0.01%	340.96	-0.04	0.18%	674.18
32	35.17	1411344	-0.03	-0.86%	321.03	-0.05	-0.40%	655.08

[0098]

34	33.94	957600	0.02	0.40%	301.76	-0.02	0.44%	659.53
36	32.68	677768	-0.04	0.12%	288.55	-0.06	-0.12%	642.68
38	31.4	493168	-0.05	-0.04%	276.97	-0.05	-0.38%	641.42
40	30.05	379240	0.05	-0.45%	263.07	0	-0.48%	625.73
42	28.67	279008	-0.03	-0.99%	251.93	-0.07	-0.84%	624.02
44	27.38	214544	-0.02	-1.32%	241.90	-0.04	-1.03%	635.41
平均			-0.02	-0.24%	316.23	-0.04	-0.18%	659.04

[0099] 表格1c—Ha11 CIF序列的模拟结果

[0100]

Foreman QCIF								
FS			PMVFAST			E-PMVFAST		
QP	PSNR	BR	PSNR 改变	BR 改 变	速度 提高	PSNR 改变	BR 改变	速度 提高
24	38.32	2816360	0	0.68%	260.10	-0.01	-0.05%	356.91
26	36.85	2122432	-0.02	0.28%	249.94	-0.02	-0.42%	343.24
28	35.47	1615528	-0.01	0.41%	242.29	-0.01	0.09%	329.89
30	34.01	1213792	-0.01	0.44%	235.40	-0.02	0.35%	322.04
32	32.61	918272	-0.04	0.57%	231.38	-0.01	-0.13%	316.56
34	31.39	710976	-0.04	0.65%	228.06	-0.05	0.05%	314.51
36	30.07	536152	-0.1	0.85%	225.33	-0.07	0.60%	311.05
38	28.81	416112	-0.04	0.31%	223.95	-0.1	-0.10%	310.02
40	27.65	328288	-0.08	0.74%	222.89	-0.08	-0.30%	311.37
42	26.5	259936	-0.12	-0.39%	221.19	-0.1	0.47%	315.72
44	25.36	207344	-0.18	1.43%	218.38	-0.25	-0.08%	321.03
平均			-0.06	0.54%	232.63	-0.07	0.04%	322.94

[0101] 表格2a—foreman QCIF序列的模拟结果

[0102]

Akiyo QCIF								
FS			PMVFAST			E-PMVFAST		
QP	PSNR	BR	PSNR 改变	BR 改 变	速度 提高	PSNR 改变	BR 改变	速度 提高
24	40.81	530832	0.01	0.01%	483.45	0.01	0.01%	1143.12
26	39.42	391936	0	-0.38%	428.54	0	-0.38%	1017.49
28	38	289096	0	-0.09%	389.95	-0.01	-0.25%	879.72

[0103]

30	36.57	208488	-0.01	0.40%	352.74	-0.03	0.33%	779.84
32	35.03	153776	-0.04	0.83%	312.72	-0.02	1.01%	699.16
34	33.66	118440	0.03	-0.05%	287.94	0.02	0.58%	644.63
36	32.27	89056	0.01	1.11%	270.48	0.03	0.88%	621.79
38	31.09	70280	-0.01	-0.01%	254.10	-0.03	-0.31%	598.06
40	30.13	57200	0	0.92%	245.76	-0.06	1.58%	612.01
42	28.77	47648	-0.02	0.22%	241.09	-0.02	0.15%	629.43
44	27.68	40608	0.02	0.55%	240.05	0.01	-0.91%	669.52
平均			0.00	0.32%	318.80	-0.01	0.25%	754.07

[0104] 表格2b-Akiyo QCIF序列的模拟结果

[0105]

Coastguard QCIF								
FS			PMVFAST			E-PMVFAST		
QP	PSNR	BR	PSNR 改变	BR 改 变	速度 提高	PSNR 改变	BR 改变	速度 提高
24	37	4456208	0	0.15%	251.59	0	0.09%	346.68
26	35.4	3298352	0	0.12%	238.53	0	0.08%	330.67
28	33.92	2386152	-0.01	0.15%	234.60	-0.01	0.24%	316.97
30	32.33	1633736	-0.02	0.06%	232.80	0	0.01%	310.88
32	30.86	1110088	0	0.22%	231.60	0	-0.17%	309.50
34	29.62	770656	-0.01	-0.38%	230.98	-0.01	-0.32%	310.31
36	28.3	511000	-0.02	0.28%	230.30	-0.01	0.50%	312.49
38	27.06	354208	0.02	-0.47%	229.61	0.01	-0.57%	314.41
40	26.08	256480	-0.08	-0.72%	229.39	-0.08	-1.08%	317.67
42	25.02	179528	-0.06	2.16%	228.04	-0.07	0.45%	325.54
44	24	131856	-0.09	-0.05%	227.30	-0.09	0.57%	335.34
平均			-0.02	0.14%	233.16	-0.02	-0.02%	320.95

[0106] 表格2c-Coastguard QCIF序列的模拟结果

[0107] 我们现在转向本发明的第二实施例,其示出了本发明的第二方面。

[0108] 如上所述,常规完全整数像素允许运动矢量在每个方向上采用-2.0,-1.0,0,1.0,2.0等等的位置值。在本发明第二实施例中,选择接近于预测器的可能位置值。对于最接近0的位置值,我们可以使用(代替1.0)另一位置值,诸如0.85,从而使得允许的位置值可以包括-2.0,-0.85,0,0.85,2.0等等。此优点在于,在统计上运动矢量趋向接近于0。因此,通过选择更加接近0的位置,我们会更加接近于真实的运动矢量,并且因而可以给出可以导致较高压缩效率的更好的运动补偿。类似地,可以改变其它位置值。例如,可以将位置值2.0改变

为1.9,从而使得允许的位置值会包括-1.9,-0.85,0,0.85,1.9等等。除了1.0的编码运动矢量位置应当解释为0.85,并且2.0的编码运动矢量位置应当解释为1.9以外,所提出的改变的优点在于可以使用相同运动矢量代码。

[0109] 半像素精度允许运动矢量采用诸如0.0,0.5,1.0,1.5,2.0等等之类的位置值。我们建议修改这些位置值,特别是那些接近于预测器的位置值。对于非常接近于0的位置值0.5,我们建议使用一个不同的值。例如,一种可能性是使用0.4代替0.5。换句话说,位置值将包括0.0,0.4,1.0,1.5,2.0。类似地,其它位置值可以被修改。例如,位置值1.0可以改变为0.9,从而使得该组新位置值将包括0.0,0.4,0.95,1.5,2.0等等。同样,这可以有助于提高压缩效率。类似地,可以修改其它位置值以提高压缩效率。然而,改变这种位置能够显著地导致编码器和解码器两者处的较高的计算效率。通常,大部分压缩效率增益来自于将位置值改变为接近于预测器。

[0110] 1/4像素精度允许运动矢量采用诸如0.00,0.25,0.50,0.75,1.00等等之类的位置值。我们可以修改位置值,特别那些接近于预测器的位置值。例如,我们可以将它们修改为0.00,0.20,0.47,0.73,0.99等。

[0111] 注意,所提出的方法允许我们在每个整数位置值之间选择任意数量的位置值。例如,在位置值0和1之间,半像素精度使用1个位置值{0.5},1/4像素精度使用3个位置值{0.25,0.50,0.75},以及1/8像素精度使用7个位置值{0.125,0.250,0.375,0.500,0.625,0.750,0.875}。提出的方法允许我们在0和1之间选择任何N个位置值。例如,我们可以选择N=2值诸如0.3和0.6。

[0112] 所提出的不均匀的亚像素运动估计和补偿不是必须要应用于每个帧的每个区域。相反,可以将某些位引入到头部中,以指示对于所述视频帧的每个区域(例如片)其是否被开启或关闭。除此之外,其可以在没有任何语法改变的情况下直接应用于现有标准,因为可以应用相同的运动矢量代码。

[0113] 采用H.264 JM82软件对所提出的不均匀的亚像素运动估计和补偿进行模拟,并且其结果在以上表格中示出,其中QP代表量化参数。该模拟在x和y方向中使用位置值(...-1,-0.75,-0.5,-0.15,0,0.15,0.5,0.75,1...)。即,与使用1/4像素空间位置值的标准方法相比较,仅仅修改了在-0.25和+0.25处的位置值。除了使用新的位置值之外,所述算法在其他方面与已知的H.264标准算法相同。如在表中所示出的,第二实施例在实现相似的PSNR的同时,明显地降低了比特率。不需要修改H.264的语法。

[0114]

序列		QP=20			QP=24		
		JM82	建议值	差异	JMS2	建议值	差异
Akiyo	PSNR	44.11	44.13	0.02	41.73	41.74	0.01
	比特	3332792	3140080	-5.78%	1766112	1674296	-5.20%
News	PSNR	42.79	42.83	0.04	40.19	40.22	0.03
	比特	6908656	6665160	-3.52%	4156840	4017968	-3.34%
weather	PSNR	43	43.01	0.01	39.75	39.58	0.01
	比特	6929368	6803496	-1.82%	4425912	4351312	-1.69%
Sean	PSNR	42.65	42.67	0.02	39.8	39.83	0.03
	比特	5009336	4755328	-5.07%	2734592	2585480	-5.45%

[0115]

序列		QP = 28			QP = 32		
		JM82	建议值	差异	JM82	建议值	差异
Akiyo	PSNR	39.33	39.34	0.01	36.66	36.69	0.03
	比特	927791	904744	-2.48%	491032	497600	1.34%
News	PSNR	37.6	37.62	0.02	34.68	34.68	0
	比特	2497384	2441216	-2.25%	1476680	1469640	-0.48%

[0116]

weather	PSNR	36.29	36.29	0	32.88	32.89	0.01
	比特	274184	272312	-0.80%	1660648	1665368	0.28%
Sean	PSNR	36.94	36.95	0.01	33.92	33.91	-0.01
	比特	1497232	1433544	-4.25%	793552	778816	-1.86%

[0117] 尽管以上描述的仅仅是本发明的几个实施例,在本发明范围内许多变化都是可能的。

[0118] 例如,以上给出的本发明的描述是用于具有一个参考帧的P帧中的固定大小的块。然而,本发明可以应用于具有多种子块大小的块,并且所述块不是必须要非重叠的。可以有一个以上的参考帧,并且参考帧可以是视频序列的相对于当前帧的过去或将来的任何块。

[0119] 对于视频,一个图像元素(像素)可以具有一个或多个分量,诸如亮度分量、红色,绿色,蓝色(RGB)分量、YUV分量、YCrCb分量、红外分量、X光或其它分量。图像元素的每个分量是可以表示为数字的符号,所述数字可以是自然数、整数、实数甚至是复数。在自然数的情况下,它们可以是12位,8位,或任何其它位分辨率。虽然在视频中的像素是具有的矩形采样网格和均匀采样周期的2维样本,但是所述采样网格不是必须为矩形并且所述采样周期

不是必须为是均匀的。

[0120] 工业实用性

[0121] 本发明的每个实施例适合于由MPEG-1、MPEG-2、MPEG-4、H.261、H.263、H.264、AVS或其他可以被修改以包括以上它的相关视频编码标准或方法的快速、低延迟和低代价软件和硬件实现来实现。可能的应用包括数字视频广播(地面、卫星、有线)、数字照相机、数字可携式摄像机(camcorder)、数字录像机、机顶盒、个人数字助理(PDA)、可使用多媒体的蜂窝式电话(2.5G、3G及以上)、视频会议系统、视频点播系统、无线局域网设备、蓝牙应用、网络服务器、低或高带宽应用中的视频流服务器、视频代码转换器(从一个格式转换到另一格式)、及其它电视通信系统等。

[0122] 参考文献

[0123] 以下参考文献的公开内容于此全面引入：

[0124] [1] Joint Video Team of ITU-T and ISO/IEC JTC 1, "Draft ITU-T Recommendation and Final Draft international Standard of Joint Video Specification (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC)", document JVT-G050r1, may 2003

[0125] [2] A.M. Tourapis, O.C. Au and M.L. Liou, "Predictive Motion Vector Field Adaptive Search Technique (PMVFAST)", ISO/IEC JTC1/SC29/WG11 MPEG2000, Noordwijkerhout, NL, March '2000

[0126] [3] R. Li, B. Zeng and M.L. Liou, "A new three-step search algorithm for block motion estimation", On Circuits and Systems for Video Technology, vol.4, no.4, pp.438-42, Aug '94

[0127] [4] Z.L. He and M.L. Liou, "A high performance fast search algorithm for block matching motion estimation", IEEE Trans. on Circuits and Systems for Video Technology, vol.7, no.5, pp.826-8, Oct '97

[0128] [5] A.M. Tourapis, O.C. Au, and M.L. Liou, "Fast Motion Estimation using Circular Zonal Search", Proc. of SPIE Sym. of Visual Comm. & Image Processing, vol.2, pp.1496-1504, Jan. 25-27, '99

[0129] [6] A.M. Tourapis, O.C. Au, M.L. Liou, G. Shen, and I. Ahmad, "Optimizing the Mpeg-4 Encoder-Advanced Diamond Zonal Search", in Proc. of 2000 IEEE Inter. Sym. on Circuits and Systems, Geneva, Switzerland, May, 2000

[0130] [7] K.K. Ma and P.I. Hosur, "Performance Report of Motion Vector Field Adaptive Search Technique (MVFAST)", in ISO/IEC JTC1/SC29/WG11 MPEG99/m81, Noordwijkerhout, NL Mar '00

[0131] [8] A.M. Tourapis, O.C. Au, and M.L. Liou, "Fast Block-Matching Motion Estimation using Predictive Motion Vector Field Adaptive Search Technique (PMVFAST)", in ISO/IEC/JTC1/SC29/WG11 MPEG2000/M5866, Noordwijkerhout, NL, Mar '00

[0132] [9] Implementation Study Group, "Experimental conditions for evaluating encoder motion estimation algorithms", in ISO/IEC JTC1/SC29/WG11 MPEG99/n3141, Hawaii, USA, Dec '99

- [0133] [10]“MPEG-4 Optimization Model Version 1.0”,in ISO/IECJTC1/SC29/WG11 MPEG2000/n3324,Noordwijkerhout,NL,Mar'00
- [0134] [11]T.Koga,K.linuma,A.Hirano,Y.lijima,and T.Ishlguro,“Motion compensated interframe coding for video conferencing” Proc.Nat.Telecommun.Conf.,New Orleans,LA,pp.G.5.3.1-G.5.3.5,Dec'81.
- [0135] [12]J.R.Jain and A.K.Jain,“Displacement measurement and itsapplication in interframe image coding”,JEEE Trans.OnCommunications, vol.COM-29,pp.1799-808,Dec'81
- [0136] [13]JVT reference software JM9.2 for JVT/H.264 FRext

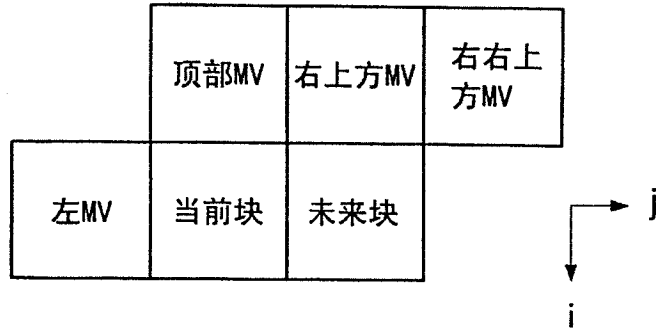


图1

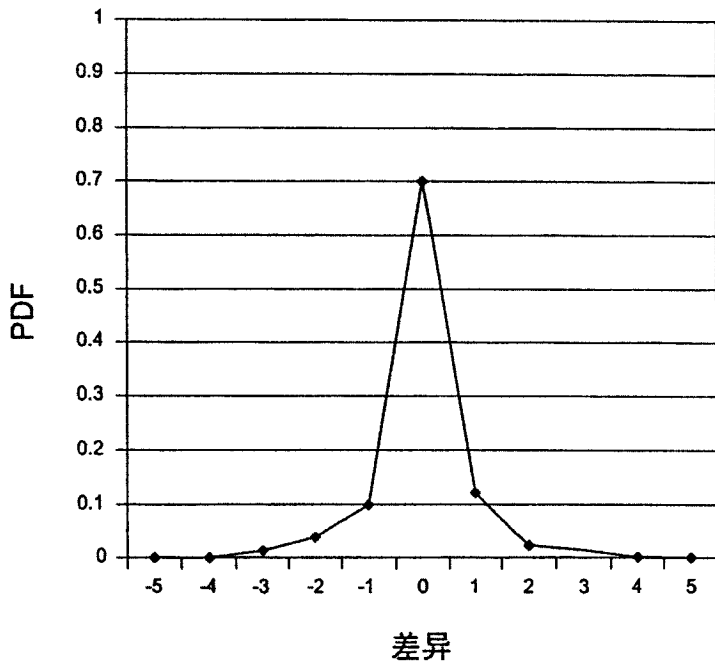


图2

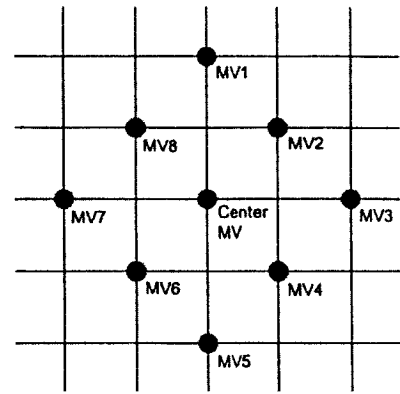


图3a

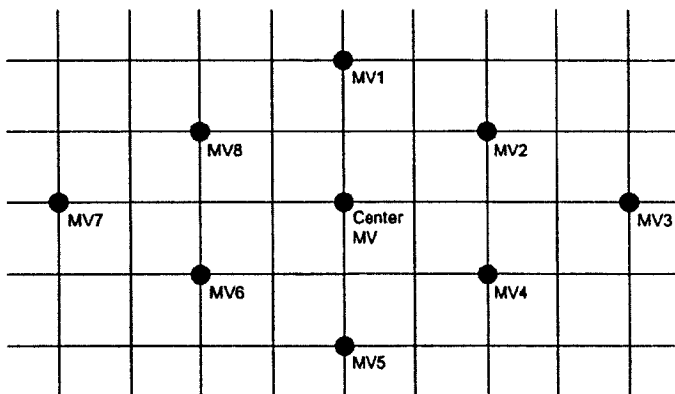


图3b

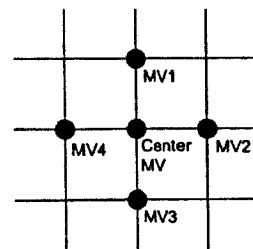


图4

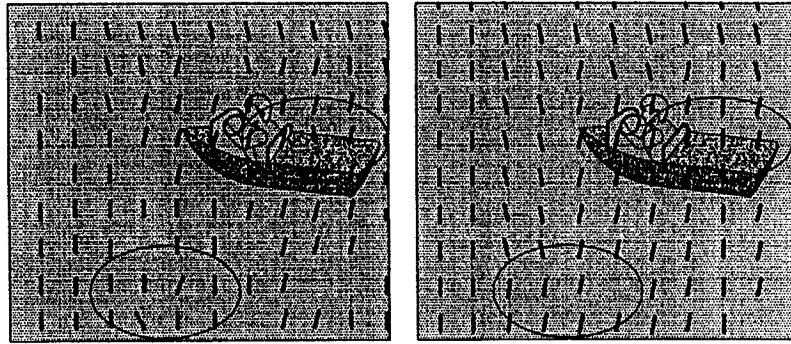


图5a

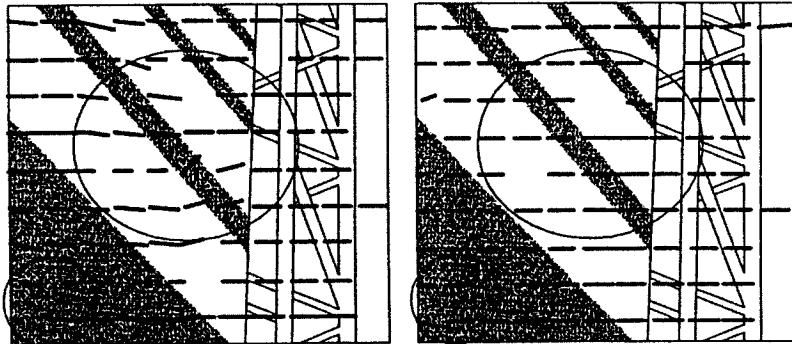


图5b

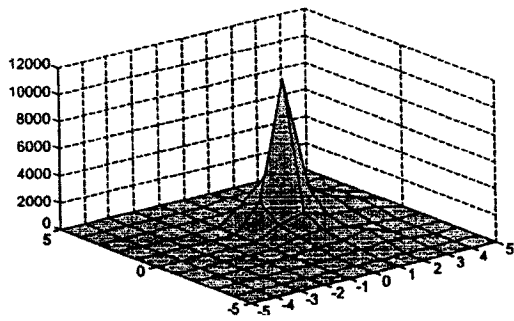


图6a

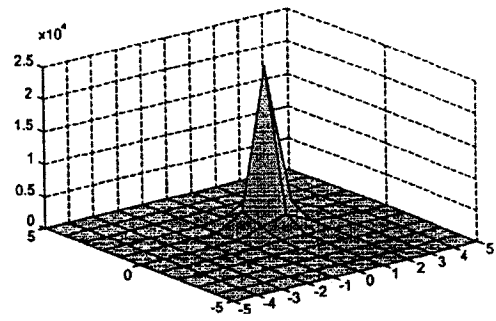


图6b

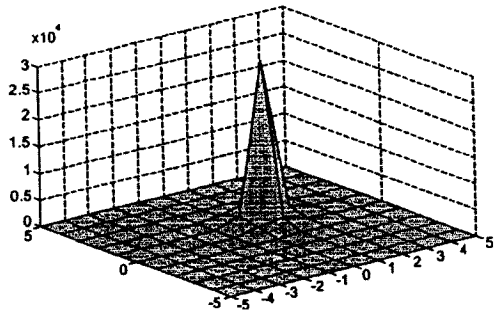


图6c

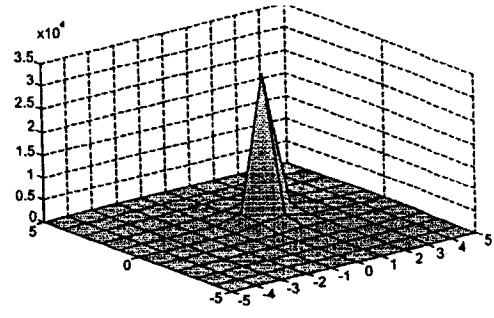


图6d

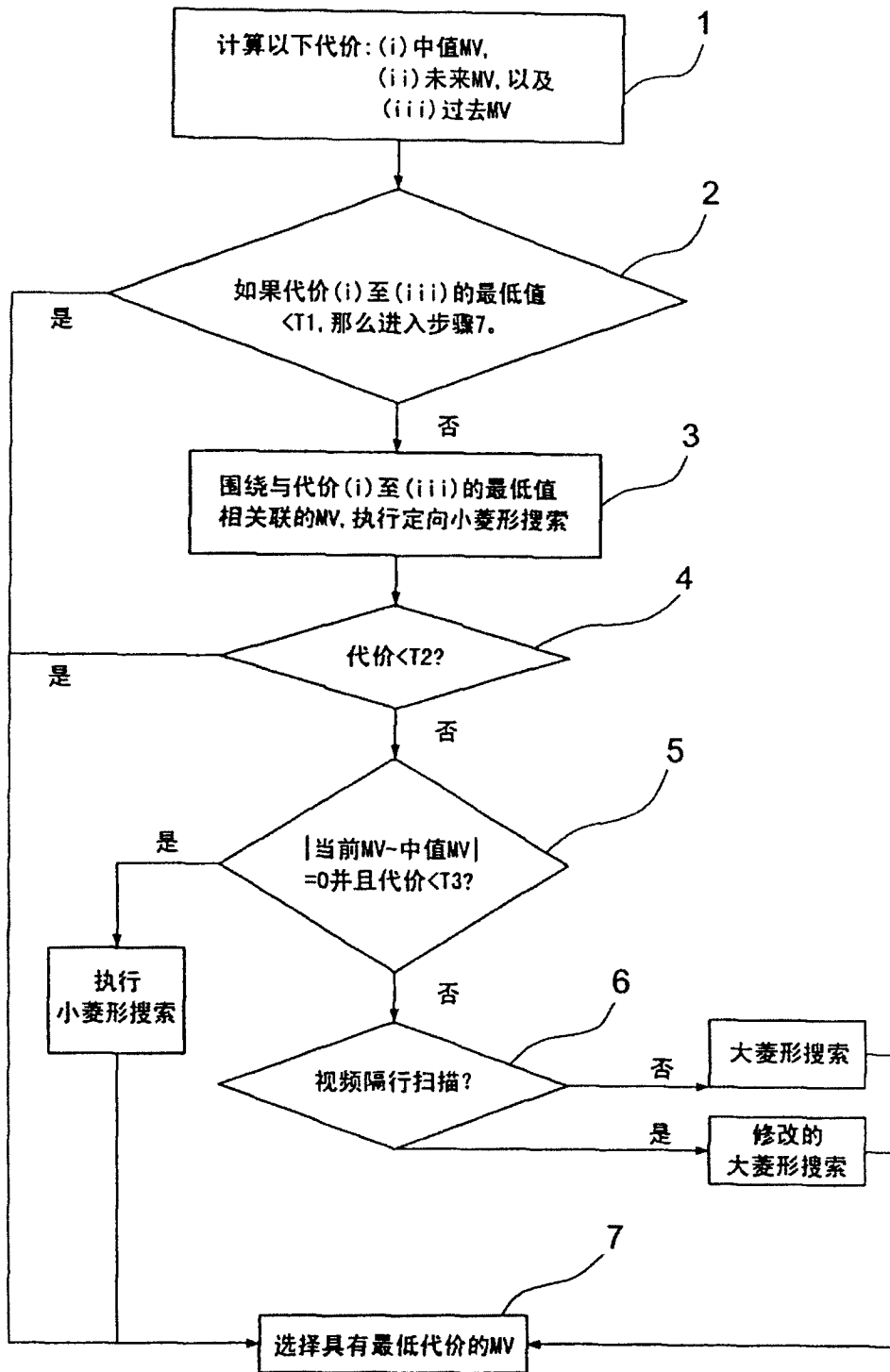


图7