

基于节点搜索的可变形块运动补偿

魏 伟, 侯正信, 郭迎春
(天津大学电子信息工程学院, 天津 300072)

摘 要: 本文讨论可变形块匹配(DBMA)的运动补偿和预测方法, 提出基于节点搜索的可变形块匹配算法(NS-DBMA), 并在此基础上提出分数像素精度预测和双模式混合预测方法. 实验结果表明, NS-DBMA 比全搜索方块匹配法(EBMA)平均改善约 2dB; 其运算量仅为基于梯度的可变形块匹配算法(GB-DBMA)的一半, 但能得到更好的主客观预测质量, 且更易于 VLSI 硬件实现.

关键词: 运动补偿; 可变形块匹配算法; 节点搜索算法

中图分类号: TN919. 81 文献标识码: A 文章编号: 0372-2112 (2005) 08-1421-04

Motion Compensation Using Nodal Search Based Deformable Block Matching

WEI Wei, HOU Zheng-xin, GUO Ying-chun

(School of electronic information engineering, Tianjin University, Tianjin 300072, China)

Abstract: Motion compensation and prediction using deformable block matching algorithm (DBMA) is discussed. This paper proposes a nodal search based DBMA (NS-DBMA) for the estimation of nodal displacement. The sub-pixel accuracy prediction and two mode hybrid method are further proposed. Experimental results show that the NS-DBMA have about 2dB better than exhaustive block matching algorithm (EBMA). NS-DBMA has only half computation amount of the gradient based search algorithm (GB-DBMA) but gets better subjective and objective quality, and is much easier to implement by VLSI.

Key words: motion compensation; deformable block matching algorithm; nodal search algorithm

1 引言

运动补偿是视频压缩编码的核心技术之一, 用于解除图像序列的时间相关性, 广泛应用于 MPEG-x、H. 26x 系列编码标准中. 目前比较成熟的算法是方块匹配算法(BMA), 它将当前帧分割成固定大小的非重叠图像块, 并假设每个像块内所有像素都做相同的平移运动. 由 EBMA、三步法(TSS)等匹配算法为每块指定一个运动矢量(MV). BMA 具有算法简单、速度较快、易于硬件实现等优点, 在以往的视频编码标准中取得了巨大成功.

但 BMA 基于平移模型, 不能适应方块作旋转、缩放或弹性形变等非平移运动. 为了弥补 BMA 的不足, 可采用可变形块模型, 但目前的 GB-DBMA 过于复杂, 无法应用在实际编码系统中. 本文提出 NS-DBMA, 具有算法简单、预测精度高、易于硬件实现等优点, 并很容易得到分数精度预测效果.

2 基于节点位移的可变形块模型

像块的运动变化可用平移、几何变换、仿射变换和双线性变换等几何关系来描述, 并可统一表示成多项式映射的形式^[1]

$$d(p) = \sum_{0 \leq l, m \leq N_1, l+m \leq N_2} a_{l, m} x^l y^m \quad (1)$$

其中 $p = [x, y]$ 是像素在当前图像中的坐标, $d(p) = [dx, dy]$ 是该像素的位移, $a_{l, m} = [c_{l, m}, d_{l, m}]$ 是多项式系数. 当 $N_1 = 1, N_2 = 2$ 时表示双线性变换, 即 $d(p) = a_{00} + a_{10}x + a_{01}y + a_{11}xy$, 且平移、几何变换、仿射变换都可看作双线性变换的特例.

在基于节点位移的可变形块模型中, 像块 D 上选定数目的控制节点可以自由移动, 且 D 上任意点的位移可由控制节点的位移内插得到^[2], 运动参数是各节点位移, 即节点 MV . 如果已知 D 上全部 K 个控制节点 n_k 的位移 $d_k (k = 1, 2, \dots, K)$, 则可用下面的插值公式得到 D 上任意点 p 的位移

$$d(p) = \sum_{k=1}^K \phi_k(p) d_k \quad (2)$$

其中权函数 $\phi_k(p)$ 称为插值核, 表示第 k 个控制节点对 p 点位移的期望贡献. 如果 D 是标准方块, 控制节点为四个顶点, 节点自由移动的结果是方形变化成任意四边形, 如图 1 所示. 按双线性变换设计插值核^[3], 则用来近似表示足够小 D 上的运动变化, 此时的插值核在归一化边长下(边长为 2, 矩形中心为坐标原点)为:

$$\begin{aligned} \phi_1(p) &= (1+x)(1-y)/4; \phi_2(p) = (1+x)(1+y)/4; \\ \phi_3(p) &= (1-x)(1+y)/4; \phi_4(p) = (1-x)(1-y)/4; \end{aligned} \quad (3)$$

3 节点位移估计方法

可变形块模型的运动参数是 K 个节点 MV , 运动估计相当于在 $2K$ 维空间搜索误差最小点, 这比只需估计一个 MV 的平移模型要复杂许多.

3.1 基于梯度的可变形块匹配算法

GB DBMA 是基于梯度的搜索算法. 按参考文献 [2] 的方法, 以误差平方和为目标函数, 先用 EBMA 给定方程的初始解, 再用 Newton

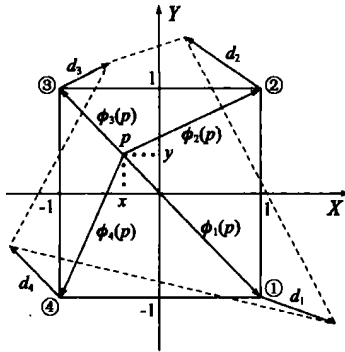


图 1 四节点可变形块模型及节点位置

Raphson 迭代算法即可得到最小误差平方和时的各节点 MV . 在迭代过程中, 不但每次迭代都需要计算矩阵逆、矩阵乘法、图像梯度、方块误差等, 过程复杂、运算量大, 而且对初始解十分敏感. 尽管采用 EBMA 也不能保证得到正确的结果, 同时 EBMA 的运算量也很大. 另外, 为了避免方块退化(节点交错或钝角顶点)和判断方程收敛性等, 还需另外附加运算. 以 16×16 的四节点模型为例, 不包括附加运算, 经 3 次迭代, 每个像块大约需要 165000 次标量乘法和 120000 次标量加法运算, 而且运算过程不易由模块化的硬件实现.

3.2 基于节点搜索的可变形块匹配算法

NS DBMA 由 BMA 得到初始 MV , 再对各节点 MV 作优化调整^①. 设 $\varphi_1(p)$ 为参考帧, $\varphi_2(p)$ 为当前帧, D 为当前编码方块, 节点最大搜索范围为 R , 初始步长为 R 的一半, 即 $R_0 = \lfloor R/2 \rfloor$. 如图 2 所示,

每个节点位移调整搜索九个待选位置(点): 节点自身点(中心点)和八个方位上的八点. 对每个待选点, 由式(2)得到 D 上各点位移并计算块误差平方和 $E = \sum_{p \in D} (\varphi_1(p + d(p)) - \varphi_2(p))^2$, 取 E 最小者作为更新点, 并按更新点位置调整节点 MV ; 用相同步长对其他节点作类似的搜索更新. 完成该步长对全部节点的调整之后, 将步长减半在上一轮各节点更新点基础上作进一步调整, 方法类似, 直到完成一个像素步长调整后结束, 得到各节点的最佳 MV .

节点 MV 的搜索更新方法类似 TSS, 但做了必要的修改, 以使其适应可变形块多节点自由移动的特征. 虽然一个块的 MV 与其他块独立, 但相同块各节点 MV 并不独立, 每一节点 MV 的确定都要考虑到其他节点的位移, 不能独立地对某一节点作最佳 MV 搜索. 在 NS DBMA 中, 初始步长搜索得到各节点位移大致位置, 步长减半搜索即考虑到了其它节点的大

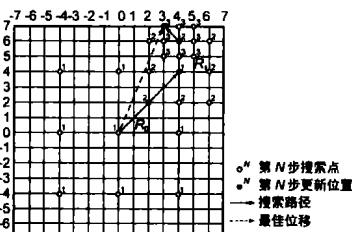


图 2 节点 MV 调整过程

致位移情况, 随步长减小各节点适应性地更新 MV , 逐步准确地得到最佳 MV . 步长逐次减半地变化, 保证节点可在 R 内自由移动, 且具有较高的预测精度.

NS-DBMA 将 $2K$ 维空间搜索分离为 K 个二维空间的搜索, 同时保持了节点相关性, 成功将搜索法应用在可变形块匹配中, 简化了匹配算法, 大大减小了运算量. 图 3 给出了算法流程图.

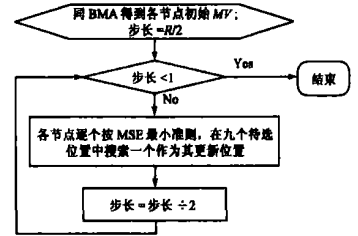


图 3 NS-DBMA 流程图

以四节点模型为例, 共 $L = \lfloor \log_2 R + 1 \rfloor$ 轮节点调整, 每轮调整 4 个节点, 每个节点搜索 9 个位置, 各节点在第一轮需计算 9 点, 以后只需计算 8 点. 总的计算点数为 $(32L + 4)$. 对 16×16 块, 调整范围 $R = 15$, 共 132 个搜索点, 每点为一次内插运算和一次误差运算, 运算量大约为 70000 次标量乘法和 65000 次标量加法, 约为 GB DBMA 的一半, 且内插和误差运算都适合由硬件模块高效执行.

3.3 分数像素精度预测

分数精度运动估计用于更精确的运动表示, 但要增加相当的运算量. $1/K$ 像素精度预测, 相当于在 K 倍内插参考帧中的整数精度预测. 以半像素精度预测为例, GB DBMA 的收敛性和收敛速度更加难以控制, 迭代次数大大增加, 运算量在 2 倍以上. 而在 NS DBMA 中, 只需将算法结束步长改为 $1/2$ 像素. 按上面分析, 总的搜索步数为 $L = \lfloor \log_2 R + 2 \rfloor$, 当 $R = 15$ 时, 运算量只提高约 25%. 如果需要更精确的运动估计, 可将算法结束步长改为 $1/4$ 像素甚至 $1/8$ 像素, 所提升的运算量也不大.

3.4 双模式混合方法

在实际的图像序列中, 平移模型在大部分情况下是有效的, 可变形块模型只需应用在平移模型不能适应的情况. 一般来说, 可把像块分为只做平移运动的平移块、做非平移简单形变的非平移块和发生复杂变化的不可运动补偿块. 双模式混合方法(TWO-MODE) 综合使用 BMA 和 DBMA, 如果像块的 BMA 预测误差小于某个预设的门限 T , 就认定为平移块, 不再做可变形块处理, 运动参数为一个 MV ; 否则由 NS-DBMA 作变形处理, 运动参数为各节点 MV ; 对不可运动补偿块, 也用 NS-DBMA 得到更好预测. 使用 TWO-MODE, 使图像中静止和平移区域不再做可变形块节点 MV 搜索, 运动参数也减为一个 MV , 既降低了运算量又增大了压缩率. 门限 T 可以看作是预测质量和运算时间的一个折衷因子, T 增大, 平移块数目增加, 则减少处理时间, 降低预测质量; 反之也成立. 结合分数像素精度预测, 可在低运算量下进一步提高预测质量.

4 测试结果和结论

测试 3 个场景内容不同的序列 Foreman、Paris 和 Mobile Calendar, 均为 CIF 格式(352×288), 预测亮度信号, 帧间距离

① 各块的调整算法一致, 省略说明哪个块正在被考虑.

为 2, 前向预测方式. 测试四节点模型, 块大小为 16×16 , 使用 MSE 准则, EBMA 搜索范围为 ± 16 , 节点调整范围为 ± 15 . 用软件实现了 EBMA、GB-DBMA、NS-DBMA、TWO-MODE、半像素精度 NS-DBMA (HP-DBMA) 和半像素精度 TWO-MODE (HP-TWO-MODE) 等算法. 为尽量保证 GB-DBMA 的收敛性, 使用 EBMA 作初始运动估计. 由于 DBMA 比 BMA 多 3 个 MV, 即 6 个运动参数, 我们还测试了加变换域 zig-zag 扫描的第 4 9 个系数作误差校正的 BMA (BMA-DCT) 的性能^①. 以峰值信噪比 (PSNR) 比较各算法的客观预测质量, 比较 NS-DBMA、HP-DBMA、TWO-MODE 和 HP-TWO-MODE 的运行时间^②.



图 4 Foreman 预测图像 (左上: 目标帧; 中上: EBMA; 右上: BMA-DCT; 左下: GB-DBMA; 中下: NS-DBMA; 右下: TWO-MODE)

图 4 为 Foreman 序列的一帧以及由各算法得到的预测帧. 可以看出 EBMA 的块效应很明显, 只在背景等平移部分预测效果较好; BMA-DCT 块效应依然存在; GB-DBMA 虽然能改善预测效果, 闭眼和颈部得到更好预测, 但并不总是有效的, 大部分区域没有得到更好预测, 整体效果不好; NS-DBMA 大大提高了预测质量, 图像各区域都得到很好预测, 视觉效果与目标帧相近; TWO-MODE 预测效果与 NS-DBMA 相近, 其中加框块为平移块. 可以看出, NS-DBMA 和 TWO-MODE 的主观质量是最好的.

图 5、图 6、图 7 分别给出了各测试序列的 PSNR 和运行时间曲线, 可以得到以下客观结论:

(1) NS-DBMA 比 BMA-DCT 平均提高 1.06dB, 比 EBMA 平均提高 1.48dB, 说明用附加比特表示节点位移比表示预测误差有效, 也说明可变形块模型比平移模型有效.

(2) GB-DBMA 不稳定, PSNR 改善的浮动较大, 预测质量均不及 NS-DBMA.

(3) 无论是整数精度还是半像素精度, TWO-MODE 的预测性能都很接近 NS-DBMA, 平均只差 0.12dB, 但平均运行时间只有 NS-DBMA 的 1/6.

(4) 半像素精度预测能有效提高预测质量, 较整数精度预测平均提高 0.69dB.

(5) NS-DBMA 运行时间很稳定(个别点跳变由 PC 运行不稳定造成), 说明该算法具有固定的计算数.

(6) HP-TWO-MODE 的性能时间比是最高的.

在测试中, 各种算法都以 EBMA 作初始运动估计, 这对 GB-DBMA 是必须的, 而 NS-DBMA 可采用其它快速匹配算法,

如 TSS 等. 在以 TSS 作初始运动估计的 NS-DBMA 测试中, PSNR 平均提高幅度更大, 达到 1.86dB, 比以 EBMA 为基础的高 0.38dB, 而 TSS 的运行时间不及 EBMA 的 1/10, 大大降低了整体运算量.

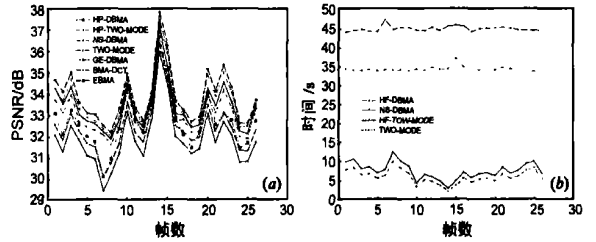


图 5 Foreman 序列各种算法的 PSNR 和运行时间曲线

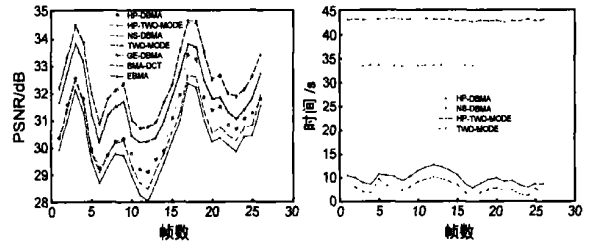


图 6 Paris 序列各种算法的 PSNR 和运行时间曲线

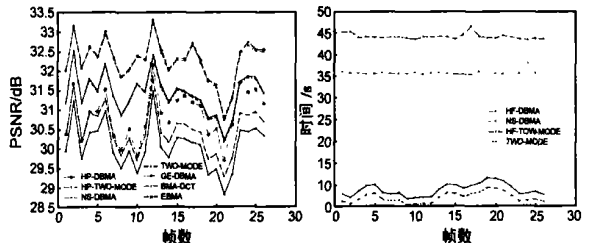


图 7 Mobile Calendar 序列各种算法的 PSNR 和运行时间曲线

对于可变形块模型, 可以采用较大的块做运动估计. 块越大, 做变形运动的可能性越大, 越适合可变形块模型; 但四节点模型的块也不宜过大. 对于更大的块, 可采用更多节点模型作预测, 如 9 节点模型. 测试中块大小为 16×16 , 减为 8×8 并不能带来预测质量的很大提高, 运算时间却要上升 4 倍. 在不全面测试中, 16×16 块 NS-DBMA 得到比 8×8 块 EBMA 更好的预测质量, 说明可变形块模型比平移模型更适合应用在运动补偿中.

如前所述, TWO-MODE 中门限 T 的选择将影响预测质量和运算时间. 选择 T 的原则是在预测质量下降不大的前提下让平移块尽量多. 设门限阈值 $T = 255 \times \text{BlockSize} \times p$, 其中 BlockSize 为块大小, p 为门限参数. 经过对多幅图像的综合测试分析, 所选 $p = 0.2$, 即阈值 $T = 255 \times 16 \times 16 \times 0.2$. 在图 4

① 在此不深入讨论变换编码、熵编码等内容, 并假设运动参数和 DCT 变换系数为同样的编码位数. 这样假设并不准确, 目的是在近似相同条件下比较 DBMA 和 BMA 的性能. 使用完整的 MPEG 标准可得到准确的结论.

② 测试用 PC 配置为赛扬 633MHz + 128M SDRAM, 软件环境为 Windows 2000 + Matlab 6.5, 不同配置机器上的运行时间会有很大差别. 运行时间不包括 EBMA 的处理时间.

(右下)中可以看出,背景和运动平缓区域多为平移块.图8给出了各序列平移块的数目(CIF格式块总数为396),平均为323块,约占81%.我们认为一般的序列也有同样的属性,即图像背景和运动平缓区域占大部分,可用平移模型

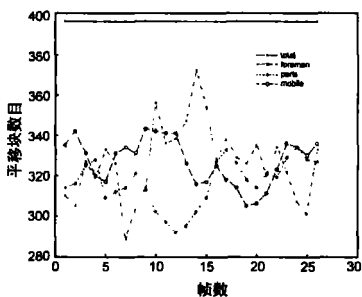


图8 各序列中平移块的数目

来预测;局部运动复杂和细节区域,应该采用更优越的可变形块模型进行预测.TWO-MODE的运算量约在EBMA基础上提高10%,约在TSS的基础上提高1倍,以现在的硬件处理速度是不难实现实时处理的.

5 结论

本文提出了NS-DBMA,并在此基础上提出分数精度预测和双模式混合方法,有效地对非平移像块进行预测,弥补了BMA的不足.测试结果表明,NS-DBMA比传统BMA有更好的主观预测质量,比GB-DBMA简单且效率更高,其中以半像素精度的双模式混合方法的综合性能最好.通过运算量和运算过程的简单分析,说明NS-DBMA可应用在实际编码系统中.

参考文献:

[1] Yao Wang. 视频处理与通信[M]. 侯正信,等.译.北京:电子工业出版社,2003.127-130.
 [2] Lee, O, Y Wang. Motion compensated prediction using nodal based deformable block matching[J]. Journal of Visual Communications and

Image Representation, 1995, 6(1): 26-34.

[3] O C Zienkiewicz, R L Taylor. The Finite Element Method[M]. Prentice Hall, Englewood Cliffs, NJ, 1989.
 [4] P E Gill, et al. Practical Optimization [M]. New York: Academic Press, 1981.
 [5] Lee O, Y Wang. A deformable block matching algorithm for motion compensated prediction[A]. Proceedings of Picture Coding Symposium [C]. Sacramento CA Sept, 1994. 223-226.
 [6] T Koga, K Iinuma, A Hirano, Y Iijima, T Ishiguro. Motion compensated interframe coding for video conferencing[A]. Proceedings of National Telecommunications Conference[C]. New Orleans, LA. 1981. C9. 6. 1-9. 6.5.

作者简介:



魏伟男,1982年生,天津大学硕士研究生,主要研究方向为数字视频编码与压缩.wei_wei_ah@163.com.



侯正信男,1945年生,博士生导师,天津市图像图形学会理事,现任天津大学电子信息工程学院教授、电视与图像信息研究所所长、学术带头人,主要研究方向包括现代数字信号与信息处理理论、数字图像处理与编码技术、数字视频处理与通信等.