

文章编号:1001-9081(2008)05-1244-03

一种改进的快速多分辨率运动估计算法

张素文, 杨富森, 汪丽丽

(武汉理工大学 自动化学院, 武汉 430070)

(wanglili0201@163.com)

摘要: 提出一种改进的快速多分辨率运动估计算法, 利用小波的多分辨率特性和小波匹配误差特性, 在部分失真消元(PDE)算法的基础上, 通过改进搜索顺序、匹配顺序和比较间隔来提高运动估计的速度。实验结果表明: 该算法能达到全搜索算法(FSA)的精度, 同时有效降低了计算复杂度。

关键词: 多分辨率; 运动估计; 部分失真消元; 小波匹配误差

中图分类号: TP391.41 **文献标志码:** A

Improved algorithm on fast multi-resolution motion estimation

ZHANG Su-wen, YANG Fu-sen, WANG Li-li

(School of Automation, Wuhan University of Technology, Wuhan Hubei 430070, China)

Abstract: In this paper, an improved algorithm on fast multi-resolution motion estimation was proposed, which made use of the multi-resolution property and wavelet matching error characteristic. Based on Partial Distortion Elimination (PDE) algorithm, we improved the speed of motion estimation by improving searching order, matching order and comparison interval. Experimental results show that the proposed algorithm can achieve the same estimate accuracy as Full Search Algorithm (FSA), while the computation complexity is reduced extremely.

Key words: multi-resolution; motion estimation; Partial Distortion Elimination (PDE); wavelet matching error

0 引言

运动估计与补偿已成为提高视频压缩编码效率的主要影响因素。传统的运动估计通常是在空域下进行, 但基于空域的运动估计存在以下不足: 1) 估计运算量大; 2) 存在块效应; 3) 编码效率低。离散小波变换(Discrete Wavelet Transform, DWT)比传统的基于块和离散余弦变换(Discrete Cosine Transform, DCT)的混合视频编码方案大大提高了编码效率, 使得小波分析技术在视频编码中占据着越来越重要的位置。随着文献[1]将小波引入运动估计, 基于小波域的运动估计方法体现出巨大的优越性。首先在最高级上的子带上进行多分辨率运动估计, 然后在递减的搜索区域中, 用先前层中获得的运动向量提炼剩余子带中的运动向量作为初始估计。

在此基础上, 小波域运动估计研究了很多成果。文献[2]提出了基于半像素精度或 $1/2$ 像素精度的多分辨率运动估计。比如在半像素精度中, 在执行运动估计前运用双线性插值。文献[3]提出了三种改进方法, 即自适应阈值多分辨率运动估计(Adaptive threshold Multi-Resolution Motion Estimation, AMRME)、双向多分辨率运动估计(Bi-directional Multi-Resolution Motion Estimation, BMRME)和快速多分辨率运动估计(Fast Multi-Resolution Motion Estimation, FMRME), 来提高小波视频编码器中的多分辨率运动估计编码的性能。在AMRME中, 如果参考帧与最佳匹配块间的差值大于设定阈值, 则这个块就被丢弃, 这种方法能提高搜索最佳匹配块的速度; 对于BMRME, 在先前或将来帧中都能找到最佳匹配

块; 在FMRME中, 把所有方向子带中的高频率子带, 即LH、HL和HH结合到一起进行MRME, 而不是对每个子带进行MRME。文献[4]提出了一种增强的MRME方法, 即在某个区域很可能存在运动的地方进行MRME, 通过此途径来减少递减搜索区域的搜索时间。文献[5]提出了一种两步MRME, 第一步进行传统的MRME; 在第二步中, 在更高频率子带中构造一个四叉树来获得更精确的运动向量, 同时减少统一区域的运动信息, 比如低频子带。文献[6]提出了一种在执行三维频率编码前进行运动补偿的方法, 通过很好地挖掘出运动补偿的时域相关性来获得高能量压缩。这些方法只是利用小波的多分辨率特性, 很好地提高了运动估计速度和精度。

除此之外, 文献[7]提出了基于小波的簇像素匹配误差部分失真(Clustered Pixel Matching Error Partial Distortion Search, CPME-PDS)算法, 利用小波金字塔分层结构中的簇像素匹配误差特性^[8], 在最高分辨率上子带中找到的自适应索引集可用于较低分辨率子带。该算法使用零运动向量作为运动预测器来决定最高分辨率级上的自适应索引集, 在先前操作中获得的索引集模式用于预测其他级上的排列顺序, 可以较早地排除不可能候选块, 提高了运动估计的速度。但是这种基于行的CPME-PDS方法丢弃了列上的系数, 损失了精度。因此本文在部分失真消元(Partial Distortion Elimination, PDE)算法的基础上, 利用小波匹配误差特性和多分辨率特性, 提出了一种改进的快速多分辨率运动估计算法, 通过改进搜索顺序、匹配顺序和比较间隔来加快估计速度, 并保证了估

收稿日期: 2007-11-21; 修回日期: 2008-01-21。

作者简介: 张素文(1963-), 女, 湖北武汉人, 副教授, 主要研究方向: 模式识别与智能系统、图像处理; 杨富森(1981-), 男, 湖北宜昌人, 硕士研究生, 主要研究方向: 图像处理与计算机视觉、视频压缩编码; 汪丽丽(1984-), 女, 湖北荆门人, 硕士研究生, 主要研究方向: 智能控制与智能自动化。

计精度。

1 PDE 算法

PDE 算法基本思想是:假设失真可由 K 个部分失真累加而得,如果第 i 次的累加部分失真大于当前的最小块误差 (Minimum Block Distortion, MBD),则该搜索点肯定不是匹配点。这样就可以省去计算剩余的部分失真的计算消耗。PDE 算法的搜索速度取决于根据累加部分失真可以多快地终止计算。经典的部分失真块运动估计算法主要以下两种:归一化部分失真搜索算法 (Normalized Partial Distortion Search Algorithm, NPDS) 和分级部分失真搜索算法 (Partial Distortion Search Algorithm, PPDS)。NPDS 在计算块失真时采用了一种中途停止技术来降低技术复杂度,而 PPDS 在其基础上使用了分级部分失真,比如第一个部分失真可以进一步分成几个更小尺度的部分失真。PDE 算法是在全部计算绝对差和 (Sum of Absolute Difference, SAD) 之前使用部分误差和消除不可能候选,如式(1)所示:

$$PSAD_m(p, v) = \sum_{n=1}^{m \leq N \times N} |S_{t,p}(O(n)) - S_{t',p+v}(O(n))| \geq SAD_{\min}(p) \quad (1)$$

其中 $SAD(p, v)$ 是 v 平移向量的第 p 个小波块的 SAD, 定义如式(2):

$$\begin{aligned} SAD(p, v) &= \sum_{i=1}^{N/2^L} \sum_{j=1}^{N/2^L} |S_{t,L,0,v}(i, j) - S_{t',L,0,p+v}(i, j)| + \\ &\quad \sum_{l=1}^L \sum_{k=1}^3 \sum_{i=1}^{N/2^l} \sum_{j=1}^{N/2^l} |S_{t,l,k,v}(i, j) - S_{t',l,k,p+v}(i, j)| \end{aligned} \quad (2)$$

其中 $S_{t',l,k,v}(i, j)$ 代表相对参考帧 t' 移动 v 像素的平移不变小波系数,如式(3):

$$S_{t',l,k,v}(i, j) = f_{t',l,k,v}(\lfloor dx/2^l \rfloor, \lfloor dy/2^l \rfloor, i + \lfloor dx/2^l \rfloor, j + \lfloor dy/2^l \rfloor) \quad (3)$$

这里 l 是小波分解的层数, k 表示 LL/LH/HL/HH 子带类型, $v = \{dx, dy\}$ 表示移动向量, $(\lfloor dx/2^l \rfloor, \lfloor dy/2^l \rfloor)$ 表示在 l 层中移动的数目, $(i + \lfloor dx/2^l \rfloor, j + \lfloor dy/2^l \rfloor)$ 是子带 k 的位置。当前帧 t 的小波系数可表示为 $S_{t,l,k}(i, j) = f_{t,l,k}(i, j)$ 。

2 小波匹配误差特性

为了提高 PDE 的计算节省量,如果在搜索窗口 W 中的匹配误差 $d_{p+v}(i, j)$ 的期望满足式(4):

$$E[d_{p+v}(i', j')] \geq E[d_{p+v}(i'', j'')] \quad (4)$$

其中 $d_{p+v}(i, j) = |S_{t,p}(i, j) - S_{t',p+v}(i, j)|$,然后假设这个匹配顺序对所有优选向量都是有效的。因此,为了满足式(4),必须用 $e_p(i, j)$ 预测误差。预测匹配误差定义如式(5):

$$e_p(i, j) = \arg \max_{e_p(i, j)} E[((e_p(i, j))^2 - (d_{p+v}(i, j))^2)^2] \quad (5)$$

因为小波系数服从拉普拉斯分布,可令 $E[S_{t',p+v}(i, j)] \approx 0$,因此可以得到:

$$\begin{aligned} e_p(i, j) &= \sqrt{E[(S_{t,p}(i, j) - S_{t',p+v}(i, j))^2]} \approx \\ &\quad |S_{t,p}(i, j)| + E[|S_{t',p+v}(i, j)|] \end{aligned} \quad (6)$$

从式(6)可以看出,预测匹配误差 $e_p(i, j)$ 包括两部分:

一部分是当前块的小波系数幅值 $|S_{t,p}(i, j)|$;另一部分是在参考块中小波系数幅值的期望值 $E[|S_{t',p+v}(i, j)|]$ 。因为在搜索窗中小波系数幅值的期望变化平缓,所以可将第二部分视为常量。因此,预测匹配误差与当前块中的小波系数幅值成比例。也就是说,在当前块中较大的小波幅值趋向于产生较大的匹配误差。这个结论与边缘和纹理对 SAD 匹配误差贡献最大,高幅值的小波系数通常对应于边缘和纹理这样的事实是一致的。

3 提出的算法

从式(1)可以看出影响 PDE 性能主要有以下 3 个因素:

- 1) 搜索顺序,即在搜索阶段测试小波块。如果 SAD_{\min} 发现得越早,则接下来的许多测试将会被跳过。
- 2) 匹配顺序,即在小波块中挑选系数计算 SAD。如果很早就得到 SAD 的最大值,则部分绝对差和 (Partial Sum of Absolute Difference, PSAD) 就会被终止。
- 3) 比较间隔,即 PSAD 与 SAD_{\min} 之间的比较。如果能在比较花费与无用差分的数目之间能达到一个较好的平衡,则比较和差分的次数都会减少。

因此本文在 PDE 算法基础上,利用小波匹配误差特性和多分辨率特性,提出了一种改进的快速多分辨率运动估计算法,通过改进基于小波多分辨率特性的搜索顺序、基于小波树组合机制的匹配顺序和基于自适应子块检查单元的比较间隔,在保证估计精度下约束运动估计时间。

1) 基于小波多分辨率的搜索顺序策略。小波系数仍然包含空间信息,LL 子带实际对应于原始帧的最低分辨率。因此,一个小块中的 LL 子带的运动信息与整个小波块紧密相关。利用这一特性,制定一个新的搜索顺序可以得到好的匹配。螺旋搜索是通过预测向量的统计分布,从搜索窗口的中心开始的螺旋向外的搜索方法。与这种方法不同,本文提出的搜索顺序使用 LL 子带中的归一化部分 SAD 作为估计绝对差和 (Estimation Sum of Absolute Difference, ESAD), 如式(7):

$$SAD_{LL}(p, v) = \sum_{i=1}^{\frac{N}{2^L}} \sum_{j=1}^{\frac{N}{2^L}} |S_{t,L,0,v}(i, j) - S_{t',L,0,p+v}(i, j)| \quad (7)$$

$$ESAD(p, v) = \frac{1}{2^L} \cdot \frac{1}{(\frac{N}{2^L})^2} SAD_{LL}(p, v) = \frac{2^L}{N \times N} SAD_{LL}(p, v) \quad (8)$$

令 $ESAD \in [0, 255]$,这样可以使用计数分类算法^[9] 得到搜索顺序 $SO = \{v_n | n = 0, \dots, W-1\}$,如式(9) 所示:

$$ESAD(p, v_0) \leq \dots \leq ESAD(p, v_n) \leq \dots \leq ESAD(p, v_{W-1}) \quad (9)$$

2) 基于小波树组合机制的匹配顺序策略。因为在较高分解层上,小波系数幅值通常较大,匹配误差也趋向于较大。因此,匹配顺序是从较高分解层到较低分解层。除此之外,具有较小幅值的重要小波系数往往会出现簇。通过观察发现,这种成簇特性也出现在匹配误差中。也就是说,在同级中具有较小幅值的匹配误差往往会出现簇。因此,本文提出

了一种小波树组合机制,即通过空间自相似特性和小波的匹配误差成簇特性把系数组合到子块中。小波块中的小波树组合机制如图 1 所示。

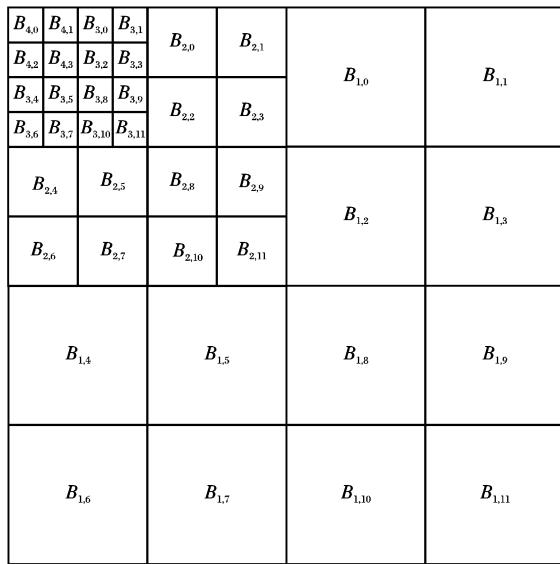


图 1 小波块中的小波树组合机制

为了确定子块的匹配顺序,每个子块 $B_{l,b_{l,m}}$ 的估计匹配误差为:

$$E_p(B_{l,b_{l,m}}) = \sum_{(i,j) \in B_{l,b_{l,m}}} |S_{t,l,b_{l,m}}/4, p(i,j)|; m = 0, \dots, M - 1 \quad (10)$$

其中 $b_{l,m}$ 是 l 层中每个子块的索引。然后以降序来整理估计的匹配误差,获得 l 层的匹配顺序。因为 M 的值小,且等于 12,所以用复杂度为 $O(n \log n)$ 的快速分类算法^[9]。

3) 基于自适应子块检查单元的比较间隔。在传统的 PDE 方法中,通常采用固定的比例间隔,比如 8 像素或 16 像素检查单元。因为在小波块中,在较高分解层中的匹配误差往往较大,所以在较高分解层中选择一个较小的比例间隔,而在较低分解层中选择一个较大的比例间隔。结合小波树组合机制,在比较花费与无用差分之间平衡,本文提出了一种自适应的比例间隔方法。在 l 分解层中每 2^{l-1} 子块作为一个检测单元,也就是在 PSAD 和 SAD_{min} 间对每个 2^{l-1} 子块进行比较。这种方法的优点在于在少量的差分步骤后就可以达到失真限制,而同时将比较成本维持在一个合理的层时就停止计算差分的部分和。

4 实验结果与分析

硬件平台是 PIV 3.0 G, 512 MB 内存计算机, 软件采用 Visual C++ 6.0 编程环境。实验中采用 Daubechies9/7 双正交小波, 分解层数 $L = 3$, 小波块尺寸是 16×16 , 搜索的窗口分别是: 对于四分之一通用中间格式 (Quarter Common Intermediate Format, QCIF) 采用 $[-15, 15]$, 对于通用中间格式 (Common Intermediate Format, CIF) 采用 $[-31, 31]$ 。将提出的算法与全搜索算法 (Full Search Algorithm, FSA) 算法和 CPME-PDS 算法进行性能比较。

从表 1 可以看出, 在每帧执行时间的加速比上, 提出的改进算法比 CPME-PDS 提高了大约 60%, 并保证了与 FSA 相同的精度。实验结果说明提出的算法较 FSA 算法在运动估计

性能上有较大的提高, 但同时也看到增加了额外的计算负载, 较 CPME-PDS 算法是可以接受的。

表 1 本算法与其他算法的每帧执行时间比较

视频 格式	视频 名称	FSA		CPME-PDS		本文算法	
		执行时 间/ms	加速比	执行时 间/ms	加速比	执行时 间/ms	加速比
QCIF (176 × 144)	Claire	767	1.00	76	10.09	48	15.98
	Suzie	821	1.00	112	7.33	79	10.39
CIF (352 × 288)	Akiyo	13943	1.00	656	21.25	286	48.75
	Container	13924	1.00	1567	8.89	1072	12.99

5 结语

本文利用小波匹配误差特性和多分辨率特性, 基于 PDE 算法, 提出了一种改进的快速多分辨率运动估计算法, 通过改进基于小波多分辨率特性的搜索顺序、基于小波树组合机制的匹配顺序和基于自适应子块检查单元的比较间隔, 在保证估计精度下约束运动估计时间。运动估计时间的减少将加快视频传输的速度, 尤其适用于视频会议、可视电话等实时性要求较高的场合, 既有质量上的保证, 又在传输速度上有了提高。实验结果显示该算法能极大地减少全搜索算法的计算复杂度, 并在保证估计精度的同时提高了传统 PDE 算法的计算效率。

参考文献:

- ZHANG Y Q, ZAFAR S. Motion-compensated wavelet transform coding for coding video compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1992, 2(3): 285 – 296.
- ZAFAR S, ZHANG Y Q, JABBARI B. Multiscale video representation using multiresolution motion compensated and wavelet decomposition[J]. IEEE Journal on Selected Areas in Communication, 1993, 11(1): 24 – 35.
- MANDAL M K, CHAN E, WANG X, et al. Multiresolution motion estimation techniques for video compression[J]. Optical Engineering, 1996, 35(1): 128 – 136.
- WEI JIE, LI ZE-NIAN. An enhancement to mmrc scheme in video compression[J]. IEEE Transactions on Circuits and Systems for Video Technology, 1997, 7(3): 564 – 568.
- KIM S, RHEE S, JEON JG, et al. Interframe coding using two stage variable block-size multiresolution motion estimation and wavelet decomposition[J]. IEEE Transactions on Circuits and System for Video Technology, 1998, 8(4): 399 – 410.
- OHM J R. Three-dimensional subband coding with motion compensation[J]. IEEE Transactions on Image Processing, 1994, 3(5): 559 – 571.
- WONG K Y, SIU W C, HUI K C. Fast motion estimation for wavelet-based video coding[C]// Proceedings of the 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing (ISIMP 2004). Washington, DC: IEEE Computer Society, 2004: 398 – 401.
- HUI K C, SIU W C, CHAN Y L. New adaptive partial distortion search using clustered pixel matching error characteristic[J]. IEEE Transactions on Image Process, 2005, 14(5): 597 – 607.
- MAHMOUD H M. Sorting: A distribution theory[M]. New York, NY: Wiley, 2000.