

基于小波变换的块匹配运动估计方法*

龚涛, 丁润涛

(天津大学电子信息工程学院, 天津 300072)

摘要 将小波变换引入低码率视频图像序列的运动估计研究中, 提出了基于小波变换的块匹配搜索方法. 该方法的特点是根据运动物体的大小和运动程度, 用可变宏块进行运动估计, 统一了匹配块和匹配点的搜索, 发展了一种改进的三步搜索算法. 实验结果表明, 该算法比三步法有更高的精度, 更适应低码率和多细节视频图像传输中的运动估计.

关键词 运动估计; 小波变换; 全搜索; 三步法

中图分类号: TN941.1 文献标识码: A 文章编号: 0493-2137(2003)03-0320-05

A Block Matching Motion Estimation Algorithm Based on Wavelet Transform

GONG Tao, DING Run-tao

(School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China)

Abstract The wavelet transform is introduced to low bit video sequences motion estimation. A block matching searching algorithm based on wavelet transform is proposed. The characteristic of the method is that the motion estimation is performed using convertible block according to the size and velocity of the moving object. It unifies the search of matching blocks and points. A modified three steps searching algorithm is developed. The simulation results show that this algorithm has higher precision than three steps searching method and is more suitable for motion estimation of low bit and more detail video transmission.

Keywords motion estimation; wavelet transform; full search; three steps search

运动估计^[1], 也称位移估计, 是一种有效降低图像序列时间(帧间)冗余度的方法. 运动估计的方法很多, 其中全搜索(full search, 简称FS)方法的精度很高, 但计算量过大. 由统计学理论可知, 运动矢量大多集中于搜索窗的中心附近, 因为在视频序列中, 静止块和小运动块的数量很多, 同一对象的宏块在运动中保持一致的概率远大于同一帧中所有宏块在运动中保持一致的概率. 这说明匹配宏块在一般运动情况下主要位于临近的宏块. 基于上述性质提出的块匹配方法对精度和复杂度进行了很好的折中而被广泛应用. 在搜索区域和匹配标准相对确定的情况下, 一些非最优的块搜索方法, 如三步搜索(three steps search, 简称

TSS)、交叉搜索(crisscross search, 简称CS)^[1]和菱形搜索^[2]等均不同程度地降低了复杂度. 这些方法有个共同点, 即先固定宏块尺寸(如MPEG-4为 16×16), 然后用邻近的宏块确定匹配宏块. 固定宏块尺寸的方法有一定的局限性, 不同尺寸的物体或含有不同程度细节的物体, 宏块尺寸应有所区别. 例如, 在MPEG-4标准中, 如果运动物体很小, 这样的大宏块很容易淹没物体边缘和内部的某些细节, 导致匹配失败或细节丢失. 另外, 对于剧烈运动, 只用邻近宏块分析容易陷入局部最优而得不到全局最优的运动矢量. 能否既兼顾全局又有效地降低搜索点数成为运动估计研究的一个热点.

* 收稿日期: 2002-05-08.

作者简介: 龚涛(1977—), 男, 硕士.

1 小波变换的性质

小波变换可得到不同频带(子带)下的子图像,其中低频平滑子图像保留了原始图像重要的视觉和统计特性^[3]. 首先,小波变换成比例地降低图像尺寸,例如原始图像中 16 × 16 的宏块,经过一级小波变换对应成 8 × 8 的宏块;子图像中 8 × 8 个点是对原始图像 16 × 16 个点特性的平均体现. 通过对低频子带图像的分析,可以获得上一级图像的平均统计特性. 同时,对子图像“点”的分析,实际上是对上一级图像“区域”的分析,考虑了全局特性,有效扩大了搜索范围. 其次,三步搜索和菱形搜索等非最优方法可以直接应用于小波子图像而不必在局部扩大步长. 分析子图像就等价于在全局范围内扩大步长,而且估计子图像得到的最优点可在上一级图像中进一步优化得到全局最优点. 最后,小波变换可以降低分辨率,对于低码率传输(如 H. 263 标准)有重要意义,而其它运动估计方法直接估计原始图像,很容易产生计算冗余.

基于小波变换的上述特点,作者提出了一种小波变换的 TSS 来降低搜索点数的运动估计方法. 为此,将分别介绍小波变换的搜索算法及其实验结果分析.

2 基于小波变换的块匹配搜索算法

本文所采用的小波是幅频特性不变(便于计算运动矢量)的 2 维离散可分离小波(DWT). 其子图像定义如下:

$$A_{j+1}^n(f, m, n) = \sum_k \sum_l A_j^n(f, k, l)h(k-2m)h(l-2n)$$

$$D_{j+1}^1(f, m, n) = \sum_k \sum_l D_j^1(f, k, l)h(k-2m)g(l-2n)$$

$$D_{j+1}^2(f, m, n) = \sum_k \sum_l D_j^2(f, k, l)g(k-2m)h(l-2n)$$

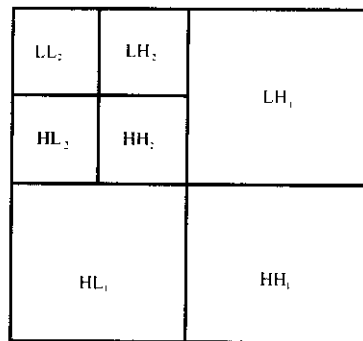
$$D_{j+1}^3(f, m, n) = \sum_k \sum_l D_j^3(f, k, l)g(k-2m)g(l-2n)$$

其中 $h(n)$ 和 $g(n)$ 是 1 维小波基的低通和高通滤波器. 以上 4 个输出分别为 LL_j 、 LH_j 、 HL_j 和 HH_j .

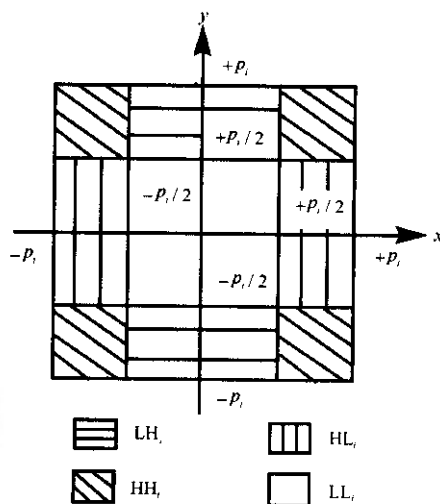
图像经过一次小波分解,就是对原图像做行和列滤波,并按 2:1 抽样,得到尺寸和频率减半的一系列子图像,2 次分解后子图像分布如图 1a 所示,频域分布见图 1b. 每次分解中 LL_j 表示的是低频平滑图像,该图像保留了原始图像基本的视觉和统计特性,其余的是细节图像. 小波变换是可逆的.

由于各级的 LL_j 子图像保留了原始图像的主要信息,运动矢量估计主要分析各级的 LL_j 子图像. 同时,各级子图像在数据向上存在着很强的相关性,如 LH_2

和 LH_1 都是原图像水平方向的细节. 所以在不同级运动估计时,应考虑同方向上的子图像. 另外,各级的运动矢量成比例,如第 2 级运动矢量是第 1 级的 1/2,这是由抽样造成的. 小波变换下的运动估计方案选取各级同方向子图像做运动估计. 这样可以在全局范围内估计出匹配块的位置,例如估计 LH_1 子图像中的块,可用 LH_2 中对应的小块来估计^[4].



(a) 空间分布



(b) 频域分布

图 1 2 次小波变换子图像分布

Fig. 1 Distribution of 2 level wavelet transform sub-images

作者根据物体运动程度,采用小波变换级数的运动估计方法决定匹配规则. 通常所用的匹配误差函数有很多,例如均方误差(MSE)和平均绝对误差(MAD)等. 本文使用 MAD 以免去复杂的平方运算,定义如下:

$$MAD(x, y) = \frac{1}{N \times N} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} |F_n(x, y) - F_{n-1}(i, j)| \quad -W \leq (x, y) \leq W$$

式中 F_n 为当前帧的图像块; F_{n-1} 为前一帧中的经过匹配后的图像块, 大小均为 $N \times N$; W 为运动矢量在水平和垂直方向上的最大分量; 搜索窗的大小为 $(2W + 1)^2$. 所采用的匹配准则为判断 MAD 是否位于给定误差容限 (EA) 内.

小波变换中各级块的尺寸由运动程度和小波变换级数共同决定. 为了准确得到运动矢量, 搜索窗内采用全搜索是必要的, 这就要求块的尺寸不宜过大, 以降低运算次数. 同时, 为了适应剧烈的运动, 要求搜索窗尽量大, 这样在子图像搜索时所跨的区域较大, 可以找到大运动的全局最优点而不至于陷入局部最优, 所以应根据经验和图像序列的具体运动情况规定宏块大小.

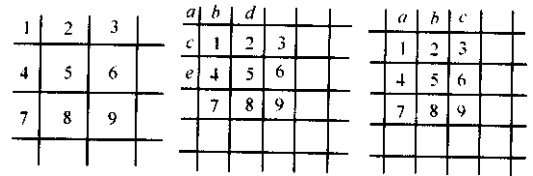
作者假设大部分运动矢量分量绝对值 ≤ 12 像素, 在原始图像中采用 4×4 的块, 以此为例, 分步骤详细描述算法, 然后再讨论剧烈运动情况.

步骤 1 给定 EA, 将连续 2 帧原始图像划分为 4×4 的小块, 取其中标为 a 的含有要预测点的块和另一帧中的对应块 a^1 . 以 a^1 块为搜索窗, 对窗里的 16 个点与预测点计算 MAD. 若找到了 MAD 小于 EA 的匹配点, 说明为微小运动, 每点最多计算 16 次即可找到匹配点, 算法结束; 若大于 EA, 说明物体运动较为剧烈, 进入第 2 步.

步骤 2 将 2 帧原始图像做 2 次小波变换, 同时调整 EA 降为 EA^1 以适应子图像的精度. 由于子图像是原图像的低频分量, EA 误差容限的精度也要适当的改变 (乘以小波系数). 这是在保证精确找到匹配点的前提下改变的, 不会引入额外误差, 同时也在一定程度上避免了小波系数带来的移变. 取 LL_2 子图像, 此时 4×4 的块 a, a^1 将收缩为点 b, b^1 , 取对应的 2 次小波变换的 LL_2 子图像, 以 b 为中心, 搜索范围为 ± 2 , 对应原图像的范围为 ± 8 , 用一种改进了的三步法^[4]搜索子图像上的点, 并与 b 点计算 MAD. 此三步法建立在单一平面假设上^[5,6], 认为搜索点远离全局最小误差匹配点时, 其匹配误差会随距离成线性增加. 根据此假设提出的算法效果很好. 三步法先按图 2a 将 b^1 周围的 9 个点标号, 再按图 3a 流程图, 取点和 b 进行 MAD 计算 (此时 $N=1$). 经过一轮计算, 找到 MAD 最小点, 若该点 MAD 小于 EA^1 , 标为 c , 进入第 3 步, 否则进入第 4 步.

步骤 3 找到小于 EA^1 的 c 点, 说明是中等程度的运动 ($4 \leq |$ 运动矢量分量 $| \leq 8$). 找到 c 点对应原图像中 4×4 的块. 即 LL_2 中的点 (m, n) 对应原图像的点 (k, l) , 其中 $4m - 3 \leq k \leq 4m + 4, 4n - 3 \leq l \leq 4n + 4$, 共 16 万方数据

个点. 找到匹配宏块后, 按照第 1 步的方法逐点搜索, 每点最多计算 16 次即可找到匹配点, 计算运动矢量, 结束算法.



(a) 编号点分布 (b) 角点情况 (c) 中间点情况

图 2 三步搜索时亚最优点分布

Fig. 2 Sub-optimum pixels distribution of TSS

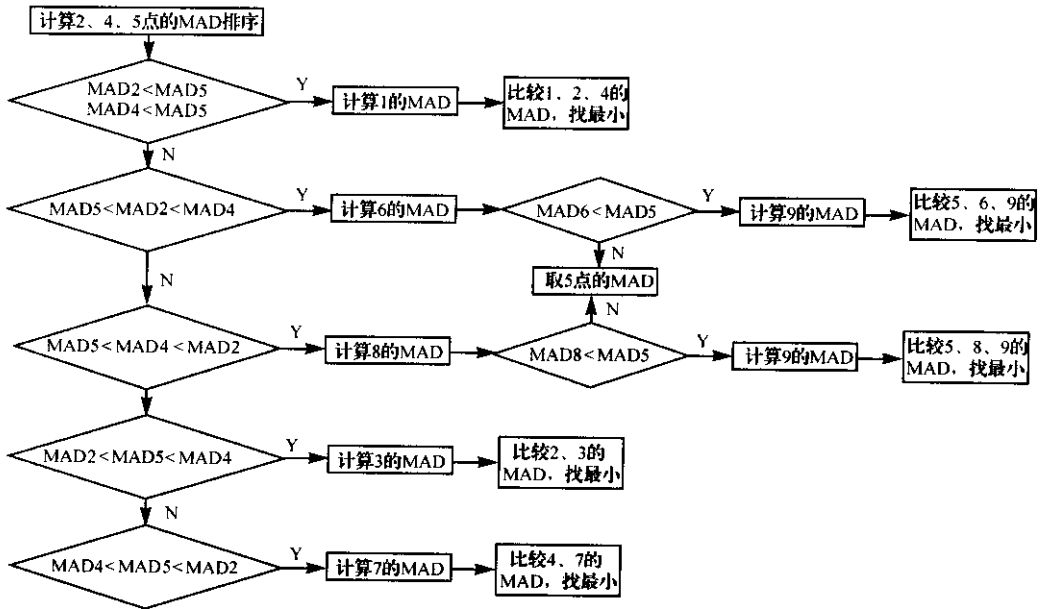
步骤 4 若经过对 LL_2 的处理没有找到匹配点, 只找到亚最优点 d , 说明为剧烈程度的运动 ($|$ 运动矢量分量 $| \geq 8$ 像素). 此时, 继续用第 2 步中的搜索方法搜索外围的点, 找到最优点 (小于 EA^1) 或亚最优点. 在第 2 步中找到的亚最优点的分布有两种情况, 为 1 (或 3、7 和 9) 这样的角点, 见图 2b, 此时搜索流程图如图 3b 所示; 为 2 (或 4、6 和 8) 这样的中间点, 见图 2c, 此时搜索流程图如图 3c 所示. 这样仍然能以较少的搜索点数快速找到最优点. 当找到最优点时, 标为 c , 回到第 3 步中继续计算, 找到匹配点.

经过以上 4 步, 对于多数运动矢量分量绝对值在 12 点之内的情况可以很好地找到匹配点. 下面看一下所用的搜索点数 (这里不考虑小波变换的点运算). 本方法在最极端的情况下, 一个点所用搜索点数: 第 1 步中用了 $4 \times 4 = 16$ 个点, 第 2 步中基本搜索最多搜索 5 个点, 若进一步搜索 (如第 4 步所示) 最多再搜索 4 个点, 共搜索了 9 个点找到匹配宏块; 步骤 3 最多搜索 16 个点找到匹配点. 由此可见, 可变宏块搜索最多搜索了 $16 + 9 + 16 = 41$ 个点, 最少用 16 个点就可以搜到匹配点. 考虑更一般的情况, 原始图像用 $N \times N$ 的块, M 次小波变换后, 块变为点 (或块, 此时可取块内点的灰度均值代替块, 即把块看成点), 则最多所用的搜索次数为 $N \times N + 9 + N \times N$, 可见搜索次数与 N 的大小有关.

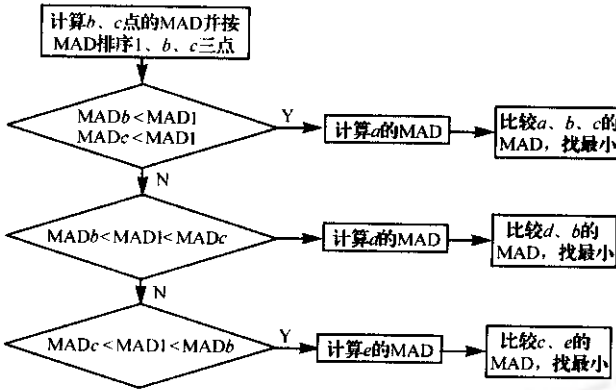
如果运动情况更为剧烈, 需要扩大原始图像中的宏块大小, 例如 8×8 , 则 2 次小波变换后块收缩为 2×2 , 搜索的运动矢量分量的范围扩大为 24 像素. 可多作一级小波变换, 再用上述三步法进行计算. 此时, 块内匹配时的计算量加大, 但所跨区域也变大, 可找到剧烈运动情况下全局最优点, 结果会更准确. 当原始图像块的大小增大到一定程度时, 现有的菱形搜索^[2]和三步

搜索等可应用于搜索窗内匹配点的搜索来降低搜索点数. 小波变换搜索匹配宏块的方法与原有的块匹配方

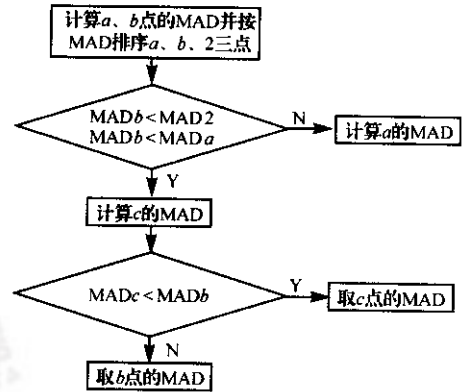
法是一致的, 小波变换实现了对点搜索与对块搜索的统一.



(a) 基本搜索算法



(b) 角点情况算法



(c) 中间点情况算法

图3 新三步法流程

Fig.3 Flow chart of new TSS algorithm

3 实验分析

Aki 序列用于分析运动物体不大、细节较多和运动剧烈的情况(见图4), 用 4×4 块划分原始图像进行计算. 作者给出全搜索算法 (16×16 块) 和菱形搜索算法 (16×16 块) 的计算结果进行比较(见表1). 其中运动矢量误差是以全搜索的结果为标准, 用分量最大误差来度量的. 可以看出, 在不考虑运算时间和增加搜索点数不多的情况下, 可以获得比菱形算法更高的精

万方数据

度, 适应了低码率和多细节图像传输的要求.

表1 算法性能的分析

Tab.1 Characteristics of algorithms

方法	全搜索 (通用 16×16 块)	三步法 (通用 16×16 块)	本文方法 (4×4 块)
平均每点搜索点数	255.00	13.89	60.83
平均每点均方误差	150.54	171.31	164.32
平均每点 MAD	12.26	13.00	12.67
运动矢量误差/ $^{\circ}$	以此为淮	2.34	1.99



图 4 Aki 序列示例

Fig.4 Aki image sequences

4 结 语

作者提出了宏块大小可变的概念,克服了现有固定宏块的搜索方式不考虑实际运动物体几何尺寸的弊端.引入小波理论,充分利用小波子图像在统计和空间上的相关性,在小波变换域实现了匹配宏块的搜索.这比现有的用中心点运动矢量代替整个宏块运动矢量来搜索匹配宏块的方法在统计上更合理.同时,应用小波理论把点的匹配搜索与块的匹配搜索统一起来,所述的三步法也可用于对匹配宏块内部点的匹配.实验结果表明,在不考虑运算时间的情况下,本方法可获得比一般块匹配方法更小的误差.

当然,运算量是本算法的一个瓶颈.但算法统一了块匹配与点匹配,模块化的算法具有很强的通用性,便于用 ANSI 或 DSP 专用芯片实现,加之小波变换快速算法的研究已取得很大进展,专用硬件电路的出现会弥补本算法的不足.

今后研究的重点将放在如下两个方面:

1)研究小波运算的快速算法,开发专用芯片实现小波变换,这不仅对运动估计领域,对其他小波应用领域也有重要意义;

2)虽然本算法中宏块是可变的,但运算时要人为预先规定宏块大小,规定原则是运动的剧烈程度和要求的细节度,目前这些指标主要靠先验经验来确定,能否量化这些指标由算法自动确定宏块尺寸,开发出自适应的运动估计算法是今后研究的方向之一.有了自适应算法,就可以根据需要选择地获得某些区域运动的细节,以用于可变码率图像传输领域中.

参 考 文 献 :

[1] Tekalp Murat A. *Digital Video Processing*[M]. USA : Prentice Hall PTR , 1995.

[2] Zhu S , Ma K K. A new diamond search algorithm for fast block matching motion estimation[A]. *International Conference on Information , Communication and Processing*[C]. Salt Lake City ,USA. 1997 , 292—296.

[3] 王 玲 , 应益荣. 小波在图像处理中的几个问题 [J]. 西北建筑工程学院学报 , 1998 , 3(2) : 66—77.

[4] 洪 波 , 庄建敏 , 余松煜. 基于时空相关性的自适应运动估计方法 [J]. 计算机工程 , 2000 , 8(2) : 50—53.

[5] 李小红 , 钱源成. 基于小波的运动估计方法的实现 [J]. 系统工程与电子技术. 2001 , 24(1) : 20—23.

[6] 施建良 , 潘 键 , 余松煜. 一种基于时空相关性的运动估计算法 [J]. 电子科学学刊 , 1999 , 3(2) : 349—354.