

一种运动估计的快速预测搜索算法¹

陆 阳 骆立俊* 邹采荣* 何振亚*

(合肥工业大学计算机与信息科学系 合肥 230009)

*(东南大学无线电工程系 南京 210096)

摘 要 本文提出了一种运动估计的快速预测搜索算法 (PSA)。该算法首先用当前块的三个邻近块运动矢量的线性加权来得到预测矢量, 然后以预测点为起始点, 采用 3×3 的搜索窗进行搜索步长为 1 的移动窗搜索, 直到搜索到达搜索域的边界或搜索的局部最小点位于搜索窗的中心时停止。该算法由于利用了序列图象的实际运动矢量与预测矢量之间距离的空间分布特性——中心偏置分布特性和时间上的相关特性, 并采用了中止判决准则, 可以明显地减少搜索次数。仿真表明这种算法减少了搜索范围和搜索次数, 提高了搜索效率, 降低了运动估计的计算复杂性。本文还详细地给出了 PSA 算法与其它常用快速算法的比较结果。

关键词 运动估计, MPEG2, 预测搜索算法

中图分类号 TN941

1 前 言

视频序列图象在时间上有很强的相关性, 利用运动估计和运动补偿技术可以有效地去除图象帧间冗余度, 实现高码率压缩比, 所以这种技术已广泛用于视频压缩的一些国际标准, 如 H.261, H.263, MPEG1, MPEG2, HDTV 中。在这些视频压缩国际标准中编码器的复杂性主要取决于运动估计算法。目前运动估计算法有很多种, 全搜索运动估计 (FS) 算法^[1] 由于精度高, 硬件易实现, 是应用最广泛的运动估计算法, 但它的计算量大, 特别是在大范围搜索时, 加大了硬件实现难度。为了减少 FS 算法的计算量, 出现了许多改进的运动估计算法, 如三步搜索法 (3SS)^[2], 共轭方向搜索法 (CDS)^[3], 二维对数搜索法 (LOGS)^[4], 交叉搜索法 (CS)^[5], 动态搜索窗调整搜索法 (DSWDS)^[6] 等, 这些运动估计算法虽然往往只能找到搜索区域的局部最优值, 但比起 FS 算法来说大大减少了运算量。在这些算法中 3SS 算法是最简单有效的, 也是最常用的, 但 3SS 算法在搜索区域为 $(\pm 8, \pm 8)$ 时初始搜索步长为 4, 对慢速运动块的估计来说太大, 所以出现了改进的三步搜索法 (N3SS)^[7], 四步搜索法 (4SS)^[8] 等。

本文以 MPEG2 编码标准为例, 提出了一种运动估计的快速预测搜索算法, 它的搜索位移为 $(-16, +15.5)$ 。该算法利用了序列图象运动矢量与预测矢量之间距离的空间分布特性——中心偏置分布特性和时间上的相关特性, 可以明显地减少搜索次数。本文在第 2 节详细阐述了预测搜索算法。在第 3 节的仿真分析中对几种不同的搜索算法, 如全搜索 (FS) 算法、三步搜索 (3SS) 算法、八步搜索 (8SS) 算法 (即相对应于在搜索区域 $(\pm 8, \pm 8)$ 上的四步搜索 (4SS) 算法) 在大范围搜索区域的性能进行了详细地比较。最后在第 4 节给出了本文的结论。

¹ 1996-12-20 收到, 1998-01-04 定稿

国家攀登计划 (神经网络) 重大项目和东南大学与非利浦公司消费电子领域合作项目资助课题

2 预测搜索算法

在 MPEG2 视频编码标准中, 运动估计以宏块为单位, 在参考帧中以相对应的宏块为中心的搜索区域内进行搜索, 寻找最佳匹配宏块, 得到相应的运动矢量。运动估计只对亮度信号进行, 色度信号借用亮度信号的估计结果。

衡量运动估计最佳匹配的准则有很多种, 如均方误差 (MSE)、归一化互相关函数 (NCCF)、平均绝对帧差 (MAD) 等, 分析表明各种匹配准则的性能差别不显著, 比较而言 MAD 的运算量最小, 便于硬件实现, 所以用得最多, 其定义为

$$\text{MAD}(i, j) = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N |S_k(m, n) - S_{k-1}(m+i, n+j)|, \quad (1)$$

其中 M 和 N 分别为宏块的水平和垂直尺寸 (以像素为单位); i, j 分别为宏块水平和垂直方向的偏移量, 在搜索区域内取值; $S_k(m, n)$ 为第 k 帧位于 (m, n) 的像素值。

MPEG2 的 Test Model 5^[10] 建议在整像素运动估计搜索完毕后进行半像素精度的全搜索, 即在规定的搜索区域内进行逐点搜索, 得到一个 MAD 最小的整像素点, 然后以该点为中心, 取其周围线性内插得到八个半像素点进行全搜索, 得到一个最佳匹配点, 本文的分析采用了半像素精度的全搜索。

在视频序列图象中, 物体运动是连续的, 所以描述物体运动的宏块的运动矢量在空间和时间上具有相关性。图 1 是两个图象序列采用 FS 算法做运动估计后运动位移统计平均分布直方图, 其中搜索区域为 $(\pm 16, \pm 16)$, "Miss America" 序列有 91 帧, "Table Tennis" 序列有 40 帧。从图 1(a) 和图 1(c) 中可以看出, 通常视频序列图象中物体运动范围和运动速度都不大, 所以宏块的运动矢量在搜索区域内不是均匀分布的, 图象序列 "Miss America" 和 "Table Tennis" 分别有 83% 和 81% 的运动位移分布在 5×5 的中心区域窗内, 分别有 52% 和 61% 的运动位移分布在 3×3 的中心区域窗内, 这些宏块可以认为是平稳或准平稳的, 而 3SS 算法在搜索区域为 $(\pm 16, \pm 16)$ 时第一步搜索采用的搜索窗步长为 8, 对运动块的运动矢量估计来说初始步长太大。4SS 算法在做运动估计中采用较小的初始搜索步长, 并对平稳或准平稳的宏块采用中止判决准则来减少搜索次数。在搜索区域为 $(\pm 16, \pm 16)$ 时 (3SS) 算法的搜索步长为 8, 4, 2, 1, 八步搜索 (8SS) 算法 (即相对应于在搜索区域 $(\pm 8, \pm 8)$ 上的四步搜索 4SS 算法) 的搜索步长分别为 2, 2, 2, 2, 2, 2, 2, 1。

在视频序列图象中, 描述物体运动的宏块的运动矢量在空间和时间上具有相关性, 当前帧中宏块的运动矢量, 不仅与上一帧中相应宏块的运动矢量有关, 而且与当前帧相邻宏块的运动矢量有关, 在大多数情况下宏块的运动矢量与帧内和帧间相邻宏块的运动矢量有着相近的大小和方向。

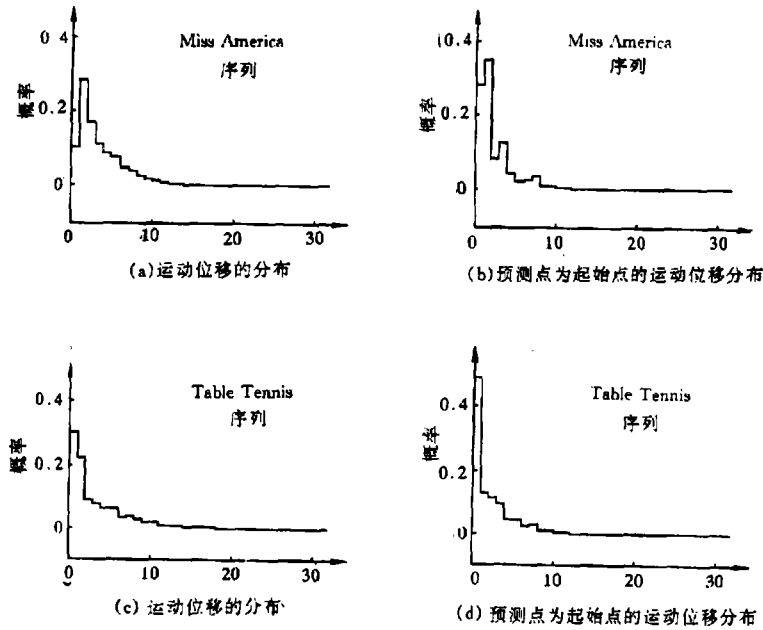


图 1 两个图象序列的运动位移分布直方图

图 2 是预测搜索 PSA 算法中用以预测的宏块的分布图。我们用三个邻近块 1, 2, 3 运动矢量的线性加权值来作为预测矢量, 可以得到图象序列 “Miss America” 和 “Table Tennis” 的预测矢量和实际的运动矢量之间距离的分布直方图分别如图 1(b) 和图 1(d) 所示。两个图象序列分别有 83% 和 82% 的运动位移分布在以预测点为中心的 3×3 区域窗内, 可以看出序列图象的实际运动矢量与预测矢量之间的距离具有很强的空间分布特性——中心偏置分布特性。

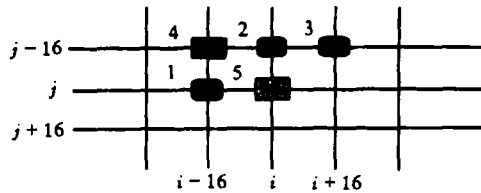


图 2 预测搜索 PSA 算法中用以预测的宏块的分布

预测搜索 PSA 算法首先利用当前帧内相邻宏块的已知运动矢量 (如图 2 中点 1, 2, 3, 4 所对应的当前帧的宏块) 和前一帧对应宏块的运动矢量 (如图 2 中点 5 所对应的前一帧的宏块) 来预测当前宏块运动矢量的搜索起始值, 然后以预测点为起始点, 采用 3×3 的搜索窗进行搜索步长为 1 的移动窗搜索, 直到搜索到达搜索域的边界或搜索的局部最小点位于搜索窗的中心时停止。在分析中发现, 点 4 的运动矢量可以用点 1, 2, 3 的运动矢量加权预测, 而点 5 对应的前一帧的运动矢量对当前宏块的预测影响并不大, 忽略它还可以减少 MPEG2 编码器中预测的复杂性。所以 PSA 算法采用点 1, 2, 3 所对应宏块的运动矢量来预测运动估计的搜索起始值。

令 $V_k = [V_{1k}, V_{2k}]^T$ 表示点 $k = 1, 2, 3$ 所对应的相邻宏块的运动矢量, $P = [P_1, P_2]^T$ 表示当前宏块运动矢量的预测值, $W = [W_1, W_2, W_3]^T$ 表示权值矢量, 当前宏块的中心点为 (i_0, j_0) , 预测点为 (i_1, j_1) , 其中

$$W = [W_1, W_2, W_3]^T = [1/3, 1/3, 1/3]^T, \quad P = [P_1, P_2]^T = \left[\sum_{k=1}^3 W_k V_{1k}, \sum_{k=1}^3 W_k V_{2k} \right]^T, \quad (2)$$

$$i_1 = i_0 + P_1, \quad j_1 = j_0 + P_2. \quad (3)$$

PSA 预测搜索算法的一种可能的搜索方式如图 3 所示, 它的具体步骤如下:

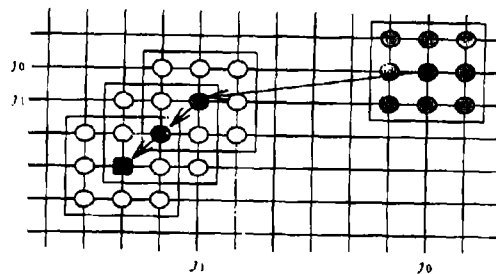


图 3 PSA 算法的一种搜索方式

步骤 1 依据式 (3) 计算运动矢量搜索初始值 (i_1, j_1) , 取搜索步长为 1, 计算以预测的初始点 (i_1, j_1) 为中心的 3×3 搜索窗上的 9 个搜索点所各自对应的 MAD 值, 以当前 MAD 值最小的点作为下一次搜索窗的中心点, 如果当前搜索窗的中心点为 MAD 最小的点, 中止搜索去步骤 3;

步骤 2 取搜索步长为 1, 计算以新的搜索中心点为中心 3×3 搜索窗上的 9 个搜索点所各自对应的 MAD, 然后以 MAD 最小的点作为下一次搜索窗的中心, 如果当前搜索窗的中心点为 MAD 最小的点, 中止搜索去步骤 3。在搜索过程中如果当前 MAD 最小的点位于搜索窗的四个角, 下一次的搜索只需要增加 5 个点, 否则的话下一次的搜索只需要增加 3 个点。步骤 2 重复直到搜索窗的中心到达区域 $(\pm 15, \pm 15)$ 的边界为止;

步骤 3 进行一次半像素精度的全搜索, 得到最终的运动矢量。

几种不同的运动估计搜索算法的搜索次数的比较如表 1 所示。

3 仿真结果

通过计算机模拟实验, 得出了 PSA 算法与 FS 算法、3SS 算法和 8SS 算法的对比实验结果, 实验在两种条件下进行:

(1) 在 MPEG2 的 MP@LL 进行, 测试序列为 Table Tennis(50 帧), Flower Garden(40 帧), 和 Football(40 帧), 图象格式为 352×240 , 色度格式为 $4:2:0$, 帧频为 25 Hz, 实验的压缩码率为 1.152Mbps; 测试序列为 Miss America(91 帧) 和 Claire(15 帧), 图象格式为 352×288 , 色度格式为 $4:2:0$, 帧频为 25Hz, 实验的压缩码率为 1.152Mbps;

(2) 在 MPEG2 的 MP@ML 进行, 测试序列为 Mobile & Calendar(33 帧), 图象格式为 720×486 , 色度格式为 $4:2:2$, 帧频为 25Hz, 实验的压缩码率为 5Mbps。

PSA 算法仿真采用的运动估计搜索区域为 $(\pm 16, \pm 16)$, 搜索位移为 $(-16, +15.5)$, 采用 MAD 准则。为了便于比较, 我们对 FS 算法、3SS 算法和 8SS 算法也进行了仿真, 仿真结果如表 1 和表 2 所示, 其中表 2 是采用不同运动估计搜索算法时原始帧和估计帧的均方误差 MSE 的平均值和解码后每帧的峰值信噪比 PSNR 的平均值的比较。从表 1 和表 2 中可以看出, 在大的搜索区域下, PSA 算法的性能比其它的快速算法都要好得多, 它的平均 MSE 是三者中最小的, 它的平均搜索次数也是最小的, 比 3SS 算法要少 20 或 21 次, 比 8SS 算法要少 7 或 8 次, 对应的运动估计所需要的时间也是最小的, 从表 1,2 中可以看出, PSA 算法非常明显地提高了运算速度, 复原图象质量与 FS 算法相比不相上下, 综合比较而言, PSA 算法的性能最接近 FS 算法。

表 1 几种不同的运动估计搜索算法的平均搜索次数的比较

序列图象	FS	3SS	8SS	PSA
Miss America	969	41	27.52	20.01
Claire	969	41	26.30	19.07
Flower Garden	969	41	29.32	22.57
Table Tennis	969	41	27.45	19.98
Football	969	41	26.56	19.36
Mobile & Calendar	969	41	29.35	21.87
平均	969	41	27.75	20.48

表 2 几种不同的运动估计搜索算法的平均 MSE 和平均 PSNR 的比较

运动估计算法 序列图象	FS		3SS		8SS		PSA	
	平均 PSNR	平均 MSE	平均 PSNR	平均 MSE	平均 PSNR	平均 MSE	平均 PSNR	平均 MSE
Miss America	41.69	10.13	41.59	11.08	41.60	10.58	41.69	10.45
Claire	43.82	6.98	43.78	7.73	43.77	7.87	43.81	7.46
Flower Garden	27.22	350.10	26.13	583.38	26.17	594.72	26.81	432.91
Table Tennis	30.85	122.73	30.39	151.67	30.46	153.59	30.46	155.17
Football	31.37	175.84	31.18	211.04	31.21	213.78	31.25	210.03
Mobile & Calendar	29.56	203.92	28.58	318.80	28.91	267.98	29.12	239.59
平均	34.09	144.95	33.61	213.95	33.69	208.09	33.86	175.94

4 结 论

本文提出了一种运动估计的预测搜索 PSA 算法。PSA 算法由于利用了序列图象的实际运动矢量与预测矢量之间距离的空间分布特性——中心偏置分布特性和时间上的相关特性, 并采用了中止判决准则, 可以明显地减少搜索次数。仿真表明 PSA 算法减少了搜索范围和搜索次数, 提高了搜索效率, 降低了运动估计的计算复杂性。在大的搜索区域下, PSA 算法的性能比 3SS 算法、N3SS 算法、8SS 算法要好得多, 它的运动估计平均 MSE 是三者中最小的, 解码后图象的平均峰值信噪比 PSNR 最接近 FS 算法的结果, 它的搜索次数也是最少的, 对应的运动估计所需要的时间也是最短的。综合比较而言, PSA 算法的性能最接近 FS 算法, 是一种简单实用的运动估计算法。

参 考 文 献

- [1] ISO/IEC 13818-2 Coding of Moving Picture and Associated Audio. 1995.
- [2] Koga T, Iinuma K, et al. Motion-compensated inter-frame coding for video conferencing. in Proc. NTC81, New Orleans, LA: 1981, C9.6.1-C9.6.5.

- [3] Kappagantula S, Rao K R. Motion compensated inter-frame image prediction. *IEEE Trans. on Comm.*, 1985, COM-33(9): 1011-1015.
- [4] Jain J R, Jain A K. Displacement measurement and its application in inter-frame image coding. *IEEE Trans. on Comm.*, 1981, COM-29(10): 1799-1808.
- [5] Ghanbari M. The cross-search algorithm for motion estimation. *IEEE Trans. on Comm.*, 1990, COM-38(7): 950-953.
- [6] Lee L W, Wang J F, Lee J Y, Shie J D. Dynamic search-window adjustment and interlaced search for block-matching algorithm. *IEEE Trans. on CASVT*, 1993, CAS VT-3(2): 85-87.
- [7] Li R, Zeng B, Liou M. A new three-step search algorithm for block motion estimation. *IEEE Trans. on CAS VT*, 1994, CAS VT-4(4): 438-442.
- [8] Po L M, Ma W C. A novel four-step algorithm for fast block motion estimation. *IEEE Trans. on CAS VT*, 1996, CAS VT-6(3), 88-91.
- [9] 范宏寅, 张春田. MPEG2 视频编码算法的几点改进. *通信学报*, 1996, 17(3), 64-69.
- [10] ISO/IEC/JTC1/SC29/WG11, "Test Model 5," Draft, Apr. 1993.

A NEW PREDICTIVE SEARCH ALGORITHM USED FOR MOTION ESTIMATION IN MPEG2

Liu Yang Luo Lijun* Zou Cairong* He Zhenya*

(Dept. of Computer and Information Science, Hefei University of Technology, Hefei 230009)

**(Dept. of Radio Eng., Southeast Univ., Nanjing 210096)*

Abstract A new Predictive Search Algorithm(PSA) is proposed in this paper, where the search displacements are between -16 to $+15.5$. It firstly derives motion vector from the linear weighting of motion vectors of three adjacent blocks, then searches from the prediction point using 3×3 search window and searches with the step of one until it reaches the boundary of search area or the local minimum point lies in the center of search window. Based on the spatial characteristics of center-biased distance between motion vector and corresponding prediction vector and temporal correlation of motion vectors, this method can significantly reduces the search times. Simulation results show that the proposed PSA is efficient to reduce the search range, search times, and computational complexity. Comparisons of PSA with other fast search algorithms are also given in detail.

Key words Motion estimation, MPEG2, Search algorithm

陆 阳: 男, 硕士生, 主要从事计算机控制、多维信号处理和视频信号处理方面的研究.

骆立俊: 男, 博士, 主要从事多维信号处理、视频信号处理和神经网络方面的研究.

邹采荣: 男, 教授, 主要从事多维信号处理、视频信号处理和神经网络方面的研究.

何振亚: 男, 教授, 博士生导师, 主要从事多维信号处理、视频信号处理和神经网络方面的研究.