

# 基于 DSP 实现的自适应搜索窗口运动估计算法

陈进成, 赵永利, 张 杰, 顾晓仪

(北京邮电大学光通信与光波技术教育部重点实验室, 北京 100876)

**摘 要:** 针对 DSP 芯片特点及运动估计精度和效率的要求, 提出一种新的自适应搜索窗口快速运动估计搜索算法。该算法充分利用视频图像的时间、空间上的相关性以及运动向量的统计特性, 在帧层和块层自适应调整搜索窗口, 利用钻石搜索法进行搜索。经实验验证, 该算法在保证运动估计精度的情况下, 数据传输量比固定搜索窗口算法平均少 50%, 搜索点数比钻石法平均少 3 个, 对不同类型视频序列都有较好的自适应能力。

**关键词:** 运动估计; 自适应搜索窗口; 峰值信噪比; 搜索点数; 块匹配

## Adaptive Search-window Motion Estimation Algorithm Based on DSP Implementation

CHEN Jin-cheng, ZHAO Yong-li, ZHANG Jie, GU Wan-yi

(Key Laboratory of Optical Communication and Lightwave Technologies, Ministry of Education,  
Beijing University of Posts and Telecommunications, Beijing 100876)

**【Abstract】** Based on DSP characteristics and motion estimation's accuracy and efficiency requirements, a new adaptive search-window fast motion estimation search algorithm is proposed. This algorithm uses the spatial and temporal correlations among video images and the statistical characteristics of motion vectors, adaptively adjusts the search-window on frame layer and block layer, and searches the window by using diamond search (DS) method. Experimental results show that, on average, it saves 50% data transmission compared with fixed-size search-window algorithm and 3 points compared with DS, maintaining high accuracy of estimation. And it suits for different types of video sequences adaptively.

**【Key words】** motion estimation; adaptive search-window; PSNR; search points; block match

随着网络的普及和视频应用需求的增加, 数字视频压缩技术得到了广泛关注。DSP 相对于通用处理器的优势使其非常适合作为视频的终端处理器。因此, 基于 DSP 的结构和特性, 研究视频压缩算法具有实际意义和实用价值。

### 1 背景介绍

在视频压缩编码中, 运动估计的运算量是影响编码速度的关键因素, 其通常占一个编码器总计算量的 60%~80%。块匹配法是目前应用最为广泛的运动估计方法, 其原理是先将当前帧分割成相互不重叠的矩形块(如大小为  $16 \times 16$  像素), 然后在参考帧中寻找最佳匹配位置作为预测, 最后仅对运动向量(motion vector)与块残差进行编码传输。由于大多数视频图像序列变化缓慢, 因此仅传输运动向量与相对较小的块残差能够很好地消除图像序列帧间冗余度, 从而降低码率。

在块匹配运动估计中, 运动估计的精度会对码率产生很大的影响。而全搜索以其极高的精度受到广泛的关注, 但由于其要求对搜索窗内所有点进行匹配, 运算量巨大, 在非专用芯片上很难适用。为了追求在较高运动估计精度下的较低运算量, 提出了各种快速算法。目前常用的快速算法有三步搜索法(TSS)、新三步搜索法(NTSS)<sup>[1]</sup>、四步搜索法(FSS)<sup>[2]</sup>和钻石搜索法(DS)<sup>[3]</sup>等。在这些算法中, NTSS在TSS的第 1 步增加 8 个邻居搜索点, 并引入了中途终止技术, 很好地消除了在小运动情况下TSS的低效; FSS使用了更小的搜索矩形模式, 能很好地适应实际视频序列运动向量的中心偏向(center-biased)特性; DS使用钻石型搜索模式, 与NTSS和FSS

相比, 减少了搜索点数。但这些算法都是基于“现实世界大部分运动向量都很小”和“残差空间的单调特性”模型, 使它们易于陷入局部极优, 且在大运动情况下性能明显下降。

同时大多数视频 DSP 芯片(如 TI 的 DM642、Cradle 的 CT3400 等)片内最高速存储空间往往不能或不适宜存放完整的参考帧原始图像数据, 在进行运动估计时需要从低速内存中将搜索窗口数据传输到高速缓存中, 然后再进行大量的估计运算。而参考窗口的传输占用了较多时间, 减小搜索窗口的大小, 能显著减少运动估计占用的时间。经验数据表明, 90%以上搜索窗口未全部使用, 存在较大的冗余。

针对上述快速搜索存在的问题和 DSP 芯片上实现的特点, 本文提出一种新的自适应搜索窗口快速算法(Adaptive Search-Window Search, ASWS)。该算法利用参考帧运动向量的统计特性和当前帧周围块搜索结果, 能自适应地确定当前块的起始搜索点和搜索窗口大小。

### 2 搜索算法

#### 2.1 算法流程

基于 DSP 实现考虑, 算法流程尽量简单。在确定初始搜

**基金项目:** 国家“863”计划基金资助项目(2006AA01Z246); 教育部新世纪优秀人才支持计划基金资助项目(NCET-05-0112)

**作者简介:** 陈进成(1982—), 男, 硕士研究生, 主研方向: 流媒体及其在智能光网络中的传输; 赵永利, 硕士研究生; 张 杰、顾晓仪, 教授、博士生导师

**收稿日期:** 2007-03-15 **E-mail:** chen\_bupt@hotmail.com

索点和窗口大小后,使用单一的钻石搜索法搜索,算法流程如图1所示。

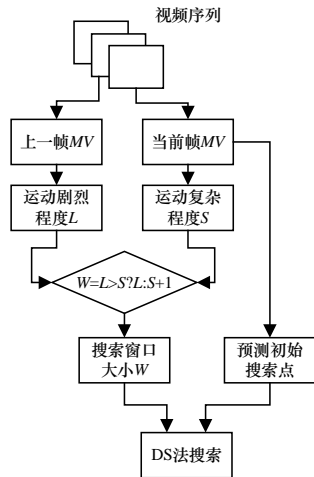


图1 ASWS 算法流程

## 2.2 初始搜索点

由于运动物体的整体性和视频运动的连续性,视频的运动必然具有时间和空间上的相关性<sup>[4]</sup>,这种相关性可用于对搜索起点进行预测。与其他点相比,预测点更加靠近最优匹配点,使得搜索点数减少,又能避免搜索点陷入局部最优。实验统计<sup>[5]</sup>表明,当前块的运动矢量与其左边、上边、右上块的运动矢量相关性较大(如图2所示),而与其他相邻块运动矢量的相关性较小,因此,考虑使用这3个块的运动矢量来预测当前块的初始搜索点。具体方法是:以当前块相邻3个块运动向量的中间值作为本块的运动向量估计值。如果当前块处在左边界就把该块左侧的运动向量的值记为0;如果当前块处在上边界就把上侧和右上侧块的运动向量值记为左侧块运动向量的值;如果当前处在右边界,则将右上侧块的运动向量置0。

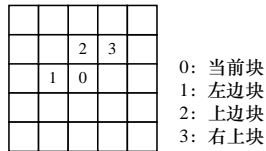


图2 预测相关块

## 2.3 自适应搜索窗口

窗口的大小根据前一帧和当前帧的周围块共同确定。一帧的MV统计值能反映运动的宏观剧烈程度,而前一帧的运动特性与当前帧有很大的相关性,因此,可以用来预测当前帧的运动剧烈程度。同时,当前帧周围块的MV能反映当前帧的局部特性,可以预测当前帧的运动细节。结合两者可以很好地预测当前帧的运动。

研究表明,运动向量的二阶矩能很好地评价当前帧相对参考帧的运动程度<sup>[6]</sup>。计算公式如下:

$$L_x = \sqrt{\frac{\sum_{i=1}^n MV_{x_i}^2}{n}}, \quad L_y = \sqrt{\frac{\sum_{i=1}^n MV_{y_i}^2}{n}} \quad (1)$$

$$L = \max(L_x, L_y) \quad (2)$$

其中,  $n$  等于一帧图像的块数;  $MV_x$ ,  $MV_y$  分别表示前一帧MV的x方向和y方向分量;  $L_x$ ,  $L_y$  分别表示x方向和y方向的运动剧烈程度;  $L$  表示整帧的运动剧烈程度。

当前帧的周围块仍选择相关性最大的左边、上边和右上

块。计算公式如下:

$$S_x = \max |MV_{x_i} - MV_{x_{middle}}|, \quad S_y = \max |MV_{y_i} - MV_{y_{middle}}| \quad (3)$$

$$S = \max(S_x, S_y) \quad (4)$$

其中,  $i=1,2,3$ ;  $|\cdot|$  表示取绝对值;  $MV_{x_i}$ ,  $MV_{y_i}$  分别表示当前帧周围3个块的MV的x方向和y方向分量;  $MV_{x_{middle}}$ ,  $MV_{y_{middle}}$  分别表示这3个MVx, MVy的中间值;  $S_x$ ,  $S_y$  分别表示块x方向和y方向的运动复杂程度;  $S$  表示块的运动复杂程度。

搜索窗口大小的确定方法如下:

$$\begin{cases} W = L & \text{if } (S < L) \\ W = S + 1 & \text{else} \end{cases} \quad (5)$$

其中,  $W$  表示窗口半径。编码的第一帧没有  $L$  值,将  $L$  设为  $MV$  最大值,以利于对视频序列做出更准确的估计。通过对全搜索的MV结果的统计,验证了自适应搜索窗口的准确性,如表1所示,其中,准确率是全搜索最佳匹配块在自适应搜索窗口中的概率。

表1 自适应搜索窗口的准确率 (%)

mobile	football	Stefan	miss_am	Suzie	foreman
100.00	99.93	99.99	99.97	99.99	99.97

## 2.4 搜索算法

基于钻石搜索法对大运动和小运动都有较好的适应性且实现简单,本文选用该单一算法对取出的自适应搜索窗口进行搜索。实验数据显示,自适应搜索窗口与钻石法的结合产生了良好的自适应性和鲁棒性。

## 3 实验结果及分析

### 3.1 实验条件

为排除其他因素影响,本文的实验只在CT3400 DSP芯片上实现运动估计。CT3400是美国Cradle公司推出的一款多核数字信号处理芯片,其计算部分内含4个通用处理核,8个数据处理核和1个数据传输核,通用处理核适用于流程控制,后两者适用于数据处理和传输。每个处理核拥有各自独立的PC(Program Counter)、PSW(Program Status Word)和高速双端口寄存器,片内还有各处理核共享的存储区和缓存区。该多核结构不单纯依赖提高时钟频率来提高芯片运算性能,随处理核数量的增长,其性能线性提高,适合实时多媒体应用。

测试的序列如表2所示,仅对亮度进行整像素运动估计,块大小为 $16 \times 16$ 像素,前一帧原始数据作为后一帧参考帧,第1帧不进行运动估计,并以SAD(Sum of Absolute Difference)为匹配衡量准则。除本文提出的ASWS算法,同时对FS, TSS, NTSS, DS进行了实现,搜索范围均为 $\pm 7$ 点。

表2 测试序列

	mobile	football	Stefan	miss_am	Suzie	foreman
格式	CIF	CIF	CIF	QCIF	QCIF	QCIF
帧数	300	90	90	150	150	400

### 3.2 实验结果与讨论

图3给出了在相同条件下的各自亮度的峰值信噪比(PSNR)比较,图4给出了各算法平均每块的搜索点数的比较,所有算法搜索点数不重复计数,也就是每块已搜索过的点不再搜索。表3给出了ASWS传输数据量,即自适应搜索窗口的数据量与固定搜索窗口的比率。

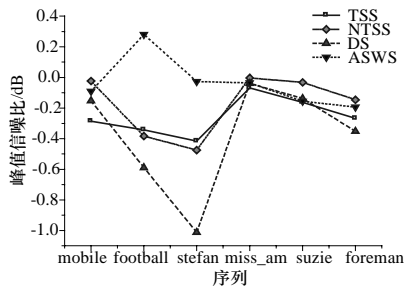


图3 各算法峰值信噪比比较

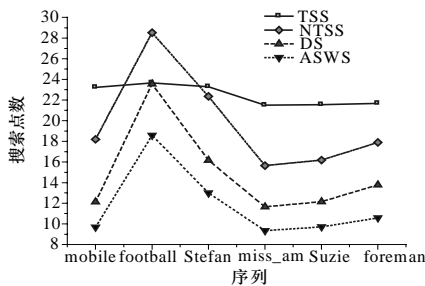


图4 各算法搜索点数比较

表3 ASWS 传输数据量 (%)

mobile	football	Stefan	miss_am	Suzie	foreman
39.61	96.88	65.97	40.85	42.25	48.45

mobile 是镜头平移和小运动序列, football, Stefan 是体育转播序列, 前者运动较剧烈, 后者相对缓慢, miss\_am, Suzie, foreman 都是头肩序列, miss\_am 是标准可视电话视频, 场景固定不变, Suzie 中人物头像所占比例较大, foreman 镜头有抖动。从实验结果可以看出, 在快速搜索算法中, 对于典型头肩序列, NTSS 有最好的 PSNR 性能, 但其搜索点数仅低于 TSS, 比 ASWS 平均每块多搜索 6 个点, 同时 ASWS 的 PSNR 仅增加 0.03 dB~0.12 dB。对于大运动复杂序列, ASWS 在 PSNR 和搜索点均有最佳性能, 其中, football 序列的 PSNR

性能甚至优于 FS, 原因是 ASWS 初始搜索点不是(0,0), 其 MV 可能超出 $\pm 7$  范围; 同时 ASWS 平均块搜索点数比 DS 平均少 4 个点。对于镜头平移序列, ASWS 能有效预测运动方向, 搜索窗口小于固定窗口的 40%, 并保持了 PSNR 和搜索点数的性能。总的来说, ASWS 平均节约了 50%左右的数据传输量, 从另一方面提高了运动估计的速度。

#### 4 结束语

本文提出了基于 DSP 实现的自适应窗口搜索算法, 对不同序列都有较好的自适应能力, 在保证 PSNR 性能接近 FS 的同时, 显著减少了平均块搜索点数, 并减少了搜索窗口数据传输量。算法简单易实现, 虽然  $L_x, L_y$  涉及浮点计算, 但结果仅取其整数部分, 且大小限制在 0~7, 容易通过查表实现。该算法对基于 DSP 实现的编码器有一定的实用价值。

#### 参考文献

- [1] Li Renxia, Zeng Bing. A New Three-step Search Algorithm for Block Motion Estimation[J]. IEEE Trans. on Circuits System Video Technology, 1994, 4(4): 438-443.
- [2] Lai-man P. A Novel Four-step Search Algorithm for Fast Block Motion Estimation[J]. IEEE Trans. on Circuits System Video Technology, 1996, 6(3): 313-317.
- [3] Zhu Shan, Ma Kaikuang. A New Diamond Search Algorithm for Fast Block Matching Motion Estimation[J]. IEEE Trans. on Image Processing, 2000, 9(2): 287-290.
- [4] 杨成军, 魏春城, 陈 铭, 等. 视频编码中基于时空信息运动估计的研究[J]. 高技术通信, 2002, 4(4): 27-31.
- [5] Hsieh C H, Lu P C, Shyn J S, et al. Motion Estimation Algorithm Using Interblock Correlation[J]. Electronics Letters, 1990, 26(5): 276-277.
- [6] 李 翔, 吴国威. 一种适用于 H.264 的基于自适应搜索范围的快速运动估计算法[J]. 中国图象图形学报, 2004, 9(4): 471-476.

(上接第 234 页)

这种点与点的关系来构造空间类别分类器, 提出了一种具体生成神经网络的算法, 并且把算法与不使用局部特性覆盖算法进行了分析、比较。通过实验可以看出: (1)SNCA 利用空间点局部特性, 可以利用这个信息构造局部同向、整体异向的空间分类器; (2)SNCA 构造的局部分类器, 相比于不使用局部区域的构造算法 SCA 而言, 测试样本正确率相对稳定, 而且拒识率相对较高, 这种类型的神经网络有利于安全性较高的应用类型。结合待识别事物特征的高维空间特征分布, 有利于提高网络构造性能。

从高维空间覆盖角度来分析和构造网络, 可以克服传统网络, 例如 BP 网络, 存在网络难于训练、中间层神经元难

于确定、知识表示缺陷等问题。如何更好地利用样本点信息构造神经网络, 还有待研究。

#### 参考文献

- [1] McCulloch W S. A Logical Calculus of the Ideas Immanent in Nervous Activity[J]. Journal of Symbolic Logic, 1944, 9(2): 49-50.
- [2] Rujan P, Marchand M. A Geometric Approach to Learning in Neural Networks[C]//Proceedings of the International Joint Conference on Neural Networks'89. Washington, D. C.: [s. n.], 1989: 105-110.
- [3] Ramacher U, Wesseling M. A Geometrical Approach to Neural Network Design[C]//Proceedings of the International Joint Conference on Neural Networks'89. Washington, D. C.: [s. n.], 1989: 147-154.