

自适应块匹配搜索算法研究

张培珍¹, 江华俊², 沈玉利¹

(1. 广东海洋大学 信息学院, 广东 湛江, 524088; 2. 中兴通信股份有限公司, 广东 深圳 518000)
(zpzhen@sohu.com)

摘要:为从视频序列提取运动矢量, 进行有效的视频检索, 对现存的运动特征的提取算法存在的问题进行了分析, 提出了基于自适应块匹配搜索的全局运动特征提取方法。实验表明该算法在大大提高搜索速度的同时, 能够获得几乎与全搜索算法相同的峰值信噪比(PSNR)。

关键词:视频检索; 全局运动; 自适应块匹配算法; 峰值信噪比

中图分类号: TP391.4 **文献标识码:** A

Research on adaptive block matching algorithm

ZHANG Pei-zhen¹, JIANG Hua-jun², SHEN Yu-li¹

(1. Information College, Guangdong Ocean University, Zhanjiang Guangdong 524088, China;
2. Zhongxing Telecommunications Equipment Corporation, Shenzhen Guangdong 518000, China)

Abstract: In order to extract motion vector to realize video retrieval efficiently, the existing problems and characters of video information were analyzed. Then a global motion feature extraction scheme based on adaptive block matching fast searching algorithm was put forward. Experiments show that the algorithm can obtain almost the same PSNR as the full search algorithm at very low computation cost.

Key words: video retrieval; global motion; adaptive block matching algorithm; PSNR

0 引言

基于内容的视频数据检索^[1]为人们检索具有相似内容的视频数据提供了新的手段。在视频序列的运动分析中, 常把运动信息分为全局运动信息和局部运动信息^[2]。由于物体各点的运动是由全局和局部运动叠加而成的, 为了正确地提取局部运动信息, 必须首先估计出全局运动信息。基于块匹配的相关性技术是最直观且被广泛应用的方法, 其基本思想是: 每个块的灰度模式在连续的帧中几乎保持不变, 且局部的纹理包含了足以互相区分的信息, 这样可以通过在一定大小的窗口中搜索出唯一匹配的灰度块来得到图像序列的运动矢量。块匹配算法的最大的不足是计算的复杂性, 计算一个帧的运动矢量场需要进行 N^2 次匹配搜索, 如果搜索限定在距离原块为 M 的范围之内, 总的时间复杂度为 $N^2 \cdot (2M+1)^2$, 这样大的计算量即使在高性能的计算机上也要相当长的时间。目前研究最多的是快速搜索算法, 如二维对数法、三步法、新三步法、四步法及菱形法等。各种块匹配搜索算法虽然加快了匹配速度, 但往往容易陷入局部最优点。综合近年来出现的各种搜索算法如 MVFAST^[3]、PMVFAST^[4]的优点, 提出了一种自适应块匹配搜索算法, 它包括搜索起点的可靠预测、搜索模式的选择以及搜索终止准则的自适应调整。该算法不仅提高了计算速度, 而且在搜索质量上可以接近于全搜索算法。

1 块匹配的基本原理^[5]

首先, 把图像划分为许多互不重叠的子块, 并认为子块内所有像素的位移都相同。假设在图像序列中, t 时刻对应于第 k 帧图像, $t - \Delta t$ 对应第 $k-1$ 帧图像。对于 k 帧的一个子块, 在

$k-1$ 帧中寻找与其匹配的子块, 即该匹配块从第 $k-1$ 帧中所处的位置运动到了第 k 帧中所处的位置, 这种位置的变化即是运动矢量。将图像分割成 $M \times N$ 的小块, 假设块内像素作相同的运动, 且只作平移运动。设子块的最大位移矢量为 (u_{\max}, v_{\max}) , 则搜索范围为 $(M + u_{\max} \times 2) \times (N + v_{\max} \times 2)$ 。图1显示了待匹配块与搜索区域的关系。

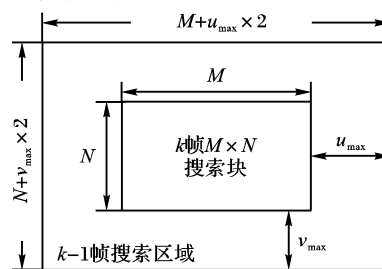


图1 待匹配块与搜索区域的关系

为了算法的方便实现, 子块的 M, N 一般取值相等, u_{\max}, v_{\max} 也取值相等。

2 自适应块匹配的搜索算法

该算法的主要思想是: 首先通过搜索起点的预测, 使当前块的初始运动矢量有可能接近其最终运动矢量, 然后根据图像局部特征简单有效地对视频进行分类并选择合适的搜索模式, 使其能根据视频运动的类型进行自适应的搜索, 最后采用搜索终止准则保证搜索结果在这个预测的起点附近结束时具有足够的精度, 从而实现快速、均匀、精度高的运动矢量搜索。

2.1 初始搜索点的预测

首先利用相邻块之间的运动相关性, 选择一个反映当前块运动趋势的预测点作为初始搜索点, 以提高搜索速度和预

测的准确性。这个预测点比(0,0)点更靠近全局最小点,使得基于中心倾向的块匹配算法更有效。然后用中值法^[6]来预测当前块的运动矢量,如图2所示, MV 为当前块运动矢量, MV_1 、 MV_2 、 MV_3 分别是其左方、上方与右上方的块的运动矢量。

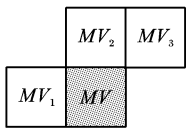


图2 相邻块运动矢量

中值法得到的预测点比(0,0)点与当前块运动矢量确实有更强的相关性,相邻帧运动矢量不仅在空间上具有相关性,在时间上也具有相关性。

2.2 搜索模板的选择

DS算法^[7]是目前快速块匹配算法中性能最优的算法之一。DS算法采用了两种搜索模板,分别是有9个检测点的大模板LDSP(Large Diamond Search Pattern)和有5个检测点的小模板SDSP(Small Diamond Search Pattern),如图3所示。

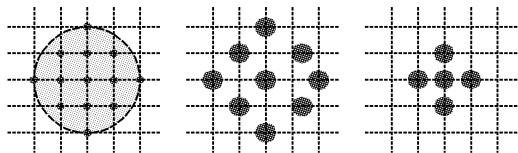


图3 DS的搜索模板

本文采用大小两种菱形搜索模板,步长为2的LDSP和步长为1的SDSP,其中LDSP用于在大范围内进行快速定位,SDSP用于在小范围内的“微调”。在最初的大小模板的选择上,是根据当前块可能的运动等级来选择的。当前块运动等级可以通过预测点的运动矢量来估计:当预测点的运动矢量幅度大于阈值 L 时,认为当前块的运动等级为高,为大运动,此时将采用LDSP,由于LDSP步长大,搜索范围广,可以进行粗定位,使搜索过程不会陷于局部最小;当预测点的运动矢量幅度小于阈值 L 时,认为当前块的运动等级为低,即小运动,此时将采用SDSP,使搜索不至于有大的起伏,所以它的性能优于其他算法。

2.3 搜索终止准则

2.3.1 块匹配的匹配准则^[8]

运动估值算法中常用的匹配准则有三种,即最小绝对差(Minimum Absolute Difference, MAD)、最小均方误差(Mean Square Error, MSE)和求和绝对误差(Sum of Absolute Difference, SAD),分别定义如下:

(1) 最小绝对差

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

(2) 最小均方误差

$$MSE(i, j) = \frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N [f_k(m, n) - f_{k-1}(m + i, n + j)]^2$$

(3) 求和绝对误差

$$SAD(i, j) = \sum_{m=1}^M \sum_{n=1}^N |f_k(m, n) - f_{k-1}(m + i, n + j)|$$

式中, (i, j) 为位移矢量, f_k 和 f_{k-1} 分别是当前帧和上一帧的像素灰度值, $M \times N$ 为宏块大小, $MAD(i, j)$ 、 $MSE(i, j)$ 或 $SAD(i, j)$ 最小的点即为要找的最优匹配点。

2.3.2 终止准则

快速算法与全搜索的本质差别在于,搜索过程能在保证搜索精度的前提下及时终止。为提高算法的速度,一些算法利用了SAD阈值作为搜索终止的条件,即当SAD小于某一阈

值时终止搜索过程。本文通过考察时间和空间相邻块的SAD的分布发现,虽然空间相邻块的SAD的分布在数值上没有很明确的相关性,但当前块的SAD与前一帧同一位置块的SAD间存在着一种数值上的相关关系。经过对多种视频序列的统计分析(如表1和表2)可以看出,相邻帧同一位置块的SAD间普遍存在着某种比例关系,其平均比值在0.9~1.1之间。

表1 序列Matrix的第20帧与第21帧的SAD

第20帧的SAD								第21帧的SAD							
897	862	794	...	1055	414	455	536	902	881	893	...	911	463	464	509
1219	1277	1380	...	983	652	581	547	1176	1204	1281	...	894	671	662	538
1174	1539	2102	...	1091	978	887	903	1204	1469	1994	...	981	989	926	971
...
2655	2712	2379	...	2132	2044	1954	1938	2613	2701	2412	...	2057	2019	1989	1910
843	1290	1106	...	1273	1158	965	833	914	1367	1050	...	1243	1266	992	879

为此,本文算法采用了如下方法确定SAD阈值:以参考帧同一位置块的SAD乘以某个系数(系统选取0.8~1.2较为理想)作为SAD阈值。这种方法不受视频运动类型的影响,具有较强的自适应性,而且能准确确定恰当的阈值使搜索过程适时终止,有助于以较小的搜索代价得到较好的搜索结果。为加快搜索过程本文算法采用了二级终止准则:若 $SAD < T1$ 表明搜索结果已经足够精确,可立即终止搜索过程;若 $T1 \leq SAD < T2$ 表明搜索结果尚可,只需在小范围进行微调,于是将搜索模板从LDSP转为SDSP。

3 实验结果

按照MPEG-4对运动估计算法评价的有关要求^[9],将本文的自适应块匹配快速搜索算法与其他算法在不同码率和搜索范围下对几个典型序列的搜索结果进行比较,如表2所示,包括平均PSNR(峰值信噪比)、搜索点数(块匹配的点数)以及搜索速度相对于FS提高的倍数。其中,Mother & Daughter(QCIF格式)为空间细节不丰富、运动变化小的A类视频;Coastguard(QCIF格式)为空间细节比较丰富、运动变化小或空间细节不丰富、运动变化中等的B类视频;Tabletennis(SIF格式)为空间细节很丰富、运动变化中等或空间细节比较丰富、运动变化大的C类视频;Basket(CCIR格式)为不仅空间细节很丰富而且运动变化大的D类视频^[10]。

表2 自适应块匹配算法实验结果

序列	码率	搜索范围	算法	PSNR	搜索点数	速度倍数
Mother & Daughter	24	16	FS	34.80	10036224	1.00
			MV	34.76	48611	206.4
			PMV	34.78	32251	311.1
			本文	34.78	30163	332.7
Coastguard	48	16	FS	28.91	10036224	1.00
			MV	28.83	63413	100.9
			PMV	28.90	50016	158.2
			本文	28.90	30163	200.6
Tabletennis	1024	48	FS	34.97	760543840	1.00
			MV	34.92	521379	1458.7
			PMV	34.95	308241	2467.3
			本文	34.97	303674	2504.4
Basket	4000	64	FS	26.71	5580207144	1.00
			MV	26.40	4208477	1325.9
			PMV	26.49	2917152	1912.9
			本文	26.55	2633648	2118.8

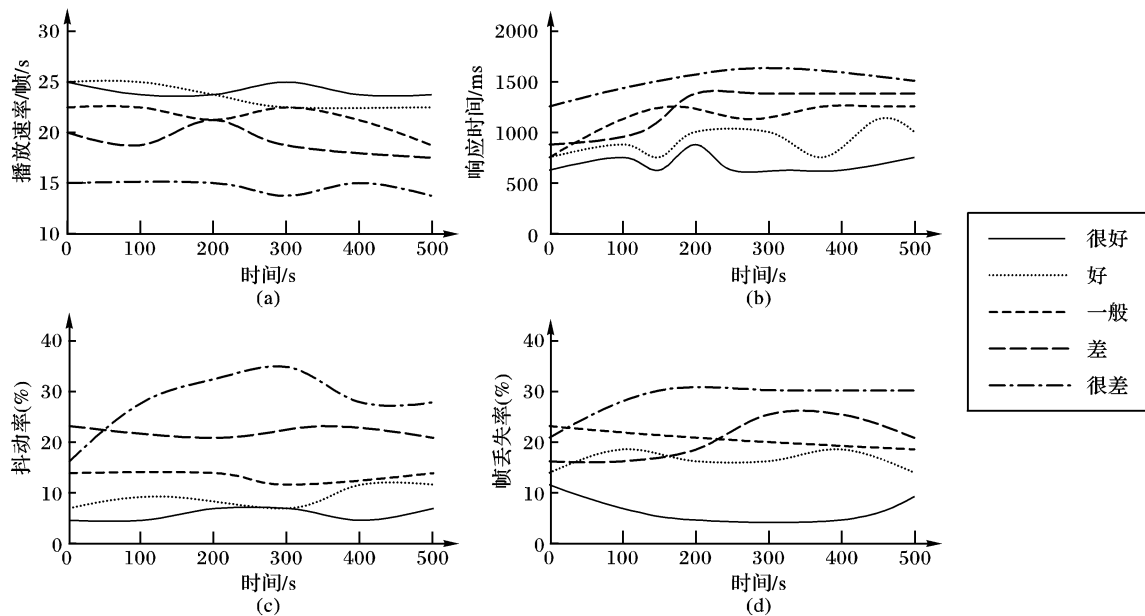


图2 实际测量的播放效果曲线

5 结语

本文基于 XVID MPEG 编码器软件包和 RTP 协议提出了 MPEG-4 视频流的服务质量 (QoS) 管理的两层分层管理方法。基于 RTP 协议提出了 QoS 保证的逻辑结构。提出基于 XVID-MPEG-4 的应用层 QoS 参数应分为媒体编码质量(类、级)、播放响应时间、抖动时间和帧丢失率等,即 $APP_QoS([C_a, C_d], [L_a, L_d], [tr_a, tr_d], [ADF_a, ADF_d], [ALF_a, ALF_d])$; 系统层 QoS 参数应分为位速率、端到端延时、延时抖动和位丢失率等,即 $SYS_QoS([r_a, r_d], [td_a, td_d], [tj_a, tj_d], [p_a, p_d])$ 。提出了应用层 QoS 参数到系统层 QoS 参数的映射公式。模拟实验结果表明本文 QoS 保证分层方法和 QoS 参数映射算法能够保证 MPEG-4 视频流的服务质量。

参考文献:

- [1] Audio-Video Transport Working Group. RTP: A Transport Protocol for Real-Time Applications [EB/OL]. <http://www.faqs.org/rfcs/rfc1889.html>, 1996-11.
- [2] 杨学良, 张占军. 分布式多媒体计算机系统教程 [M]. 北京: 电

子工业出版社, 2002.

- [3] 陶蒙华, 黄孝建, 沈树群. MPEG 视频流的两种基于 RTP 的建模及其比较 [J]. 北京邮电大学学报, 2002, 25(1): 53.
- [4] LI WP. Overview of Fine Granularity Scalability in MPEG-4 Video Standard [J]. IEEE Transactions Circuits and Systems for Video Technology, 2001, 11(3): 894.
- [5] GONZALEZ RC, RICHARD E. Digital Image Processing [M]. Second Edition. 北京: 电子工业出版社, 2003.
- [6] RICHARDSON IEG. H. 264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia [M]. 北京: 电子工业出版社, 2003.
- [7] SCHAFER R, WIEGAND T. The Emerging H. 264/AVC Standard [J]. EBU Technical Review, 2003, 38(1): 25.
- [8] KIKUCHI Y, NOMURA T, FUKUNAGE S. RTP Payload Format for MPEG-4 Audio/Visual Streams [EB/OL]. <http://rfc3016.x42.com>, 2000-11.
- [9] 汤庸, 杨学良, 区海翔, 等. QoS 映射与自适应协商机制的研究与实现 [J]. 软件学报, 2001, 12(05): 751.
- [10] 张占军, 韩承德, 杨学良. 基于自适应节的多媒体流 QoS 保证的研究 [J]. 计算机学报, 2000, 23(12): 1320.

(上接第 798 页)

实验结果表明, 与全搜索相比, 本文算法在大幅度提高搜索速度 (200 ~ 9000 倍按搜索点数计算) 的情况下得到了与其相当的效果。与快速运动估计算法 MVFAST 和 PMVFAST 相比, 本文算法的平均速度分别提高了 45% 和 8%, 而且平均 PSNR 更高。

参考文献:

- [1] 章毓晋. 基于内容的视觉信息检索 [M]. 北京: 科学出版社, 2003.
- [2] JEANNIN S, JASINSCHI R, SHE A, et al. Motion description for content-based video representation [J]. Signal Processing: Image communication, 2000, 16(1-2): 59-85.
- [3] HOSUR PI, MA KK. Motion Vector Field Adaptive Fast Motion Estimation [A]. Second International Conference on Information, Communications and Signal Processing (ICICS '99) [C]. Singapore, 1999.
- [4] TOURAPIS AM, AU OC, LION ML, et al. Fast block matching motion estimation using predictive motion vector field adaptive search technique (PMVFAST) [A]. In: ISO/IEC JTCL/SC29/WG11

MPEG2000/m 5866 [C]. Noordwijkerhout, Netherlands, 2000.

- [5] CHEN OT-C. Motion Estimation Using a One-Dimensional Gradient Descent Search [J]. IEEE Transactions Circuits and Systems for Video Technology, 2000, 10(4): 608-616.
- [6] PEKER KA, DIVAKARAN A, PAPATHOMAS TV. Automatic measurement of intensity of motion activity of video segments [A]. Proceedings of SPIE Conference on Storage and Retrieval from Multimedia Databases [C]. San Jose, CA, 2001. 341-351.
- [7] ZHU S, MA KK. A new diamond search algorithm for fast block-matching motion estimation [J]. IEEE Transactions on Image Processing, 2000, 9(2): 287-290.
- [8] TOURAPIS AM. Core experiment on block based motion estimation [Z]. ISO/IEC JTCL/SC29/WG11 MPEG2000/M5867, Noordwijkerhout, 2000.
- [9] Experimental conditions for evaluating encoder motion estimation algorithms [A]. In: ISO/IEC JTCL/SC29/WG11 N3141 [C]. Hawaii, USA, 1999.
- [10] MPEG-4 testing and evaluation procedures document [A]. In: ISO/IEC JTCL/SC29/WG11 N0999 [C]. Tokyo, July, 1995.