

文章编号:1001-9081(2010)06-1626-03

基于 H. 264 的多参考帧运动估计快速算法

严浩, 何小海, 卿粼波, 吕瑞, 曾强宇

(四川大学 电子信息学院, 成都 610064)

(goddn@126.com)

摘要: H. 264/AVC 中引入多参考帧运动补偿来提高视频编码性能, 由此产生的多参考帧运动估计 (MRF-ME) 却带来了巨大的运算代价。为提高编码速度, 降低计算复杂度, 提出一种基于空间域相关性的运动估计算法——缩小的菱形算法 (DDS)。先运用前向主矢量选择法不断修正预测运动矢量, 再根据最佳参考帧位置的统计特性对不同参考帧使用不同模板进行搜索。实验结果表明, 与 H. 264 参考模型 JM10.2 相比, 该算法保持了较好的图像质量且码率变化很小, 运动搜索点数平均减少接近 80%, 并能有效地降低编码器复杂度。

关键词: 多参考帧; 运动估计; 空间域相关性; 菱形搜索

中图分类号: TN919.81 **文献标志码:** A

Fast multi-frame motion estimation algorithm for H. 264

YAN Hao, HE Xiao-hai, QING Lin-bo, Lü Rui, ZENG Qiang-yu

(College of Electronics and Information Engineering, Sichuan University, Chengdu Sichuan 610064, China)

Abstract: The multiple reference frame motion compensation was introduced into H. 264/AVC to improve video coding performance. However, the resulted Multiple Reference Frame Motion Estimation (MRF-ME) leads to much more computational cost. To improve the encoding speed and reduce the computational complexity, a new algorithm called DDS based on spatial correlation was proposed. First, the algorithm updated predictive motion vector constantly using Forward Dominant Vector Selection (FDVS), then searched different reference frame by different template according to the statistical characteristic of the location of optimal reference frame. The experimental results show that the algorithm can decrease the search points by 80% and reduce the complexity of the encoder considerably while maintaining nearly unchanged Peak Signal-to-Noise Ratio (PSNR) of pictures and the bit rate compared with the H. 264 reference software JM10.2.

Key words: multiple reference frame; motion estimation; spatial correlation; diamond search

0 引言

新一代视频压缩编码标准 H. 264/AVC^[1] 提出了多参考帧运动估计, 在多参考帧模式下, 对一个宏块或子块进行运动补偿时, 编码会从过去的若干个已编码帧中选定一帧作为参考帧, 寻求当前编码宏块或子块的最佳匹配块, 以获得更好的预测效果。多参考帧运动估计不仅使预测的精度得以提高, 压缩率上升, 而且也增强了抗误码干扰的能力, 增强了率失真性能。

目前 H. 264/AVC 参考模型 JM10.2^[2] 提供了 4 种算法来实现多参考帧的预测。一种方法是全搜索, 虽然它的效果最好, 但搜索所花费的时间却非常惊人, 在实际使用中并不被采用。另外 3 种方法都是关于运动搜索的快速算法 (Fast Motion Estimation, FME), 虽然在单参考帧的情况下能大幅度提高运动搜索的速度, 但在多参考帧模式下, JM10.2 中仍然是对每个参考帧使用相同的方法搜索, 计算量随着参考帧数目的增加而线性增加。FME 相对全搜索而言, 搜索速度有了很大的提高, 但计算量仍然很大。据统计, 即使采用 FME, 如果 7 种宏块模式全开并且采用 5 个参考帧的话, 运动估计的时间还要占整个编码时间的 60% 以上。因此, 对其进行进一步优化, 以降低计算复杂度具有重要意义。

文献[3]在分析运动矢量空间分布特性基础上, 将菱形算法 (Diamond Search, DS)^[4] 扩展到多参考帧情形上, 并运用前向主矢量选择法 (Forward Dominant Vector Selection, FDVS) 不断修正搜索模板中心点的位置, 提出了三维菱形法 (Three-Dimensional Diamond Search, TDDS)。该算法在保持高信噪比和低比特率的同时, 能有效地降低编码器复杂度, 但计算复杂度仍然很高。本文根据最佳参考帧位置的统计特性提出了一种新的基于空间域相关性的运动估计算法——逐渐缩小的菱形算法 (DDS), 在保持运动估计没有显著下降情况下减少更多的搜索点, 降低了计算量。

1 多参考帧分析

1.1 多参考帧运动矢量相关性

在视频序列图像中, 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。物体运动的连续性同样也导致多参考帧中运动矢量域的相关性^[5]。如图 1 所示。

沿着运动轨迹进行追踪, 就能得到相邻帧运动矢量之间的关联性^[6], 即:

$$MV_t^{-i} = MV_t^{-i-1} + MV_{t-i+1}^{-1} \quad (1)$$

其中 MV_t^{-i} 表示从第 i 帧指向 $t-i$ 帧的运动矢量。采用运动矢

收稿日期: 2009-12-03; 修回日期: 2010-01-25。

作者简介: 严浩 (1985-), 男, 四川达州人, 硕士研究生, 主要研究方向: 视频压缩编码、图像通信; 何小海 (1964-), 男, 四川成都人, 教授, 主要研究方向: 通信与信息处理、图像处理与通信、机器视觉、智能系统; 卿粼波 (1982-), 男, 四川简阳人, 讲师, 主要研究方向: 模式识别、图像处理、图像通信; 吕瑞 (1982-), 男, 四川绵阳人, 博士研究生, 主要研究方向: 视频压缩编码、图像通信; 曾强宇 (1982-), 男, 四川自贡人, 博士研究生, 主要研究方向: 视频压缩编码、图像通信。

量相关性进行多参考帧预测,需将每帧相对于前一帧的运动矢量都保存下来,每个宏块需要保存7种块类型的运动矢量,这对于多参考帧情形下的快速搜索很有帮助,可有效定位编码块在每一帧的预测起点,从而减少搜索点数,节约运动估计时间^[5]。

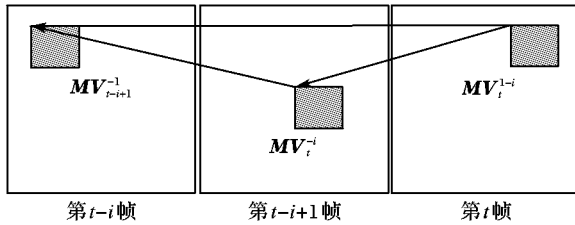


图1 运动矢量相关性

参考模型JM10.2在多参考帧模式下加快的只是搜索每一个参考帧的过程,每次仍需遍历所有的参考帧,没有利用参考帧之间的相关性缩小搜索的范围。由于视频图像采样频率的原因,大部分运动图像序列相邻帧之间的运动变化是很小的,有的几乎没有变化,随着帧间间隔越远,运动变化越大,因此,在运动搜索的时候,通常情况下距离当前帧越近的图像帧其相关性越高,作为参考帧预测的效果就越好。所以,在多参考帧搜索模式下,距离当前帧越近的图像帧成为最佳参考帧的几率越大。

1.2 最佳参考帧位置的统计分析

为了验证在多参考帧搜索模式下,距离当前帧越近的图像帧成为最佳参考帧的几率越大,本文在参考软件JM10.2上对不同的视频序列进行了实验统计。表1为参考帧个数为5、QP为28、使用全搜索算法编码100帧时各序列中参考帧被选中的概率。其中REF(0)表示与当前待编码帧最相邻的前一参考帧,REF(1)表示REF(0)的前一帧,以此类推。

表1 参考帧概率统计结果

格式	序列	REF(0)	REF(1)	REF(2)	REF(3)	REF(4)
QCIF	Coastguard	84.73	7.96	1.03	1.85	1.43
	Container	93.22	2.77	2.03	1.28	0.69
	Foreman	70.23	5.91	9.83	3.04	2.99
	M&D	90.46	4.42	3.35	0.76	1.00
	Mobile	51.31	13.79	14.59	9.93	10.37
	News	93.90	3.25	1.82	0.47	0.57
	Silent	92.09	3.71	2.48	0.74	0.97
	平均值	82.28	7.12	5.45	2.58	2.57
CIF	Coastguard	85.26	6.72	4.22	2.20	1.60
	Container	90.44	4.33	2.85	1.33	1.06
	Foreman	74.73	12.03	8.55	2.38	2.61
	M&D	89.82	4.93	3.58	0.72	0.94
	Mobile	47.26	16.76	17.02	9.85	9.11
	News	93.93	2.99	2.04	0.52	0.53
	Silent	91.86	3.81	2.63	0.75	0.94
	平均值	81.90	7.37	5.84	2.24	2.40
总平均值		82.09	7.24	5.64	2.56	2.49

注:以上平均值均只保留两位小数(四舍五入)。

从表1可看出,一般情况下,在所有参考帧中距离当前帧越近的参考帧被选择为最优参考帧的概率就越高,尤其是第一参考帧REF(0)。REF(0)成为最优参考帧的概率平均为82.09%,最高的序列达到了93.93%,而其余参考帧被选中的概率远远低于REF(0),平均概率仅为17.91%。因此,在多参考帧搜索模式下,最佳参考帧的位置在空间上也具有中心偏置的特性,即多参考帧下最佳参考帧的位置大都位于当前帧的附近参考帧上,如果对每个参考帧都逐一搜索,即使使用

FDVS算法快速定位预测起点,也会造成计算时间的浪费。所以,对于不同参考帧应该使用不同的搜索方式。

由于从参考帧REF(0)中获得的运动搜索结果对整个编码性能有着重要影响,在进行多参考帧选择时,对该参考帧应该进行大范围地搜索,得到最优的运动矢量^[5]。相比较而言,由于其余的参考帧被选中的概率很低,而且每增加一个参考帧就会增加很大一部分计算量,所以对于离当前帧较远的参考帧进行小范围的搜索,尽量减少对它们的搜索,降低视频编码计算复杂度。

2 逐渐缩小的菱形搜索算法

2.1 菱形搜索模板

本文算法根据运动矢量分布的中心偏移性和最佳参考帧位置的统计特性,设计了3个菱形搜索模板。菱形模板不仅能很好地适应运动矢量概率分布上的中心偏置特性,而且菱形模板搜索时各步骤之间有很强的相关性,模板移动时只需要在几个新的检测点处进行匹配计算,因此也提高了搜索速度,硬件实现上也比较容易。

具体模板如图2所示:模板1是在模板0的基础上减少最外围的4个搜索点形成的;模板2是在模板1的基础上减少最外围的4个搜索点形成的。本文的搜索算法在对距离当前帧最近的参考帧使用大模板进行大范围的搜索以得到较好的运动矢量。在前面参考帧所得到的运动矢量的基础上,在后面的参考帧中都使用逐步减小的模板进行搜索。

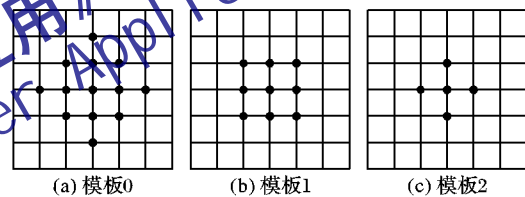


图2 逐渐缩小的菱形模板

2.2 算法描述

首先,将每帧相对于前一帧的运动矢量都保存下来,可以有效地定位编码块在每一帧的预测起点,从而减少搜索点数节约运动估计时间。每一个参考帧都会对前面所得的运动矢量进行修正,以求得到最接近最优运动矢量。然后,将所有参考帧分为4个部分:最接近当前帧的参考帧、比较接近当前帧的参考帧、偏离当前帧的参考帧和远离当前帧的参考帧。对于每个部分使用不同的搜索模板,对于距离当前帧最近的参考帧使用模板0所示的大菱形模板进行搜索;比较接近当前帧的参考帧使用菱形模板1进行搜索;偏离当前帧的参考帧使用菱形模板2进行搜索;远离当前帧的参考帧只对所预测的中心点进行搜索。算法步骤具体如下。

第1步 计算预测运动矢量。

从第2个参考帧开始,每一个参考帧的预测矢量都是根据前一参考帧的运动矢量计算得到的。设ref_Num为参考帧数, MV_{t-i-1} 为参考帧 $F(t-i)$ ($0 < i < \text{ref_Num} + 1$)中每个 4×4 块通过前一参考帧 $F(t-i-1)$ 预测得到的运动矢量,当前的块E尺寸为 $M \times N$,由 MV_{t-i-1} 能够得到参考帧 $F(t-i)$ 中块E的每个像素点对应于前一参考帧 $F(t-i-1)$ 的运动矢量 $MV_{m,n,(t-i-1)}$,所以当前块E在 $F(t-i)$ 到 $F(t-i-1)$ 中的平均运动矢量为:

$$\overline{MV}_{t-i-1}(E) = \frac{1}{M \times N} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} MV_{m,n,(t-i-1)} \quad (2)$$

则当前块E在参考帧 $F(t-i)$ ($i = 2, 3, \dots, \text{ref_Num}$)的搜索起始点为:

$$MVP_{t-i} = MVP_{t-i+1} + \overline{MV}_{t-i}(E) \quad (3)$$

第2步 运动搜索。

设 ref 为当前参考帧,另设两个变量 t 和 s ,且 $t = ref_Num/4, s = ref_Num\%4$;

1) 若 $0 \leq ref < t$,当前参考帧属于最接近当前帧的参考帧,使用模板0进行搜索;

2) 若 $t \leq ref < (2t + s/3)$,当前参考帧属于比较接近当前帧的参考帧,使用模板1进行搜索;

3) 若 $(2t + s/3) \leq ref < (3t + s/3 + s/2)$,当前参考帧属于偏离当前帧的参考帧,使用模板2进行搜索;

4) 若 $(3t + s/3 + s/2) \leq ref$,当前参考帧属于远离当前帧的参考帧,经过前面参考帧的运动搜索,每帧都对运动矢量进行修正,可以认为此时的运动矢量已经很接近最佳参考点,所以只对所预测的中心点进行计算。

第3步 判断是否已到最后一个参考帧。若是,结束算法;若不是,返回第1步。

3 实验结果与分析

为了验证新算法的性能,使用 H. 264/AVC 参考模型 JM10.2 在相同条件下和本文算法进行了计算机仿真对比实验。实验中采用 2 个 QCIF 格式的视频序列和 4 个 CIF 格式的视频序列。主要的编码参数设置如下:搜索范围为 16,5 个参考帧,帧率为 30 fps, QP 值为 28,编码结构为 IPPP,对每一个序列编码帧数为前 50 帧,CAVLC 熵编码,Hadamard 滤波,率失真优化(RDO)。实验结果如表 2 所示,其中 PSNR 为 Y 分量。

从表 2 可以看出,与参考模型相比,本文所提出的算法可平均减少 79.84% 的运动搜索的点数,从而大幅度减少运动搜索时间,提高编码速度。同时 PSNR 平均只下降了 0.07 dB。本文提出的快速搜索算法有以下两个特点。

1) 本文充分利用了最佳参考帧位置的统计特性,将参考帧分成 4 种类型,对不同类型的参考帧使用不同的搜索模板,从而有效地减少运动搜索的点数,在图像质量和比特率变化很小的情况下,可以较大幅度提高编码速度。

2) 本文的多参考帧运动搜索的基本搜索模式采用菱形搜索模板,其搜索方式相对简单,易于硬件实现,而参考模型搜索模式较多,判断复杂,硬件实现较困难。

表2 本文算法与参考模型的实验结果比较

格式	序列	搜索点数			PSNR/dB		
		UMH	本文算法	变化率/%	UMH	本文算法	变化量
QCIF	Coastguard	38 044 318	6 782 469	-81.65	34.02	33.94	-0.08
	Silent	17 559 837	4 239 501	-65.19	35.95	35.92	-0.03
CIF	Bus	162 673 808	32 120 036	-79.82	34.86	34.76	-0.10
	Coastguard	171 476 941	29 945 481	-82.80	34.35	34.31	-0.04
	Silent	69 504 474	17 524 088	-64.18	35.96	35.92	-0.04
	Stefan	122 597 329	26 662 909	-76.64	35.65	35.52	-0.13
平均值		96 976 118	19 545 747	-79.84	—	—	-0.07

注:平均值只保留两位小数(四舍五入)。

4 结语

本文根据最佳参考帧位置的统计特性,提出了一个低复杂度的快速多参考帧搜索算法。实验结果表明,本文提出的算法在保持图像质量和码率的前提下平均可以节省接近 80% 的搜索点数,大幅度提高编码速度,这对 H. 264 编码的实时和硬件实现很有利。

参考文献:

- [1] WIEGAND T, SULLIVAN G J, LUTHRA A. Draft ITU-T recommendation and final draft international standard of joint video specification. Part 1: s. n.], 2003.
- [2] Joint Video Team (JVT) reference software[EB/OL]. [2009-08-25]. <http://bs.hhi.de/~suehring/tml/download/>.
- [3] 黄军浩,毕笃彦,许悦雷,等.基于 H. 264 的多参考帧运动估计快速算法[J].计算机工程,2006,32(12):238-290.
- [4] ZHU SHAN, MA KAI-KUANG. A new diamond search algorithm for fast block-matching motion estimation [J]. ICICS IEEE Transactions on Image Processing, 2000, 9(2): 287-290.
- [5] ZHANG SHUFANG, WANG YANLING, KANG JIN-HUA, et al. A new approach to fast multiple reference frame motion estimation for H. 264 [C]// ISCSCT '08: International Symposium on Computer Science and Computational Technology. Washington, DC: IEEE Press, 2008, 2: 254-258.
- [6] SU Y-P, SUN M-T. Fast multiple reference frame motion estimation for H. 264/AVC[J]. IEEE Transactions on Circuits and System for Video Technology, 2006, 16(3): 447-452.

(上接第 1625 页)

- [3] 杨琬,吴乐华,李淑云,等.基于感兴趣区域的图像质量评价方法[J].计算机应用,2008,28(5):1310-1312.
- [4] 王宇庆,刘维亚,王勇.一种基于局部方差和结构相似度的图像质量评价方法[J].光子·激光,2008,19(11):1546-1553.
- [5] 李航,路羊,崔慧娟,等.基于频域的结构相似度的图像质量评价方法[J].清华大学学报:自然科学版,2009,49(4):559-562.
- [6] CHEN GUAN-HAO, YANG CHUN-LING, PO LAI-MAN, et al. Edge-based structural similarity for image quality assessment [C]// Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing. Washington, DC: IEEE, 2006: 14-19.
- [7] CHEN GUAN-HAO, YANG CHUN-LING, XIE SHENG-LI. Gradient-based structural similarity for image quality assessment [C]// Proceedings of IEEE International Conference on Image Processing. Washington, DC: IEEE, 2006: 2929-2932.
- [8] WANG ZHOU, SHANG XIM-LI. Spatial pooling strategies for perceptual image quality assessment [C]// Proceedings of IEEE International Conference on Image Processing. Washington, DC: IEEE, 2006: 2945-2948.
- [9] 叶盛楠,苏开娜,肖创柏,等.基于结构信息提取的图像质量评价[J].电子学报,2008,36(5):856-861.
- [10] SHEIKH H R, WANG ZHOU, CORMACK L, et al. LIVE Image Quality Assessment Database Release 2 [EB/OL]. [2010-05-10]. <http://live.ece.utexas.edu/research/quality>.
- [11] SHEIKH H R, SABIR M F, BOVIK A C. A statistical evaluation of recent full reference image quality assessment algorithms [J]. IEEE Transactions on Image Processing, 2006,15(11):3440-3451.
- [12] ROUSE D M, HEMAMI S S. Understanding and simplifying the structural similarity metric [C]// Proceedings of 15th IEEE International Conference on Image Processing. Washington, DC: IEEE, 2008: 1188-1191.
- [13] Video Quality Experts Group. Final report from the video quality experts group on the validation of objective models of video quality assessment [EB/OL]. [2000-03-05]. <http://www.vqeg.org/>.