H.264 整像素快速运动估计算法研究

杨云飞[1],韩彦芳[1],张定会[1],李 竹[2]
YANG Yun-fei[1],HAN Yan-fang[1],ZHANG Ding-hui[1],LI Zhu[2]

1.上海理工大学 光学与电子信息工程学院,上海 200093
2.山西师范大学 物理与信息工程学院,山西 临汾 041004
1.College of Optical and Electronic Information Engineering,University of Shanghai for Science and Technology,Shanghai 200093,China
2.College of Physics and Information Engineering,Shanxi Normal University,Linfen,Shanxi 041004,China
E-mail:yunfeiy@foxmail.com

YANG Yun-fei,HAN Yan-fang,ZHANG Ding-hui,et al.Study on integer fast motion estimation for H.264.Computer Engineering and Applications,2009,45(35):196-198.

Abstract: A new fast motion estimation algorithm,Direction Based Diamond-T Search(DBDTS),is proposed based on analyzing the integer motion estimation algorithm in JM11.0 of H.264.The algorithm efficiently reduces the time of motion estimation with minimal loss in bitrate and reconstructed quality.Then,it is tested by typical video sequences and comparisons with FS,EPZS and UMHexagonS are made.Experimental results show that the improved algorithm reduces the time of motion estimation efficiently while keeping similar visual quality.

Key words: H.264;block based motion estimation;search model;fast searching algorithm;Direction Based Diamond-T Search(DBDTS)

摘 要:通过对 H.264 参考模型 JM11.0 中整像素运动估计算法的分析,提出了一种新的快速运动估计搜索算法:基于方向预测的菱形-T 形搜索算法(Direction Based Diamond-T Search,DBDTS)。改进算法在保证视频序列各分量信噪比和输出比特率的同时有效地减少了运动估计时间。分别应用 FS、EPZS、UMHexagonS 和改进算法对典型序列进行测试。测试结果表明,新算法在保证编码性能的同时,有效地提高了运动估计速度。

关键词:H.264 标准;块运动估计;搜索模型;快速搜索算法;基于方向预测的菱形-T 形搜索(DBDTS)算法
DOI:10.3778/j.issn.1002-8331.2009.35.059 **文章编号**:1002-8331(2009)35-0196-03 **文献标识码**:A **中图分类号**:TP919.81

1 引言

H.264/AVC 是由 ISO/IEC MPEG 和 ITU.T VCEG 联合制定的最新的视频编码标准[1]。在标准中,基于块的运动估计算法(Block Matching Algorithm,BMA)是其中重要的组成部分,用来消除运动序列的时间冗余度。运动估计明显地提高了编码的效率,但却耗费了大量的运算时间。

目前,基于块匹配运动估计算法中,最细致的搜索方法是全搜索(Full Search,FS),即在搜索区内逐点搜索,每搜索一点计算一次准则函数值,例如选择 MAD 为准则函数,当 MAD 达到最小值时,求得最佳匹配像素块。该算法能得到最优的运动矢量,但运算量相当大,不适合实时应用。为了减少搜索次数,早期研究人员提出了多种快速搜索方法,如三步搜索法(TSS)[2]、二维对数搜索法(2DLOG)[3]、四步搜索法(FSS)[4]、新三步搜索法(NTSS)[5]、预测搜索法[6]等。除了原来的矩形窗搜索模式外,还出现了菱形模式、六边形模式、十字形模式以及这些模式相结合的混合模式等[7-8]。

菱形搜索算法相对全搜索算法而言,在保证视频质量的同时能有效地减少搜索时间。但在一些诸如明显的全局运动或场景变换的情况下,菱形搜索性能下降明显,并且容易陷入局部最优,缺少鲁棒性。文章在菱形搜索的基础上,考虑到自然视频序列运动的惯性特性,提出了一种基于运动方向的菱形-T 形搜索(Direction Based Diamond-T Search,DBDTS)算法,并对算法进行实验及性能分析。

2 JM 模型中整像素搜索算法

JM 模型中提供 4 种可选整像素运动搜索算法[9],分别为全搜索(FS)算法、UMHexagonS 算法、Simplified UMHexagonS 以及 EPZS 算法。Simplified UMHexagonS 是对 UMHexagonS 的简化,以下分析 UMHexagonS 和 EPZS 的搜索模板及算法特点。

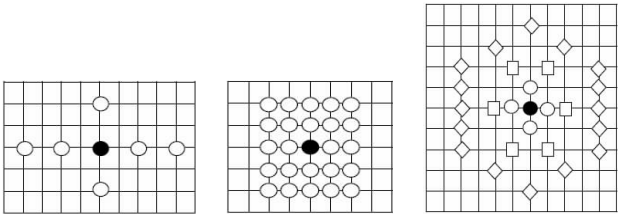
基金项目:山西省高校科技开发研究项目(the University Science and Technology Research of Shanxi Province Foundation of China under Grant No.2007128)。

作者简介:杨云飞(1984-),男,硕士研究生,主要研究领域为视频图像的获取与处理;韩彦芳(1974-),女,博士研究生,硕士生导师,主要研究领域为模式识别,图像处理;张定会(1964-),男,博士生,教授,硕士生导师,主要研究领域为小波分析,信息处理及故障检测;李竹(1975-),男,硕士,主要研究领域为通信与信息系统,信号处理。

收稿日期:2008-07-09 **修回日期**:2008-10-21

2.1　UMHexagonS 算法

UMHexagonS 算法的搜索模板有非对称十字形模板、正方形螺旋搜索模板、中六边形及超六边形模板和小菱形搜索模板，如图 1 所示。



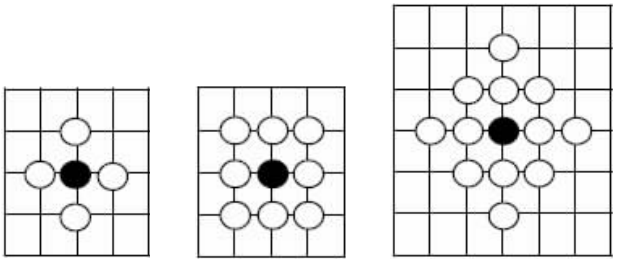
(a)非对称性十字模板　(b)正方形螺旋搜索模板　(c)超六边形、中六边形、小菱形模板

图 1　UMHexagonS 算法搜索模板

UMHexagonS 算法起点预测准确，因为其采用的起点预测综合利用了帧内、帧间相邻块的运动矢量相关性，以及 H.264 采用的宏块划分技术所带来的不同尺寸块的运动矢量相关性，因此能选出较好的初始搜索点，准确率高。并且该算法采用了内容自适应的搜索模板和搜索方式，对不同内容的块进行了不同的搜索，搜索性能得到进一步改善[10]。

2.2　EPZS 算法

EPZS 算法使用 3 种搜索模板，包括小菱形(DS)、正方形 EPZS 和扩展 EPZS。如图 2 所示。



(a)DS　(b)EPZS²　(c)extEPZS

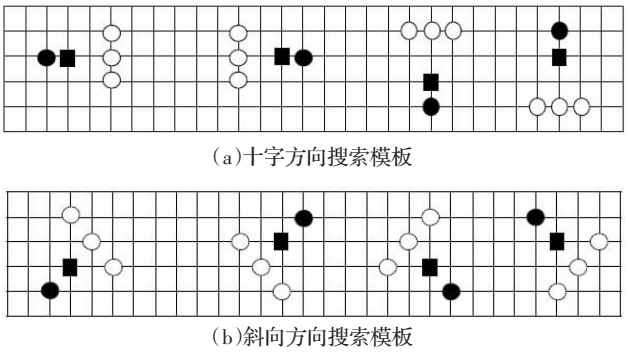
图 2　EPZS 算法搜索模板

EPZS 算法由局部开始逼近最佳点，该算法虽然可以较好地避免落入局部最优，但是仍有其局限性，即对于搜索范围比较小、运动比较平缓的情况效果较好，但在搜索范围较大和运动剧烈的情况下，很容易在搜索初期就落入局部最优。因此如何避免落入局部最优，在较大范围内找到最优的匹配点，仍然是运动估计研究的重点[10]。

3　基于方向预测的菱形–T 形搜索(DBDTS)算法

3.1　DBDTS 搜索模板

对于自然图像序列，由于物体运动的惯性，各图像块的最佳匹配块呈现明显的方向性(梯度方向)。在对参考图像进行匹配块搜索时，考虑到这种方向性特点，下一个候选像素在已搜索最佳像素点的基础上确定，具体实现可通过设置两种大小和方向不同的 T 形搜索模板，如图 3 所示。考虑到大部分的自然图像序列在水平和垂直方向上的运动情况概率大，若当前最佳匹配像素点在上一匹配点的水平或垂直(Cross)方向，则利用大 T 形模板进一步搜索。反之，若当前最佳匹配像素点在上一匹配点斜向方向，则利用斜向的小 T 形模板进一步搜索。



(a)十字方向搜索模板

(b)斜向方向搜索模板

图 3　DBDTS 搜索模板

3.2　DBDTS 搜索

步骤 1　小菱形搜索。从 5 个点中搜索具有最小 SAD 值的像素点。如果最佳匹配点在菱形中心，则结束搜索返回，如图 4(a)所示。否则算法进入步骤 2。

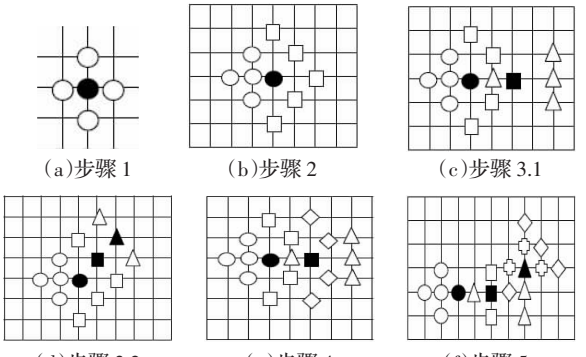
步骤 2　大菱形搜索。以上一步得到的最佳匹配点为中心进行大菱形搜索，如图 4(b)所示。如果最佳匹配点在中心(起始点)则跳到步骤 5。否则，如果最佳匹配点在起始点的水平或垂直方向，则跳到步骤 3.1 进行大 T 形搜索；如果最佳匹配点在斜向方向，跳到步骤 3.2 进行小 T 形搜索。

步骤 3.1　大 T 形搜索。以上一步得到的最佳匹配点为中心选择如图 3(a)中的相应模板进行大 T 形搜索。如果最佳匹配点是黑色圆圈点，则跳到步骤 5；在黑色矩形点则跳到步骤 4；否则判断方向重复步骤 3，如图 4(c)所示。

步骤 3.2　小 T 形搜索。以上一步得到的最佳匹配点为中心选择如图 3(b)中的相应模板进行小 T 形搜索。如果最佳匹配点是黑色矩形或黑色圆圈点，则跳到步骤 5；否则判断方向重复步骤 3，如图 4(d)所示。

步骤 4　进行一次大菱形搜索，如图 4(e)所示。在起始点跳到步骤 5。否则返回步骤 3 判断方向重复搜索。

步骤 5　用小菱形搜索四个点后结束，如图 4(f)所示。



(a)步骤 1　(b)步骤 2　(c)步骤 3.1

(d)步骤 3.2　(e)步骤 4　(f)步骤 5

说明：● 第一步获得的最佳点
■ 第二步获得最佳点
▲第三步获得的最佳点
○ 第一步 □ 第二步 △ 第三步 ◇ 第四步 ⌑ 第五步

图 4　DBDTS 搜索步骤

4　实验结果与分析

4.1　实验结果

使用 JM11.0 作为测试模型，对 qcif 格式的视频序列进行测试。参数设置：量化参数 *QP*=28，选用 5 个参考帧，帧率为 30

frames/s，搜索范围为 16 个像素点，编码序列为 I-P-P-P...格式，采用 Hadamard 变换和 CABAC 熵编码，编码 60 帧。实验平台：2.66 GHz CPU，256 MB 内存，VC++6.0 开发工具。

为了客观评价该文算法的性能，考虑以下 3 个因素：*Y* 分量峰值信噪比(PSNR)，用来表明压缩后图像质量；运动估计时间(MET)用来考察运动搜索时间；输出比特率(kbit/s)，用来考察压缩率。实验测试结果如表 1~表 3 所示。表 1 中的性能指该文算法相对前 3 种算法 PSNR 减少值；表 2 中的性能指该文算法相对前 3 种算法比特率增加值；表 3 中性能指该文算法相对前 3 种算法运动估计时间减少的百分比。

表 1 PSNR 性能比较 dB

测试序列	FS	UMHexagonS	EPZS	DBDTS
silent.qcif	36.120	36.070	36.120	36.09
Sign_Irene.qcif	37.410	37.400	37.430	37.36
foreman.qcif	35.910	35.920	35.910	35.86
container.qcif	36.310	36.310	36.310	36.30
claire.qcif	39.930	39.870	39.910	39.78
highway.qcif	37.960	37.900	37.910	37.87
平均	37.273	37.245	37.265	37.21
性能	−0.063	−0.035	−0.055	0

表 2 输出比特率性能比较 kbit·s^{-1}

测试序列	FS	UMHexagonS	EPZS	DBDTS
silent.qcif	53.990	53.730	53.990	54.41
Sign_Irene.qcif	79.720	80.010	79.610	81.09
foreman.qcif	75.400	74.700	74.700	75.96
container.qcif	27.000	27.120	27.210	27.06
claire.qcif	22.780	22.700	22.700	22.75
highway.qcif	40.730	42.110	41.350	44.19
平均	49.937	50.062	49.927	50.91
性能	0.973	0.848	0.983	0

表 3 运动估计时间(MET)性能比较 s

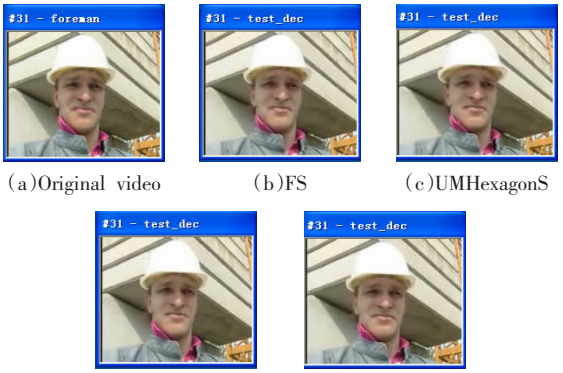
测试序列	FS	UMHexagonS	EPZS	DBDTS
silent.qcif	153.316	66.975	78.014	43.564
Sign_Irene.qcif	161.940	71.443	79.789	37.796
foreman.qcif	173.312	80.998	88.389	57.524
container.qcif	175.266	62.072	81.690	38.448
claire.qcif	184.509	64.726	74.972	27.587
highway.qcif	174.995	64.783	73.490	43.798
平均	170.556	68.499	79.390	41.452
性能	75.695%	39.484%	47.786%	0

为从主观考察视频质量，利用测试模型中的解码器分别对各中算法编码生成的码流进行解码，得到 YUV 格式视频。图 5 为各种算法的的解码效果(显示第 31 帧)。

4.2 性能分析

从实验结果可以看出，相对于 FS、UMHexagonS 和 EPZS，该文算法在输出视频质量(PSNR 及主观效果)及压缩率都没有明显变化的情况下(PSNR 降低在 0.01 dB 以下，输出比特率增加在 1 kbit/s 以下)，运动搜索所用时间大幅度下降：分别下降了 FS 的 75%，UMHexagonS 的 40%，EPZS 的 48%。

然而，DBDTS 算法还有一些不足之处。由于采用基于方向的搜索模板，在某些场合下也会出现局部最优点的情况。例如，



(a)Original video (b)FS (c)UMHexagonS

(d)EPZS (e)DBDTS

图 5 编码性能的主观对比

对于 highway 序列，其输出比特率为 44.19 kbit/s，相对 FS 算法的 40.73 kbit/s 增加了近 4 kbit/s，在很多程度上影响了视频的有效传输。此算法的局部最优点预防是今后研究的方向。

5 结论

文章在分析 JM 测试模型的基础上提出了一种新的快速整像素运动估计搜索(DBDTS)算法，与常用的标准算法进行实验对比，结果证明该文算法有效地减少了运动估计时间，提高了 H.264 的编码效率。

参考文献：

[1] Wiegand T.Overview of the H.264/AVC video coding standard[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003,13(7):560-576.

[2] Koga T,Iinuma K,Hirano A,et al.Motion compensated interframe coding for video conferencing[C]//Proc Nat Telecommon Conf.New Orleans:LA,1981.

[3] Jain J R,Jain A K.Displacement measurement and its application in inter frame image coding[J].IEEE Trans Commun,1984(29): 1799-1808.

[4] Po L M,Ma W C.A novel four-step search algorithm for fast block motion estimation[J].IEEE Trans Circuit Syst,1996(6):313-317.

[5] Li R,Zeng B,Liou M L,A new three-step search algorithm for block motion estimation[J].IEEE Trans Circuits Syst,1994(4):438-443.

[6] Xu J B,Po L M,Cheung C K.A new prediction model search algorithm for fast block motion estimation[C]//Proc ICIP 1997,1997, 3:610-613.

[7] Zhu S,Ma K K.A new diamond search algorithm for fast block-matching motion estimation[C]//Proc Int Conf Inform Commun Signal Process,1997:292-296.

[8] Li W,Salari E.Successive elimination algorithm for motion estimation[J].IEEE Trans Image Processing,1995,4(1):105-107.

[9] Zhibo C,Peng Z,Yun H.Fast integer pixel and fractional pixel motion estimation for JVT,JVT-F017.doc[C]//JVT 6th Meeting,Awaji,Japan,2002.

[10] 付金良，张佳申.H.264/AVC 中快速运动估计算法研究[J].软件导刊,2008,7(4):41-43.

[11] Jeon G,Kim J,Jeong J.Enhanced cross-diamond search algorithm for fast block motion estimation[J].Image Analysis and Recognition,2007,4633:481-490.