

# 运动估计快速半像素级搜索方法的实验研究\*

陈维强 高文 赵建斌 傅宇卓

(哈尔滨工业大学, 哈尔滨 150001)

**摘要** 提出了3种运动估计快速半像素级搜索方法。实验分析表明这3种方法具有较好的MSE性能和解码图象质量;3种快速方法中最慢的可以节省约29.17%运算次数,最快的可节省近58.33%运算次数。本文所提出的快速方法可作为运动估计单元中半像素级搜索实时实现的候选算法。

**关键词** 运动估计,半像素级搜索,视频编码

## 1 引言

在MPEG-2<sup>[1]</sup>视频编码器的实时实现中,最具挑战性的是运动估计单元,其原因主要表现为:输入编码图象的分辨率高,如MPEG-2 MP@ML广播级支持ITU-R-601格式(即:720pixels × 480lines × 30frames/sec或相当),是MPEG-1所支持的SIF格式的4倍;允许多种运动估计类型并存,即允许宏块同时进行16×16、16×8块以及dual prime运动估计(当编码序列中没有定义B类型图象时)。

一般情况下,运动估计的搜索是由2部分构成:整像素级搜索和半像素级搜索。显然,整像素级搜索是视频编码运动估计单元实现的主要问题。对于这个问题,到目前为止,已经进行了许多研究<sup>[2,3]</sup>,并且还在继续进行。另一个问题是半像素级搜索的实现,实验表明进行半像素级搜索能明显提高运动补偿的效果(可参见本文表1和表2实验结果),因此,一般说来编码时半像素级搜索是不可缺少的,尤其对MPEG-2视频编码器(追求高的解码图象质量)来讲更是如此。

半像素级搜索虽然远比不上整像素级搜索那么复杂,但对MPEG-2视频编码(如MP@ML广播级)而言,还是具有一定的难度。例如在文献<sup>[4]</sup>中,采用4片数字信号处理芯片(DSP)并行工作以完成

MPEG-2 MP@ML广播级的半像素级搜索。因此,研究快速的半像素级搜索方法是非常必需的,而且具有实际应用价值。然而,到目前为止并未见有任何关于快速半像素级搜索方法的研究报道。

## 2 快速半像素级搜索方法提出

在MPEG-2视频编码标准中,编码图象通常被分成3种类型:即I图象、P图象、B图象,其中I图象一般只做帧内编码,P图象做前向预测编码,B图象做双向预测编码。显见对P和B图象进行编码时需要采用运动估计。下面首先简述MPEG-2测试模型5(TM5)<sup>[4]</sup>中的运动估计半像素搜索,然后给出快速半像素级搜索方法。

### 2.1 运动估计的半像素级搜索

半像素级搜索是指在整像素级搜索的基础上,为了寻找更好的匹配而进一步在像素的插值空间上所进行的搜索。由于图象像素点之间存在采样时间间隔,因而最好的匹配块有可能不是整像素块,而是存在于插值空间上,因此有必要进行半像素搜索。在TM5中,首先在整像素级上利用搜索算法(采用全搜索算法FSA),寻找到最佳匹配块,例如图1所示,设以D0为左上角的块是待估计块在整像素级上的最佳匹配块,假设此时的TAD(Total Absolute Dif-

\* 本文得到国家863计划智能计算机系统主题(863-306-01)和国家“九五”攻关项目(96-B07-3-1)的支持  
收稿日期:1997-11-13;收到修改稿日期:1998-01-04

ference)值记为  $tad_0$ ; 然后沿  $D_0$  周围的 8 个象素点 ( $D_1 \sim D_8$ ) 依次作半象素搜索(本文称这种方法为半象素级全搜索, 记为 FHPS):

(1) 做插值形成参考数据块记为  $data'[,] [.]$ : 当搜索到  $D_2$  或  $D_6$  时, 即与  $D_0$  同在一列上, 则在列方向做均值插值形成参考数据块  $data'$ ; 当搜索到  $D_4$  或  $D_8$  时, 即与  $D_0$  同在一行上, 则在行方向做均值插值形成参考数据块  $data'$ ; 当搜索到  $D_1, D_3, D_5$  或  $D_7$  时, 即与  $D_0$  既不在同在一列上又不在同一行, 则在行方向列方向同时做均值插值形成参考数据块  $data'$ .

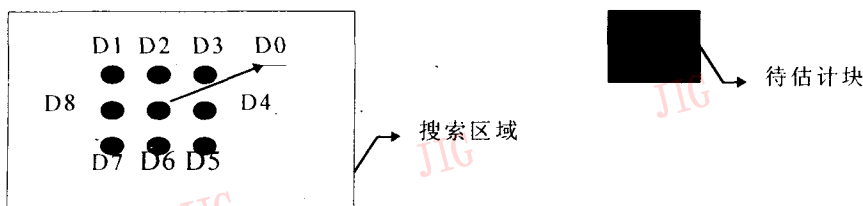


图1 半象素搜索

## 2.2 快速半象素级搜索方法

由上小节可知, TM5 所采用的半象素级搜索方法是一种完全搜索方法, 因此, 相对而言计算量较大. 假设在小范围内搜索时, TAD 值是单调的. 考虑到在搜索中对位于对角线“×”上的点(即  $D_1, D_5, D_3$  和  $D_7$ )的计算量要大于位于“+”上的点(即  $D_2, D_6, D_4$  和  $D_8$ )的计算量. 根据这一点和上述假设, 我们提出了以下 3 种快速半象素级搜索方法.

**方法 1:** 只搜索“+”方向上的 4 个点 ( $D_2, D_6, D_4$  和  $D_8$ ), 比较相应的  $tad_2, tad_4, tad_6, tad_8$  和  $tad_0$ , 并以其中 TAD 值最小者所对应的块为最佳匹配块. 该方法节省了对 4 个“×”方向上点的搜索.

**方法 2:** 首先, 搜索“+”方向上的 4 个点 ( $D_2, D_6, D_4$  和  $D_8$ ), 比较相应的  $tad_2, tad_4, tad_6$  和  $tad_8$ , 得到其最小值所对应的点记为  $D'$  ( $D'$  可能是  $D_2, D_4, D_6$  或  $D_8$ ), 假设  $D' = D_2$ ; 然后, 比较与  $D'$  ( $D_2$ , 其它情况依此类推) 既不在同一行又不在同一列的 2 个点的大小, 如  $D_4$  和  $D_8$ , 假设  $D_4$  小于  $D_8$ , 搜索  $D_3$  (“×”方向上的点), 比较  $tad_2, tad_3$  和  $tad_0$ , 以其中 TAD 值最小者所对应的块为最佳匹配块, 如果  $D_8$  小于  $D_4$  同理. 该方法节省了对 3 个“×”方向上点的搜索.

**方法 3:** 首先, 与方法 2 相似, 搜索“+”方向上

(2) 计算并记录每次搜索的 TAD 值, 设为  $tad_1 \sim tad_8, tad_0 \sim tad_8$  中的最小值者所对应的参考块为最佳匹配块.

(3) 运动向量的表示: 运动向量是指参考块(最佳匹配块)相对于待估计块的位移, 例如待估计块左上角位置是  $(x, y)$ , 参考块左上角位置是  $(x+k, y+l)$ , 则运动向量为  $(k, l)$ . 在 MPEG 标准中, 为了表示半象素, 都是将  $k, l$  左移一位, 然后, 用最低位表示半象素, 最低位为 ‘1’ 表示有半象素, ‘0’ 则表示没有.

的 4 个点 ( $D_2, D_6, D_4$  和  $D_8$ ), 比较相应的  $tad_2, tad_4, tad_6$  和  $tad_8$ , 得到其最小值所对应的点记为  $D'$  ( $D'$  可能是  $D_2, D_4, D_6$  或  $D_8$ ), 假设  $D' = D_2$ ; 然后, 比较  $D'$  ( $D_2$ , 其它情况依此类推)、与  $D'$  相邻的 2 个“×”方向上的点 ( $D_1$  和  $D_3$ ) 以及  $D_0$  的 TAD 值, 并以其中 TAD 值最小者所对应的块为最佳匹配块. 该方法节省了对 2 个“×”方向上点的搜索.

## 3 比较实验与分析

为了检验所提出的快速半象素级搜索方法性能, 以下从 2 个方面进行实验和分析比较: 即运动估计的平均平方误差 MSE、解码图象的峰值信噪比 PSNR 的实验比较和运算次数分析比较. 参与比较的方法有: 半象素全搜索方法 (FHPS)、不带半象素搜索 (记为 NHPS) 以及本文提出的方法 1、2、3.

### 3.1 MSE 与解码峰值信噪比 PSNR 比较

实验是在 3 组代表不同情况的标准测试序列上进行, 即在 football, flower garden 和 mobile & calendar 的各自前 31 帧图象上进行的. 运动估计与编码时, 图象组 GOP 的结构设为:  $N=15, M=3$ , 且搜索窗的大小为  $32 \times 32$  pixels, 编码时比特流速率设为 6.4 Mbits/s, 整象素级搜索是采用全搜索算法

FSA。这些比较实验都是在 TM5 测试模型基础上进行的。

表 1 是 MSE 的平均值,表中最后一栏是 3 个测试序列上的 MSE 总平均值,由此可以得到,3 种快速方法中,方法 3 的 MSE 值最低,只高出 FHPS 约 0.07,方法 1 的 MSE 值最高,约高出 FHPS 方法 13.58,但低于 NHPS 方法 42.72,明显优于 NHPS。

表 1 MSE 的平均值结果

测试序列	FHPS	NHPS	方法 1	方法 2	方法 3
football	145.56	164.86	149.34	146.09	145.61
flower garden	78.90	123.12	90.34	79.12	78.93
mobile&calendar	120.42	222.81	142.93	121.06	120.54
总平均值	114.96	170.26	127.54	115.42	115.03

表 2 是 PSNR 的平均值,表中最后一栏是 3 个测试序列上的 PSNR 总平均值,显见,3 种快速方法中,方法 3 的 PSNR 值最高,只低于 FHPS 约 0.01dB,方法 1 的 PSNR 值最低,约低于 FHPS 方

法 1.25dB,但高于 NHPS 方法 2.12dB,明显优于 NHPS,并且主观实验也证实了这一点。

表 2 PSNR 的平均值结果(dB)

测试序列	FHPS	NHPS	方法 1	方法 2	方法 3
football	36.35	35.14	36.09	36.31	36.34
flower garden	35.28	31.89	33.93	35.25	35.28
mobile&calendar	34.21	28.71	32.07	34.13	34.20
总平均值	35.28	31.91	34.03	35.23	35.27

为了更清楚地展现所提出方法的性能,以下在图 2 和图 3 中分别给出了 football 序列的 MSE 和 PSNR 差值曲线,其中 I 图象的 MSE 差值统一给定为零。MSE 差值是 NHPS、方法 1、方法 2、方法 3 与 FHPS 的对应帧 MSE 之差;而 PSNR 差值是 FHPS 与 NHPS、方法 1、方法 2、方法 3 的对应帧 PSNR 之差。由图 2 与图 3 可知,方法 3 与 FHPS 的 MSE 和 PSNR 最接近,然后依次是方法 2、方法 1,最后是 NHPS,并且方法 1 明显优于 NHPS。

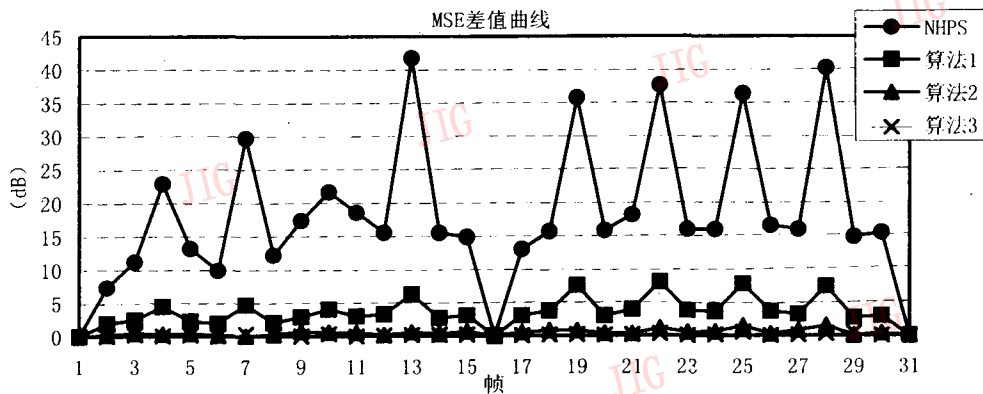


图 2 MSE 差值曲线 (football)

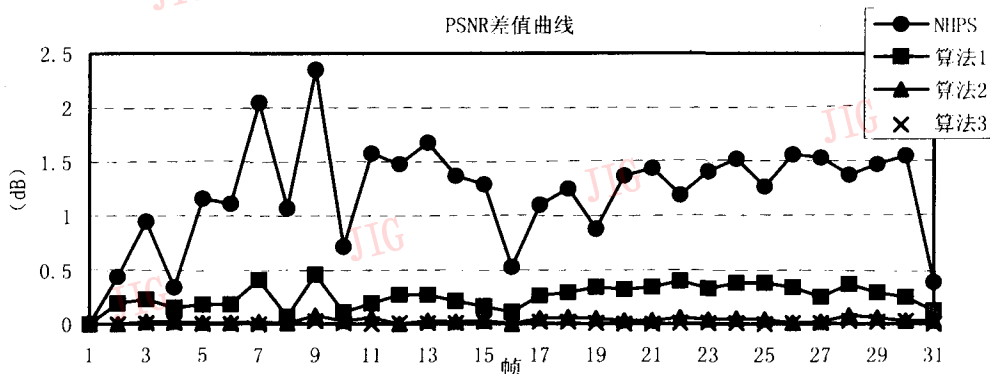


图 3 PSNR 差值曲线 (football)

### 3.2 运算次数分析对比

根据本文第 2 节可知,在进行半像素级搜索时,搜索对角线(“×”)方向的点要比搜索十字(“+”)方向的点运算量要大。设进行运动估计的块尺寸大小为  $W \times H$ ,对 MPEG-2 而言, $W=16, H=16$  或 8。则搜索一个对角线方向点的运算次数为: $7 \times W \times H - 1$ ;搜索一个十字方向点的运算次数为: $5 \times W \times H - 1$ 。设  $W=16, H=16$ ,表 3 为 FHPS 和方法 1~3 的一次半像素级搜索的运算次数估计。

表 3 运算次数估计

FHPS	方法 1	方法 2	方法 3
12 288	5 120	6 914	8 704
0.00%	58.33%	43.73%	29.17%

表 3 中最后一栏表示方法 1~3 相对于 FHPS 的节省运算次数占 FHPS 总运算次数的百分比。由表 3 可以得到:节省运算次数最多的是方法 1,超过一半;其次是方法 2;节省运算次数最少的是方法 3,但也能节省近 30%的运算次数。

## 4 结 论

本文提出了 3 种快速半像素级搜索方法,经实验及分析主要可得到如下结论:所提出的方法 3 的 MSE 和 PSNR 性能最好,其 MSE 只高出半像素级全搜索方法(FHPS)约 0.07,PSNR 低于 FHPS 约

0.01dB;其次是方法 2,MSE 高出 FHPS 约 0.46,PSNR 低于 FHPS 约 0.05dB;最差的是方法 1,其 MSE 高出 FHPS 约 13.58,PSNR 低于 FHPS 约 1.25dB,但方法 1 明显优于不进行半像素搜索情况。相对于 FHPS,节省运算次数最多的是方法 1,节省约 58.33%;其次是方法 2,节省约 43.73%;方法 3 节省最少,为 29.17%。

总之,从应用角度来看,方法 2 和方法 3 可以作为 MPEG-2 运动估计单元实时实现的半像素级搜索候选方法,方法 1 也可应用于可视电话或电视会议等场合。故本文工作为运动估计的半像素级搜索提供了快速算法,并且填补了对快速半像素级搜索方法研究的空白。

### 参 考 文 献

- 1 ISO/IEC 13818-2. Information Technology—Generic Coding of Moving Pictures and Associated Audio, Video. ISO/IEC JTC1/SC29/WG11,1994.
- 2 Pan S B, Chae S S, Park R H. VLSI architectures for block matching algorithm using systolic arrays. IEEE Transactions on Circuits and Systems for Video Technology,1996, 6(1): 67~73.
- 3 Komarek T, Pirsch P. Array architectures for block matching algorithms. IEEE Transactions on Circuits and Systems,1989, 36(10): 1301~1308.
- 4 MPEG2 Test Model 5. Document ISO/IEC JTC1/SC29/WG11/93-400. Test Model Editing Committee, 1993.
- 5 Akiyama T, Aono H et al. MPEG-2 video codec using image compression DSP. IEEE Transaction on Consumer Electronic, 1994, 40(3): 466~472.



**陈维强** 1968 年生,1993 年获哈尔滨工业大学自动化仪表专业硕士学位,现为哈尔滨工业大学讲师、计算机应用专业在职博士生,目前主要研究领域是运动图象压缩编码技术、系统控制。



**高文** 1956 年生,1988、1991 年分别获哈尔滨工业大学计算机应用和日本东京大学电子学博士学位。现任哈尔滨工业大学教授,博士生导师,兼任国家 863 计划信息领域智能计算机主题专家组组长。主要研究领域为多媒体压缩与处理技术,智能计算机接口技术,虚拟现实技术,以及人工智能应用等。



**赵建斌** 1973年生,1996年获哈尔滨工业大学模式识别与智能控制学士学位,现为哈工大计算机系统与组织结构专业硕士研究生,目前主要从事运动估计专用结构的研究与实现。



**傅宇卓** 1968年生,1997年获哈工大计算机系统与组织结构专业硕士学位,现为该专业博士研究生,目前主要从事图象压缩专用结构的研究。

## Experiment Study on Fast Half Pixel Search Methods of Motion Estimation

Chen Weiqiang, Gao Wen, Zhao Jianbin, Fu Yuzhuo

(Harbin Institute of Technology, Harbin 150001)

**Abstract** In this paper, three methods of fast half pixel search of motion estimation have been proposed. Experiments show the three methods have good MSE performance and decoded picture quality; Analysis show that the slowest method can save compute times about 29.17% and the most fast can save about 58.33%. The proposed methods can be as candidate algorithms for the real time implementation of motion estimation half pixel search.

**Keywords** Motion estimation, Half pixel search, Video coding

## Designed for HP DesignJet 首批认证的软件商

1. 蒙泰科技发展有限公司
2. 北京凯思博宏应用工程有限公司
3. 北京市宇思机电公司
4. 华康计算机技术有限公司
5. 怡泓电子制版有限公司
6. 北京远潮信息技术有限责任公司
7. 北京侏罗纪技术发展有限责任公司
8. 中国建筑科学研究院电子计算中心
9. 北京高华计算机有限公司
10. 北京大凯电子技术有限责任公司
11. 浪潮 CAD 系统工程公司
12. 西安经纬正直软件有限责任公司
13. 北京北方德赛计算机技术有限责任公司
14. 武汉开目集成技术有限责任公司
15. 武汉华中软件公司
16. 北京华远软件工程有限公司
17. 北京市京渝天河计算机软件技术有限责任公司
18. 奥发软件工程(武汉)有限公司
19. 北京市正华电子技术公司
20. 武汉中地信息工程有限公司
21. 金豪(北京)电子系统有限公司