



# 一种新的快速两级块匹配运动估计算法<sup>1)</sup>

陈维强

(哈尔滨工业大学计算机科学与工程系 哈尔滨 150001)

高文

(中国科学院计算技术研究所 北京 100080)

**摘要** 针对 MPEG-2 视频编码运动估计, 提出了一种新的快速两级块匹配算法. 分析表明该算法的数据流量和计算量更小, 且更易于 VLSI 实现. 实验表明该算法具有很好的质量性能.

**关键词** MPEG-2 视频编码, 运动估计, 块匹配.

## A NEW FAST ALGORITHM FOR TWO STAGE BLOCK MATCHING MOTION ESTIMATION

CHEN Weiqiang

(Department of Computer Science and Engineering, Harbin Institute of Technology, Harbin 150001)

GAO Wen

(Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100080)

**Abstract** In this paper, we focus on MPEG-2 video coding motion estimation, and propose a new fast two-stage block matching algorithm. Analysis shows that the data rate and computing complexity of the proposed algorithm are smaller than other existing two-stage algorithms, and the new algorithm can be more easily implemented by VLSI technology. Moreover, experimental results show that the proposed algorithm has good quality performance.

**Key words** MPEG-2 video coding, motion estimation, block matching.

## 1 引言

MPEG-2<sup>[1]</sup> 视频编码运动估计的实时实现是具有挑战性的工作. 结合 MPEG-2 运动估计实现问题, 现有的研究存在如下不足: 1) 在文献[2]中, 报道了一个 MPEG-2 运动估

1) 国家自然科学基金、国家“八六三”计划智能计算机系统主题和国家“九五”攻关计划资助课题.

收稿日期 1997-07-03 收修改稿日期 1998-05-26

计单元的实现,采用了6个全搜索运动估计芯片并行工作,分别完成同一宏块的 $16 \times 16$ 块、 $16 \times 8$ 顶、底块的前向和后向运动估计,显然,如此实现的 MPEG-2 视频编码运动估计单元对总线与存储器速度要求很高,因而性能价格比较低;2) 文献[3]中提出了一种易于并行实现的快速两级块匹配算法,但运动估计的数据流量很大,与对应的 FSA 算法相等;3) 针对文献[3],文献[4]提出了一种基于投影技术搜索的 systolic 阵列,但其规模与搜索区域的大小相关,这意味着阵列的规模固定后将不能适应搜索区域的变化(变大或变小),缺乏灵活性.同时 MPEG-2 视频编码追求宽的搜索区域,这就是说如果采用文献[4]的阵列结构,那么阵列的规模将很大;4) 文献[3,4]中没有考虑 MPEG-2 视频编码运动估计的特点,即一个宏块需同时完成多种类型的运动估计.针对这些不足,本文提出一种新的适用于 MPEG-2 视频编码运动估计的快速两级块匹配算法,该算法的数据流量和计算量更小,且更易于 VLSI 实现.

## 2 一种新的快速两级块匹配算法

### 2.1 MPEG-2 视频编码运动估计

MPEG-2 视频编码标准支持三种编码方式:1) 帧/场编码方式;2) 场编码方式;3) 帧编码方式. 鉴于场编码方式与帧/场编码方式所支持的运动估计相类似,其主要差别在运动估计数据的组织方式上,而帧编码方式所支持的运动估计是帧/场编码方式的一个子集;另外 dual prime 运动估计的计算量主要是在同侧参考场中搜索,而这一搜索已经在 $16 \times 8$ (顶、底)块运动估计中完成.为简明起见,本文仅以帧/场编码方式下的运动估计(不含 dual prime)为例进行讨论.

$16 \times 8$ (顶、底)块运动估计的数据组织方式是将帧结构块数据按顶、底场的分割,相应的搜索区域也应进行同样的分割,并且 $16 \times 8$ (顶、底)块运动估计的结果将包括两个运动向量与两个场指示.用  $MB$  表示宏块,通过分割得到的 $16 \times 8$ 顶块与底块分别记为  $MB_{top}$ ,  $MB_{bot}$ .用  $SA$  表示宏块的相应搜索区域,通过分割得到的顶场与底场搜索区域分别记为  $SA_{top}$ ,  $SA_{bot}$ .显然有下式成立

$$MB = MB_{top} \cup MB_{bot}, SA = SA_{top} \cup SA_{bot}. \quad (1)$$

那么, $16 \times 8$ (顶、底)块运动估计的搜索包括:1)  $16 \times 8$ 顶块即  $MB_{top}$  在顶场搜索区域  $SA_{top}$  中的搜索与在  $SA_{bot}$  中的搜索;2)  $16 \times 8$ 底块即  $MB_{bot}$  在搜索区域  $SA_{top}$  中的搜索与在  $SA_{bot}$  中的搜索. 设经搜索,  $MB_{top}$  与  $MB_{bot}$  得到的运动向量依次记为  $mv_{top \rightarrow top}$ ,  $mv_{top \rightarrow bot}$  与  $mv_{bot \rightarrow top}$ ,  $mv_{bot \rightarrow bot}$ .

### 2.2 相关性分析

#### 2.2.1 $16 \times 16$ 块与 $16 \times 8$ (顶、底)块之间的运动相关性

设整个宏块  $MB$  在搜索区域  $SA$  中的搜索得到的运动向量记为  $mv_{MB \rightarrow SA}$ . 那么对一个宏块来讲,在完成 $16 \times 6$ 块与 $16 \times 8$ (顶、底)块运动估计的过程中,将产生五个运动向量,即  $mv_{top \rightarrow top}$ ,  $mv_{top \rightarrow bot}$ ,  $mv_{bot \rightarrow top}$ ,  $mv_{bot \rightarrow bot}$  与  $mv_{MB \rightarrow SA}$ . 这些运动向量之间是存在着相关性的,如果已知运动向量  $mv_{top \rightarrow top}$  的值,依据运动向量之间的相关性(线性相关与非线性相关),可预测出其它四个运动向量的值. 相关关系可统一用式

$$mv = (V_x + \Delta V_x, \lambda \times V_y + \Delta V_y) \quad (2)$$

描述. 上式中  $\Delta V_x, \Delta V_y$  为偏差,  $\lambda$  为常数, 当  $mv$  表示  $mv_{MB \rightarrow SA}$  时  $\lambda=2$ , 其它情况  $\lambda=1$ .

同样, 设  $MB_{top}, SA_{top}$  分别进行隔列抽样得到  $MB'_{top}, SA'_{top}, MB'_{top}$ , 在  $SA'_{top}$  中搜索产生的运动向量为  $mv'_{top \rightarrow top} = (V'_x, V'_y)$ , 则  $mv_{top \rightarrow bot}$  与  $mv_{top \rightarrow top}$  关系见下式

$$mv_{top \rightarrow top} = (V_x, V_y) = (2V'_x + \Delta\tilde{V}'_x, V'_y + \Delta\tilde{V}'_y), \quad (3)$$

其中  $\Delta\tilde{V}'_x, \Delta\tilde{V}'_y$  为非线性因素所产生的影响.

显然, 由(2)与(3)式可导出  $mv_{top \rightarrow top}, mv_{top \rightarrow bot}, mv_{bot \rightarrow top}, mv_{bot \rightarrow bot}, mv_{MB \rightarrow SA}$  与  $mv'_{top \rightarrow top}$  关系的统一表达式

$$mv = (2V'_x + \Delta V'_x, \lambda \times V'_y + \Delta V'_y), \quad (4)$$

其中  $\Delta V'_x = \Delta V_x + \Delta\tilde{V}'_x, \Delta V'_y = \Delta V_y + \lambda + \Delta\tilde{V}'_y$ .

由(4)式可知, 在已知运动向量  $mv'_{top \rightarrow top} = (V'_x, V'_y)$  的情况下, 宏块的运动向量  $mv_{top \rightarrow top}, mv_{top \rightarrow bot}, mv_{bot \rightarrow top}, mv_{bot \rightarrow bot}$  与  $mv_{MB \rightarrow SA}$  可用  $mv'_{top \rightarrow top}$  与一个偏差向量表示. 事实上偏差量  $\Delta V'_x, \Delta V'_y$ , 是比较小的. 因此可以依据  $mv'_{top \rightarrow top}$  确定搜索中心, 然后在此中心的一个小范围内搜索(FSA)便能得到  $\Delta V'_x, \Delta V'_y$ , 并由此可得到  $mv_{top \rightarrow top}, mv_{top \rightarrow bot}, mv_{bot \rightarrow top}, mv_{bot \rightarrow bot}$  与  $mv_{MB \rightarrow SA}$ .

### 2.2.2 16×16块与16×8(顶、底)块运动估计的计算相关性

不失一般性, 设进行运动估计的块尺寸为  $W \times H$ , 运动向量的范围是: 水平方向为  $[-K, +K]$ , 垂直方向为  $[-L, +L]$ . 运动估计时采用平均绝对值误差和(MAD)作为衡量块匹配程度的判据, 失配函数定义为

$$F(k, l) \Big|_{k \in [-K, +K], l \in [-L, +L]} = \frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H |G(i, j) - H(i+k, j+l)|, \quad (5)$$

这里  $G, H$  分别表示当前帧与参考帧图像,  $(k, l)$  代表候选块相对于当前待估计块的位移. 相应的运动向量可由下式决定

$$mv = (k, l) \Big|_{\min F(k, l)}. \quad (6)$$

对一个宏块来讲, 其  $W=16, H=16$ , 并假设  $K=L=2 \times SR0$ ; 它的两个场块为  $16 \times 8$ , 即  $W=16, H=8$ , 这时  $K$  也为  $2 \times SR0$ , 而  $L$  应为  $SR0$ . 对于某一宏块, 如果  $16 \times 16$  块与  $16 \times 8$ (顶、底)块运动估计都采用全搜索算法(FSA), 那么根据(1)与(5)式可知,  $16 \times 16$  块与  $16 \times 8$ (顶、底)块运动估计之间存在计算相关, 即其中一种运动估计的中间计算结果可被另一种运动估计直接利用.

设宏块进行  $16 \times 16$  块匹配时, 得到的  $MAD$  值记为  $MAD_{16 \times 16}$ , 进行  $16 \times 8$ (顶、底)块匹配得到的  $MAD$  值分别记为  $MAD_{top \rightarrow top}, MAD_{bot \rightarrow bot}, MAD_{top \rightarrow bot}$  与  $MAD_{bot \rightarrow top}$ . 当宏块进行一次  $16 \times 16$  块匹配时, 搜索区域中的候选块数据或顶场在前底场在后, 或底场在前顶场在后. 根据(5)式得出如下关系:

当候选块的数据顶场在前底场在后时

$$MAD_{16 \times 16} = MAD_{top \rightarrow top} + MAD_{bot \rightarrow bot}; \quad (7)$$

当候选块的数据底场在前顶场在后时

$$MAD_{16 \times 16} = MAD_{top \rightarrow bot} + MAD_{bot \rightarrow top}. \quad (8)$$

显见, 在完成一种运动估计过程中, 只要增加少许计算量就可以完成另一种运动估计. 因此, 利用  $16 \times 16$  块与  $16 \times 8$ (顶、底)块运动估计的计算相关性能降低近50%的计算

量.

### 2.3 算法提出

#### 1) 第一级.

i) 对当前宏块及其搜索区域的顶场分别进行子采样得到各自的子采样块,这里定义子采样是指对顶场图像数据进行隔列抽样;

ii) 当前宏块的子采样块在搜索区域的子采样块上,采用 *MAD* 全搜索策略(FSA)得到初步运动向量( $V'_x, V'_y$ ),设搜索区域的大小为 $(8+2 \times SR) \times (8+2 \times SR)$ .

#### 2) 第二级.

i) 小搜索区域确定,根据初步运动向量( $V'_x, V'_y$ ),可求出当前宏块的 $16 \times 16$ 块运动估计的搜索中心相对位置为 $(2V'_x, 2V'_y)$ ,设 $16 \times 16$ 块运动估计的搜索区域大小为 $(16+2 \times sr) \times (16+2 \times sr)$ ,*sr*取值较小,故称为小搜索区域;

ii) 在小搜索区域中采用 *MAD* 全搜索算法(FSA),并利用 $16 \times 16$ 块与 $16 \times 8$ (顶、底)块运动估计的计算相关性,可在一次搜索过程中同时得到 $16 \times 16$ 与 $16 \times 8$ 顶、底块的偏差运动向量,然后根据(4)式可以得到运动向量  $mv_{top \rightarrow top}$ ,  $mv_{top \rightarrow bot}$ ,  $mv_{bot \rightarrow top}$ ,  $mv_{bot \rightarrow bot}$ , 与  $mv_{MB \rightarrow SA}$ ;

iii) 根据  $MAD_{top \rightarrow top}$ ,  $MAD_{bot \rightarrow bot}$ ,  $MAD_{top \rightarrow bot}$  与  $MAD_{bot \rightarrow top}$  的值,进一步形成 $16 \times 8$ (顶、底)块运动估计的结果.

## 3 算法性能比较

本文算法的第一级与第二级都是采用 *MAD* 全搜索算法(FSA),而基于 *MAD* 的全搜索算法(FSA)的实现技术非常成熟,所以本文所提出的新快速两级块匹配算法更易于 VLSI 实现.限于篇幅,下面只给出算法的比较结果.参与对比的算法是,全搜索算法(FSA)与文献[3]提出的两级搜索算法(记为 TSBMA).

### 3.1 计算量比较

设 FSA, TSBMA 与本文算法的 $16 \times 16$ 块运动向量在水平与垂直方向的最大位移绝对值相等且为  $V_{max}$ .为了便于计算量的比较,这里分两种情况,即宏块只做 $16 \times 16$ 块运动估计和宏块做 $16 \times 16$ 块并同时做 $16 \times 8$ 块运动估计.计算量以运算次数为单位,即进行一次加(减)法、绝对值或比较运算都被认为是一次运算.

设  $V_{max} = 24$ ,即  $sr = 4, SR = 10$ .根据计算可得 FSA, TSBMA 与本文算法完成宏块的一次单向运动估计计算量,比较如下:1) 当宏块只做 $16 \times 16$ 块运动估计时,本文算法的计算量为 FSA 的 8.0%,为 TSBMA 的 12.2%;2) 当宏块同时做 $16 \times 16$ 块与 $16 \times 8$ 块运动估计时,本文算法的计算量仅为 FSA 的 4.0%;3) 对本文算法来讲,宏块同时做 $16 \times 16$ 块与 $16 \times 8$ 块运动估计和只做 $16 \times 16$ 块运动估计的计算量几乎相等.

### 3.2 MSE 与 PSNR 比较

计算机模拟实验是在标准测试序列 football 和 flower garden 的各自前 31 帧图像上进行.比较指标是采用 *P, B* 图像的运动估计均方误差 MSE 和解码图像的峰值信噪比 PSNR.设进行运动估计与编码时图像组 GOP 的长度  $N = 15$ ,相邻参考图像的距离  $M = 3, V_{max} = 24$ ,本文算法的  $SR = 10, sr = 4$ ,MPEG-2 编码器的输出比特速率为 6.4 Mbits/s.

根据实验结果可得如下结论:1) 宏块同时做 $16 \times 16$ 块与 $16 \times 8$ 块运动估计的效果明显优于只做 $16 \times 16$ 块运动估计的效果;2) 当宏块只做 $16 \times 16$ 块运动估计时,本文算法与TSBMA的性能相当;3) 宏块同时做 $16 \times 16$ 块与 $16 \times 8$ 块运动估计时,本文算法与FSA相比性能只有些许下降,综合football与flower garden的PSNR实验结果,可以得到 $\widetilde{PNSR}$ 平均只下降约0.17dB. 另外,主观实验也表明本文算法的解码图像质量优良.

### 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 Akiyama T, Aono H *et al.* MPEG-2 video codec using image compression DSP. *IEEE Transaction on Consumer Electronic*, 1994, **40**(3):466—472
- 3 Kim J S, Oark R H, LEE B U. A tow-stage fast block matching algorithm using integral projections. *Visual Communication Image Representation*, 1993, (4):336—348
- 4 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

**陈维强** 1968年出生,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、在职博士生,目前主要研究领域是运动图像压缩编码技术、系统控制.

**高 文** 1956年出生,1988年、1991年分别获哈尔滨工业大学计算机应用博士学位和日本东京大学电子学博士学位. 曾任日本东京大学客座研究员,美国CMU(1993年),MIT(1995年)客座教授. 现任中国科学院计算技术研究所所长、哈尔滨工业大学教授,博士生导师,兼任国家“八六三”计划信息领域智能计算机主题专家组组长等职. 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、虚拟现实技术及人工智能应用等.