

文章编号: 2095-4980(2017)04-0663-07

基于 AVS 高清编码器 IME 算法设计和实现

王佳薇, 冯光朗, 杨海钢

(中国科学院 电子学研究所, 北京 100190)

摘 要: 在一个典型的视频编码系统中, 其中运动估计(ME)的运算量大约占总运算量的 60%~90%, 尤其未来的数字电视将以高清和超高清标称, 意味着在保持一定的信噪比条件下更大的搜索范围。本文提出了一种基于以上算法优势特点的运动估计算法及其硬件实现——分层全搜索运动估计(HFSME), 该算法支持 AVS 标准的多参考帧技术(B,P 帧均采用 2 幅图像作为参考帧)、率失真优化(RDO)和 4 种宏块分块模式($16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8$), 并满足 AVS 高清实时编码器的需求。从性能上该算法及其硬件实现支持 AVS Baseline 级高清 1080P, 帧率达到 30 fps, 搜索面积达到 234×98 像素, 相较于同样性能下的全搜索算法, 峰值信噪比(PSNR)相差不大, 但运算量只有全搜索算法运算量的 1/4。同时, 相较于文献[12]中的结构, 本文设计的基于 HFSME 算法的 IME(整像素运动估计)结构设计, 在处理单元(PE)规模上是文献[12]结构的 2 倍, 但是搜索范围能力是其 4 倍, 并且吞吐率是其 10 倍, 具有最优的性能, 带来了良好的性价比。

关键词: AVS 高清编码器; 整像素运动估计(IME); 处理单元(PE)

中图分类号: TN919.81

文献标志码: A

doi: 10.11805/TKYDA201704.0663

A new IME algorithm and architecture based on AVS HD encoder

WANG Jiawei, FENG Guanglang, YANG Haigang

(Institute of Electronics, Chinese Academy of Science, Beijing 100190, China)

Abstract: In a classic video encoder, Motion Estimation(ME) comes as a high computational complexity, about 60%–90% in a whole system. An efficient algorithm named as Hierarchical Full Search ME(HFSME) is proposed and implemented on the circuit. Based on AVS HD encoder, it supports the AVS variable block size match in which both B and P frames adopt two pictures as the reference frames, Rate Distortion Optimization(RDO) and four kinds of macro-block partition modes concurrently. It totally produces 18 motion vectors for each macro-block. Experimental results show that the architecture can achieve 30 fps under the conditions of two reference frames, 234×98 search range and $1\ 920 \times 1\ 080$ picture size. The algorithm & architecture proposed in this paper can meet the requirements of AVS HD Real-Time Encoder. The computational complexity of HFSME proposed in the paper is only a quarter of that of full search. At the same time, HFSME can provide the highest computational capability which is about 10 times higher than that of reference [12].

Keywords: AVS HD encoder; Integer Motion Estimation(IME); Processing Element(PE)

数字音视频系统与产品需要以若干技术标准作为支撑, 其中信源编码标准和信道传输标准为最重要的两大技术标准。信源编码技术对数字音视频系统的整体效率影响最大, 由 ISO/IEC 推出的 MPEG 系列压缩标准, ITU-T 推出的 H.26X 系列压缩标准以及 ISO/IEC 和 ITU-T 联合成立的联合视频编码组(Joint Video Team, JVT), 致力于 H.26L 的推出^[1]。基于以上国际标准的知识产权授权问题, 确定了拥有自主知识产权的音视频标准——AVS 标准^[2]的必要性, 其在技术和性能上处于国际先进水平: 编码效率比 MPEG-2 国际标准高 2~3 倍, 与 MPEG-4 和 H.264 标准编码效率相当或更高, 并在一定程度上兼容。一个视频编码器的优势体现在两个方面: a) 视频编码器具有足够的压缩比; b) 视频信号压缩编码后, 经解码重构后的视频质量。在一个典型的视频编码系统中, 其中运动估计(ME)的运算量大约占到总运算量的 60%~90%^[3], 尤其未来的数字电视将以高清和超高清标称, 意

收稿日期: 2015-12-14; 修回日期: 2016-02-02

λ (拉格朗日)^[11]是 RDO 输出参数作为权重因子, 则权重值:

$$\sigma = \lambda \cdot (R(mv_x - pmv_x) + R(mv_y - pmv_y)) \quad (3)$$

式中: R 是通过运动矢量预测值, 即 pmv ^[2]和搜索点 mv 的距离, 经查表得到熵编码比特长度。

2.2 HFSME 算法中下采样层

本文研究的算法为分层全搜索, 分为 2 层: 下采样层和原始层分别进行全搜索, 下采样层是将待编码的当前和参考帧进行下采样为 1/4 原始图像大小的缩小帧, 16×16 宏块经下采样后变成 8×8 宏块, 在下采样层中的固定搜索窗口内进行全搜索, 将搜索结果保存后作为原始层的搜索起始点。

如图 2 所示 HFSME 算法分为 2 层, Level 0 层为原始层, Level 1 层为 HFSME 算法的下采样层, Level 1 层下采样方式不同于文献[7]中推荐的下采样方式:

$$I^{(l+1)}(i, j) = \frac{1}{2} \sum_{n=0}^1 I^{(l)}(i \times 2, j \times 2 + n) \quad (4)$$

i 和 j 代表一帧中像素的位置, l 代表 Level, 公式(4)表示 Level 1 层 (i, j) 的像素值等于 Level 0 层的 (i, j) 和 $(i, j+1)$ 的均值。本文推荐的 HFSME 算法的下采样滤波器按照公式(5)的方式进行下采样:

$$I^{(l+1)}(i, j) = \frac{1}{4} \sum_{m=0}^1 \sum_{n=0}^1 I^{(l)}(i \times 2 + m \times 2, j \times 2 + n \times 2) \quad (5)$$

经过下采样后的全搜索具有以下优势:

1) 搜索面积能够满足高清编码器的需求, 搜索窗达到 $P_h = [-2 \times 56, 2 \times 52]$, $P_v = [-2 \times 20, 2 \times 20]$, 经过 RM52C 验证, 90%以上的高清码流运动矢量在此搜索窗范围内;

2) 搜索运算量是 FS 算法的 1/4;

3) 在下采样层只需要进行 8×8 宏块的代价函数运算, 而不是 16×16 的运算单元, 并且不需要进行 8×4, 4×8 和 4×4 的代价函数运算, 结构上进一步简化。

2.3 数据加载

对于硬件实现上, 数据带宽是硬件结构的制约瓶颈。在 AVS 高清编码器中, 参考帧和当前帧的数据量大, 通常放在外部存储器中存储。如何提高带宽利用率, 降低运动估计对存储器的数据访问量, 成为设计编码器的关键。按照高清分辨率(1 920×1 080)@30 Hz, 系统工作于 200 MHz 的设计需求, 硬件复杂度分析如下:

$$\frac{1}{frequency} \cdot cycles = \frac{1}{\frac{width \cdot high}{MB} \cdot frame_rate} \quad (6)$$

将设计参数代入公式(6), 计算得出每个宏块处理周期为 823 个周期(cycles)。

数据加载量, 按照 BP 双参考帧, 下采样层的搜索窗 $P_h = [-(2 \times 56), (2 \times 52)]$, $P_v = [-2 \times 20, 2 \times 20]$ 计算:

$$data_bandwidth = frame_num \times \frac{\frac{1}{2} R_h}{\frac{1}{2} MB} \times \frac{1}{2} P_v \quad (7)$$

数据带宽代入公式(7), 计算得出每个宏块搜索窗的数据访问量为 1 080 个数据, 不能满足公式(6)的设计需求。因此提高带宽利用率, 降低访存量成为设计的关键。基于下采样层每个宏块的搜索窗中心为(0,0)点, 那么 1 帧图像中 2 个相邻宏块的参考像素范围只差 1 个宏块宽度, 如图 3 所示, 大部分参考像素可以重用, 基于此前提, 设计采用最大化重用已载入数据的技术, 降低系统访存带宽, 同时采用流水线方式。当处理宏块 MB_n 时, 即搜索范围是“正在使用”部分, 加载相邻宏块 MB_{n+1} 搜索窗所需要的额外的参考像素。

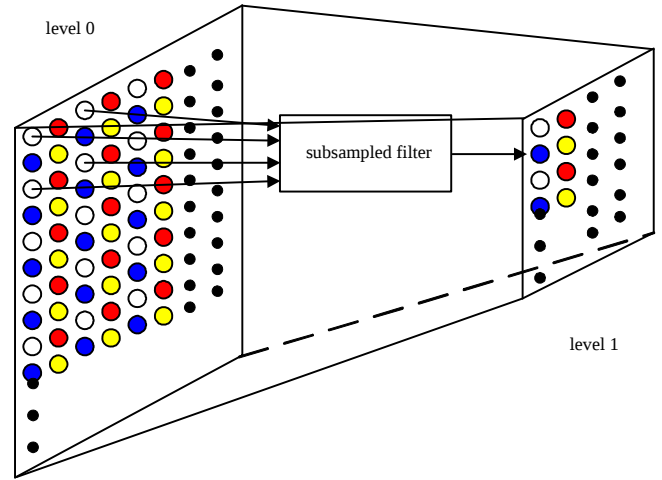


Fig.2 Subsampled method based hierarchy architecture in HFSME
图 2 HFSME 算法中分层结构的下采样方式

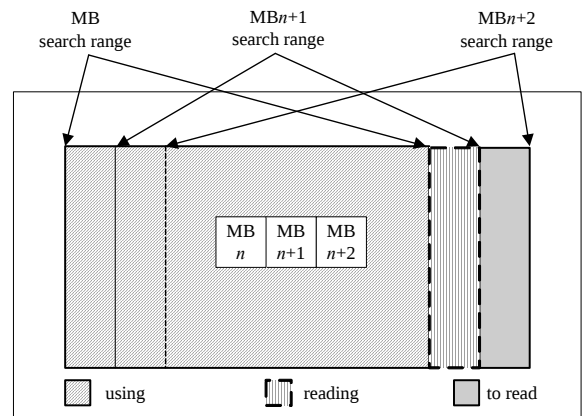


Fig.3 Diagram of reference data loading
图 3 参考像素数据加载示意图

按照图 3 的数据加载方式, 数据加载量由公式(8)代替公式(7):

$$data_bandwidth = frame_num \times \frac{1}{2} P_v \quad (8)$$

计算得出数据带宽为 80 个数据, 完全满足公式(6)的设计需求。

2.4 试验结果

为了评估本文设计 ME 算法的性能, 采用高清编码器 C 模型, 该模型具有:

- 1) 按照硬件结构进行搭建, 用来进行 RTL 代码调试和验证;
- 2) 可以在高清编码器 C 模型上实现不同算法, 比较不同算法间的性能;
- 3) 高清编码器 C 模型还有另外一个重要作用, 即为各个模块产生测试激励需要的测试向量。

采用 C 模型来评估算法的性能, 在同样的搜索面积和同样的量化值(QP)下, 通过信噪比(PSNR)和码率(bit rate)进行不同算法的性能比较。信噪比越高越好, 码率越低越好。表 1 所示为在搜索面积为 234×98 , QP 为 32 条件下的 PSNR 和码率。从表 1 可知, 分层全搜索在性能上与全搜索算法相差不多, 但计算量上却只是全搜索的 $1/4$ 。基于 HFSME 算法实现的 IME 设计, 在 FPGA 实现后经仿真结果验证, 与 C 模型结果一致, 再经板级调试, 视频压缩后的编码经解码器解码后正常播放高清图像。

表 1 算法性能比较

sequence	FS		HFSME	
	PSNR/dB	bit rate	PSNR/dB	bit rate
Foreman	33.51	229.01	33.45	231.41
Coast guard	31.48	514.31	31.46	516.65
Flower	30.99	946.19	30.97	949.70
Mobile	30.14	892.14	30.12	891.71

3 硬件结构

基于本文提出的 HFSME 算法, AVS 高清编码器 IME 设计采用流水线架构, 分为 4 级流水结构: Level 1 层加载数据、Level 1 层搜索处理、Level 0 层加载数据、Level 0 层搜索处理, 如图 4 所示。

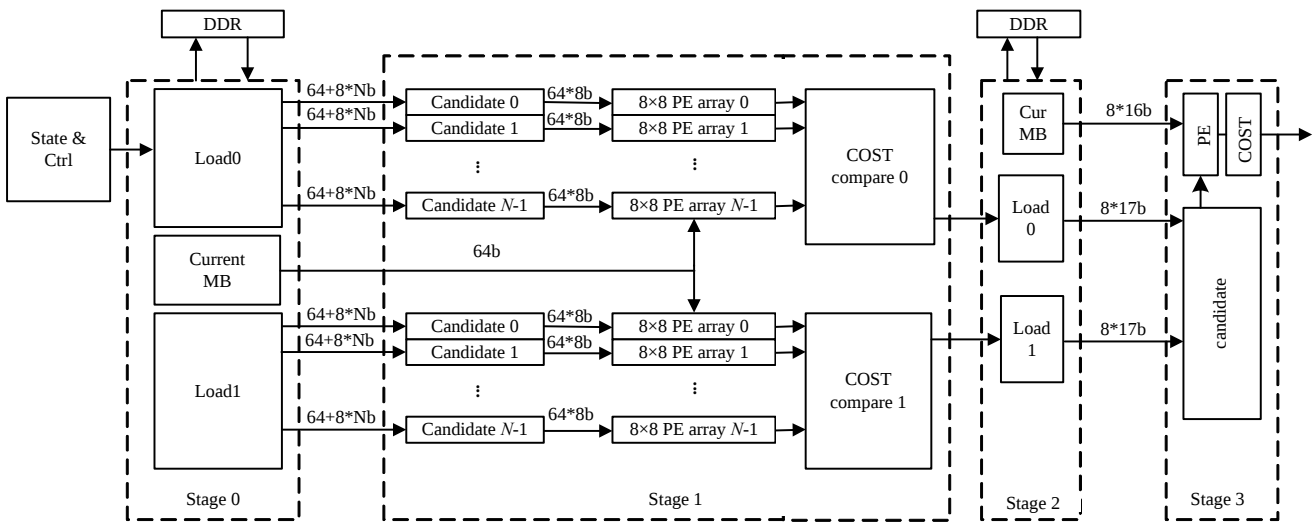


Fig.4 Top architecture of IME proposed in this paper

图 4 AVS 高清编码器 IME 流水线架构

图 4 中, 下采样后的当前帧和 2 帧参考帧以及原始图像的当前帧和 2 帧参考帧动态地存储在片外存储器中, 其中:

Stage 0: 将下采样层(L_1)2 个参考帧中 $P_b=[-56,52], P_v=[-20,20]$ 的搜索窗数据(Load 0 和 Load 1)和当前宏块数据按照 2.3 方式加载到片内存储空间 RAM 中;

Stage 1: 为 L_1 层的 2 个参考窗搜索处理, 经过 8×8 宏块的代价函数运算比较后, 将最优位置转换为原始层(L_0)的运动矢量(MV)输出, 作为 L_0 层的搜索起点;

Stage 2: 以 L_1 层输出 MV 为中心, 将 L_0 层 2 个参考帧中 9×9 的搜索窗数据(Load 0 和 Load 1)和当前宏块数据加载到片内存储空间;

Stage 3: 为 L_0 层的 2 个参考窗搜索处理, 基于 AVS 支持 $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8$ 共 4 种可变块大小模式。

3.1 candidate 并行处理结构

如图 4 所示, candidate 的作用是从加载参考像素的 RAM 中读取像素, 经过整理后传递给 PE(处理单元)单元。基于本文推荐算法中提到的 L_1 层的搜索范围($P_h=[-(2 \times 56), (2 \times 52)], P_v=[-2 \times 20, 2 \times 20]$)计算, 共 108×40 个搜索点, 按照每个流水级控制在公式(6)提出的 823 个周期是不能达到指标要求的。因此, 如图 4 所示的 Stage 1 级的 candidate 结构采用并行处理方式, 满足公式(9)的指标要求:

$$\frac{1}{2} \frac{P_h}{N} \frac{1}{2} P_v \leq 823 \text{ cycles} \quad (9)$$

将本文推荐的搜索范围和并行程度 N 代入公式(9), 得到本级流水控制在 723cycles, 满足基于 AVS 的高清编码器的实时性需求。并且基于公式(9), 利用结构上的可重用性, 提供了灵活的搜索范围和并行程度。

3.2 搜索遍历结构

每个搜索点都需要从相应的 RAM 读取宏块大小的像素数据, 这增加了带宽负荷。考虑到每个搜索点的数据重用性, 本文设计了一种基于移位寄存器硬件结构遍历方式, 如图 5 所示。基于图 5 和表 2 描述的 candidate 遍历方式:

- 1) 当 $up_down=a$ 时, $shift_0$ 从相应 RAM 中读取位宽为 $(64+N \times 8)$ bits 的数据, 同时向下依次更新 $shift_1 \sim shift_7$ 的数据, 并将 $shift_0 \sim shift_7$ 中低 64 bit 数据输出。
- 2) 当 $up_down=b$ 时, $shift_0 \sim shift_7$ 的数据进行右移 $(8 \times N)$ bit, 并将 $shift_0 \sim shift_7$ 中低 64 bit 数据输出。
- 3) 当 $up_down=c$ 时, $shift_7$ 从相应 RAM 中读取位宽为 $(64+N \times 8)$ bit 的数据, 同时向上依次更新 $shift_6 \sim shift_0$ 的数据, 并将 $shift_0 \sim shift_7$ 中低 64 bit 数据输出。

表 2 Candidate 数据流
Table 2 Candidate data flow

up/down	Shift_0	shift_1	shift_2	shift_3	shift_4	shift_5	shift_6	shift_7
a	$r(0,0), r(0,13)$	—	—	—	—	—	—	—
a	$r(1,0), r(1,13)$	$r(0,0), r(0,13)$	—	—	—	—	—	—
a	...	...	...	...	...	...	...	...
a	$r(7,0), r(7,13)$	$r(6,0), r(6,13)$	$r(5,0), r(5,13)$	$r(4,0), r(4,13)$	$r(3,0), r(3,13)$	$r(2,0), r(2,13)$	$r(1,0), r(1,13)$	$r(0,0), r(0,13)$
:	:	:	:	:	:	:	:	:
a	$r(1/2P_v+7,0),$ $r(1/2P_v+7,13)$	$r(1/2P_v+6,0),$ $r(1/2P_v+6,13)$	$r(1/2P_v+5,0),$ $r(1/2P_v+5,13)$	$r(1/2P_v+4,0),$ $r(1/2P_v+4,13)$	$r(1/2P_v+3,0),$ $r(1/2P_v+3,13)$	$r(1/2P_v+2,0),$ $r(1/2P_v+2,13)$	$r(1/2P_v+1,0),$ $r(1/2P_v+1,13)$	$r(1/2P_v+0,0),$ $r(1/2P_v+0,13)$
b	$r(1/2P_v+7,6),$ $r(1/2P_v+7,13)$	$r(1/2P_v+6,6),$ $r(1/2P_v+6,13)$	$r(1/2P_v+5,6),$ $r(1/2P_v+5,13)$	$r(1/2P_v+4,6),$ $r(1/2P_v+4,13)$	$r(1/2P_v+3,6),$ $r(1/2P_v+3,13)$	$r(1/2P_v+2,6),$ $r(1/2P_v+2,13)$	$r(1/2P_v+1,6),$ $r(1/2P_v+1,13)$	$r(1/2P_v+0,6),$ $r(1/2P_v+0,13)$
c	$r(1/2P_v+6,6),$ $r(1/2P_v+6,19)$	$r(1/2P_v+5,6),$ $r(1/2P_v+5,19)$	$r(1/2P_v+4,6),$ $r(1/2P_v+4,19)$	$r(1/2P_v+3,6),$ $r(1/2P_v+3,19)$	$r(1/2P_v+2,6),$ $r(1/2P_v+2,19)$	$r(1/2P_v+1,6),$ $r(1/2P_v+1,19)$	$r(1/2P_v+0,6),$ $r(1/2P_v+0,19)$	$r(1/2P_v-1,6),$ $r(1/2P_v-1,19)$
:	:	:	:	:	:	:	:	:

3.3 16×16 PE 结构

如图 6 所示, L_0 层的 16×16 的 PE 由 4 个 8×8 PE 组成, 分别为 8×8 PE0, 8×8 PE1, 8×8 PE2, 8×8 PE3, 每个 8×8 PE 计算得到最小块 8×8 模式的代价(σ)结果, 并能同时计算出 $8 \times 8, 8 \times 16, 16 \times 8, 16 \times 16$ 四种模式共 9 个 σ 值, 支持双参考帧共 18 个 σ 值。

$$\begin{aligned} \sigma(8 \times 8 \text{PE}0) &= \sigma_{8 \times 8_0}; \\ \sigma(8 \times 8 \text{PE}1) &= \sigma_{8 \times 8_1}; \end{aligned}$$

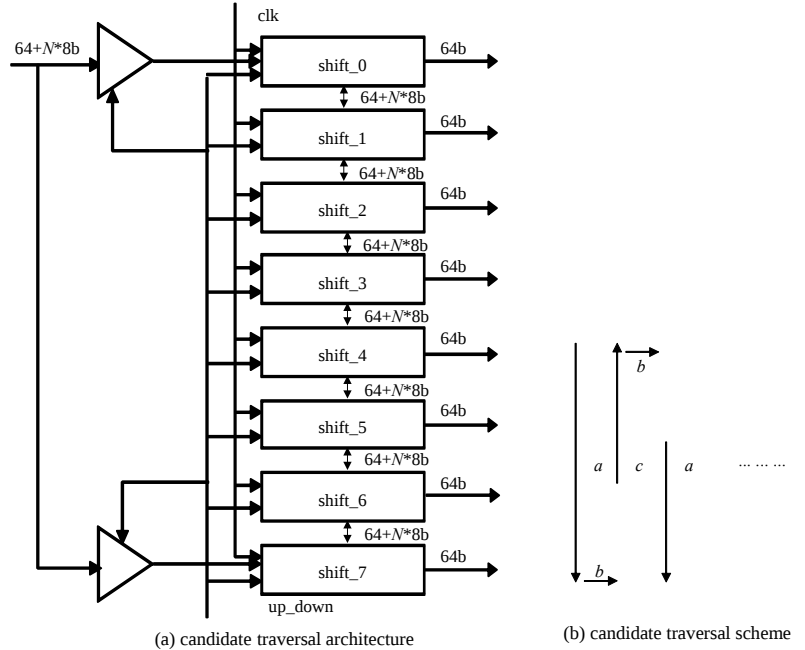


Fig.5 Candidate traversal in search
图 5 Candidate 搜索遍历

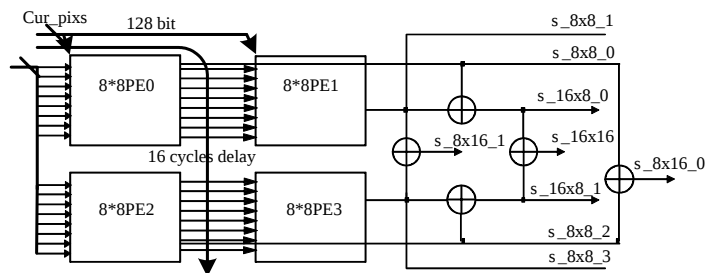


Fig.6 PE architecture of 16×16 method
图 6 16 × 16 PE 结构

```
σ(8×8PE2)=σ_8×8_2;  
σ(8×8PE3)=σ_8×8_3;  
σ(8×8PE0)+σ(8×8PE1)=σ_8×16_0;  
σ(8×8PE2)+σ(8×8PE3)=σ_8×16_1;  
σ(8×8PE0)+σ(8×8PE2)=σ_16×8_0;  
σ(8×8PE1)+σ(8×8PE3)=σ_16×8_1;  
σ(8×8PE0)+σ(8×8PE1)+σ(8×8PE2)+σ(8×8PE3)=σ_16×16。
```

每种模式下的最小 COST(min_COST)，以第 1 个搜索点对应的 COST 为初值：

```
if(COST< min_COST)  
begin  
min_COST<= COST;  
min_mv_x<= mv_x;  
min_mv_y<= mv_y;  
end
```

4 设计结果

HFSME 设计的 IME 支持基于 AVS 的高清编码器，输入为所需要的 2 帧参考图像和当前图像，运动矢量以数据包形式进行打包，通过流水级间 FIFO 进行输出。

同时，基于本设计的高清编码器 C 模型验证了 HFSME 算法 IME 设计的功能和性能。ISE 下使用 V4LX160 型号芯片进行布局布线后，资源占用和时钟频率结果如表 3 所示。

性能上 HFSME 设计的 IME 及其硬件实现支持 AVS Baseline 级高清 1 080P，帧率达到 30 fps，搜索面积达到 234×98 像素，工作频率达到 220 MHz，完全满足设计需求。经过系统联调，结果显示能够正确编码，并且码率低，信噪比高。

表 3 布局布线后资源占用

Table3 Resource consumption of place & route

module	LUT	register	DSP	BRAM	Freq
IME	98 601(72%)	65 138(48%)	7(7%)	86	220 MHz

表 4 结构对比

Table4 Architecture comparison

architecture	PE/(Nums)	search range	frequency/MHz	resource/(gate count)	throughput/(pixel/s)
in [12]	512	65×65	250	400 k	396 981 000
HFSME proposed in this paper	1 024	234×98	220	980 k	5 572 476 000

文献[12]中进行了各种 VBSME 实现结构的比较，包括 PE 规模、搜索范围能力、频率、资源和吞吐率对比。从表 4 看到，对比文献[12]中的结构，本文设计的基于 HFSME 算法的 IME 结构设计，在 PE 规模上是文献[12]结构的 2 倍，但搜索范围能力是其 4 倍，吞吐率是其 10 倍，具有最优的性能，带来良好的性价比。

5 结论

基于 AVS 的高清编码器实际工作在 200 MHz 时钟频率，本设计在时钟频率上的冗余保证了基于 AVS 的高清编码器的工作稳定性。从数据流角度，本文提出的 HFSME 算法和结构设计保证了每个 MB 在 823 cycles 内吞吐率。因此，本文提出的基于 HFSME 算法实现的 IME 设计，支持 AVS 标准的双参考帧技术，率失真优化 RDO 和可变块大小模式(16×16,16×8,8×16,8×8)。从性能上看，该算法及其硬件实现支持 AVS Baseline 级高清 1 080P，帧率达到 30 fps，搜索面积达到 234×98 像素，相较于同样性能下的全搜索算法，峰值信噪比相差不大，但运算量大幅下降，只有相同搜索面积下全搜索算法运算量的 1/4，满足 AVS 高清实时编码器的需求。

参考文献：

[1] 尹显东,李在铭,姚军,等. 图像压缩标准研究的发展与前景[J]. 太赫兹科学与电子信息学报, 2003,1(4):326–331. (YIN Xiandong,LI Zaiming,YAO Jun,et al. A survey on trends of image compression coding standards[J]. Journal of Terahertz Science and Electronic Information Technology, 2003,1(4):326–331.)

[2] Audio Video coding Standard Workgroup of China(AVS). Advanced Audio Video Coding Standard Part 2: Video: Doc. AVS-N1063[S]. AVS, 2005.

[3] 曹文锋,张颖,张兆扬,等. 一种适用于 H.264 的整像素运动估计算法[J]. 上海大学学报(自然科学版), 2004,10(4): 331–336. (CAO Wenfeng,ZHANG Ying,ZHANG Zhaoyang,et al. An integer pixel motion estimation algorithm applicable to H.264[J]. Journal of Shanghai University(Natural Science), 2004,10(4):331–336.)

- [4] 毕厚杰. 新一代视频压缩编码标准——H.264/ACV[M]. 北京:人民邮电出版社, 2009:31-32. (BI Houjie. A new generation of video compression coding standard—H.264/ACV[M]. Beijing:Posts & Telecom Press, 2009:31-32.)
- [5] WARRINGTON Stephen, CHAN Wai-Yip, SUDHARSANAN Subramania. Scalable high-throughput architecture for H.264/AVC variable block size motion estimation[C]// 2006 IEEE International Symposium on Circuits and Systems. Island of Kos, Greece:IEEE, 2006:3830-3833.
- [6] YIN Haibing, JIA Huizhu, QI Honggang, et al. A hardware-efficient multi-resolution block matching algorithm and its VLSI architecture for high definition MPEG-like video encoders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010,20(9):242-249.
- [7] SONG Byung Cheol, CHUN Kang-Wook. Multi-resolution block matching algorithm and its VLSI architecture for fast motion estimation in an MPEG-2 video encoder[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004,14(9):1119-1137.
- [8] CHEN Ching-Yeh, CHIEN Shao-Yi, HUANG Yu-Wen, et al. Analysis and architecture design of variable block-size motion estimation for H.264/AVC[J]. IEEE Transactions on Circuits and Systems, 2006,53(2):578-593.
- [9] ZHANG L, XIE D, WU D. Improved FFSBM algorithm and its VLSI architecture for AVS video standard[J]. Journal of Computer Science and Technology, 2006,21(3):378-382.
- [10] YE Xin, DING Dandan, YU Lu. A hardware-oriented IME algorithm and its implementation for HEVC[C]// 2014 IEEE Visual Communications and Image Processing Conference. Valletta, Malta:IEEE, 2014:205-208.
- [11] WIEGAND T, GIROD B. Lagrangian multiplier selection in hybrid video coder control[C]// 2001 International Conference on Image Processing. Thessaloniki, Greece:IEEE, 2001:542-545.
- [12] DENG Lei, XIE Xiaodong, GAO Wen. A real-time full architecture for AVS motion estimation[J]. IEEE Transactions on Consumer Electronics, 2007,53(4):1744-1751.

作者简介:



王佳薇(1980-), 女, 天津市人, 硕士, 工程师, 主要研究方向为基于 AVS 高清视频编码压缩、数字信号处理、大规模集成电路设计.
email:jwwang@mail.ie.ac.cn.

冯光朗(1979-), 男, 成都市人, 硕士, 工程师, 主要研究方向为通信技术、大规模集成电路设计.

杨海钢(1960-), 男, 上海市人, 博士, 研究员, 主要研究方向为可编程逻辑芯片技术、传感器和数模混合信号 SOC 设计、大规模集成电路自动化设计技术.

(上接第 655 页)

作者简介:



苏敏(1984-), 女, 湖北省仙桃市人, 在读博士研究生, 主要研究方向为计算电磁学、电磁兼容.email:sumin102@163.com.

刘培国(1969-), 男, 山东省章丘市人, 教授, 博士生导师, 主要研究方向为电磁防护、电磁兼容、计算电磁、生物电磁等.