

SVC 空域增强层快速运动估计算法

封 颖, 吴成柯

(西安电子科技大学 综合业务网理论及关键技术国家重点实验室, 陕西 西安 710071)

摘要: 基于分级编码空域增强层运动估计的特点, 提出了一种增强层快速运动估计算法. 该算法利用基本层信息产生增强层无残差预测运动估计的整像素 SAD 门限, 由该门限对是否进行基于残差预测的运动估计进行终止判决, 由此使大量的编码宏块避免了不必要的基于残差预测的运动估计. 并且对基于残差预测的运动估计, 直接在层间预测模式下, 以层间预测矢量为预测值进行运动估计, 进一步加快了编码速度. 实验结果表明, 该算法在保持信噪比几乎不变的同时, 使增强层的运动估计时间比原联合编码模型减少了 35%, 从而降低了编码器的运算复杂度.

关键词: 分级编码; 层间预测; 运动估计

中图分类号: TP919.8 **文献标识码:** A **文章编号:** 1001-2400(2007)05-0697-05

Fast motion estimation algorithm in spatial enhancement layer based on SVC

FENG Ying, WU Cheng-ke

(State Key Lab. of Integrated Service Networks, Xidian Univ., Xi'an 710071, China)

Abstract: Based on the analysis of the characteristics of motion estimation in the spatial enhancement layer, a new algorithm is proposed. The algorithm estimates the prediction threshold of pixel SAD with the information from the base layer. It makes a decision of whether the motion estimation with the residual-prediction is terminated, so a lot of macroblock can avoid the undesirable motion estimation with the residual-prediction. And it makes use of the upsampled motion vectors to do the motion estimation with the residual-prediction if it is unavoidable. Simulation results indicate that the proposed algorithm can decrease the coding time of the motion estimation in the enhancement layer by 35%, and the computational complexity of encoder is significantly reduced.

Key Words: scalable video coding; inter-layer prediction; motion estimation

随着网络和终端设备的发展, 视频应用对视频压缩技术不断提出新的要求. 分级编码(SVC)^[1]就是为了满足不同用户终端需求的视频编码标准, 它通过对一个视频序列采用时空域相结合的编码技术, 产生一个可以满足不同用户需求的码流, 各个用户都可以在自身解码条件下得到最佳的解码图像.

目前 SVC 标准是通过通过对 H. 264/AVC^[2]进行扩展实现的, 它能够提供完整的时域、空域和编码质量分级. SVC 利用层次 B 帧来实现时域的可分级, 解码不同层次的 B 帧可得到不同时域分辨率的视频重建图像. 对于空域的可分级性, SVC 采用对输入的视频序列进行空域下采样, 将其分解为不同分辨率的空域层序列, 对每个空域层采用独立的运动估计和补偿. SVC 增强层编码支持自适应的层间预测技术, 编码器可以选择是否进行层间运动信息和残差预测, 以此达到更高的编码效率, 但是这使得增强层的编码复杂度显著地增加, 所以需要增强层的运动估计引入优化算法, 提高其编码效率. 对于运动估计, 已经存在一些算法来降低运动估计的复杂度. 文[3]通过设定绝对残差和(SAD)阈值来减少运动估计的运算量, 文[4]根据相邻块动态调整搜索范围, 并利用 SAD 和运动矢量率失真来提前完成多参考帧的运动估计. 针对 H. 264/AVC, 已被

收稿日期: 2006-12-28

基金项目: 国家自然科学基金重点项目(60532060, 60372043)

作者简介: 封 颖(1981-), 男, 西安电子科技大学博士研究生.

联合视频编码组(JVT)采纳的算法 UMHexagonS^[5,6]效率最高,该算法通过设定特定的搜索形状、构造多个预搜索预测运动矢量和全零块判决等方法使得运动估计消耗的时间大大减少. SVC 的参考软件 JSVM^[7]中采用与之相似的快速运动估计算法. 该算法在对增强层宏块进行编码时,对所有的编码模式先进行残差预测模式下的运动估计,然后再进行无残差预测下的运动估计. 且每个编码模式的运动估计可能还需要进行两次,首先是采用基于层内预测运动矢量的编码,当参考帧与层间预测矢量的参考帧相同时还要采用基于层间预测运动矢量的编码. 由于该算法没有考虑到分级编码的特点,对所有可能采用的层间预测技术进行全遍历,使得运算量很大. 针对上述问题,笔者提出了一种改进的快速运动估计算法. 该算法在空域增强层运动估计时,根据无残差预测运动估计整像素点的 SAD 对层间残差预测技术进行提前判决,快速而准确地得到增强层宏块的最佳编码模式和运动矢量.

1 SVC 的层间预测技术

SVC 在增强层编码时引入了层间预测技术,即根据编码视频序列分辨率的比例,将基本层的运动矢量和残差数据进行上采样. 在对运动矢量和残差数据编码时,将这些值作为预测值,以此来达到降低增强层码率的目的.

图 1 是一个典型的层间预测示意图^[8],空域增强层和基本层的比例为 2 : 1. A 为基本层的一个宏块,则经过上采样后与增强层 B 中的 4 个宏块相对应. B 中每个宏块的编码模式在基本层中都有一个对应的 8×8 块子模式.

对于运动矢量的上采样,根据不同的层间分辨率比例,SVC 设定了不同的算法. 如果层间分辨率比为 2 : 1,设基本层的运动矢量为 \mathbf{B}_{MV} ,相应的参考帧为 BRefX(X 为 0 或 1),则通过下式可以得到增强层对应宏块 MV 的一个预测值 \mathbf{P}_{MVP} 为 2 倍的 \mathbf{B}_{MV} .

如果增强层宏块在运动估计时得到的运动矢量和层间预测值具有相同参考帧,并且两者的数值相同或者相差很小,该宏块的运动信息编码就采用基本层预测模式,可以统称为基本层运动(BLMM)模式^[9].

如果基本层的重构残差不全为零并且和增强层的残差数据比较相近,可以将基本层残差数据的上采样作为增强层的编码预测值,可以称之为基本层参差模式(BLRM). SVC 支持多种残差数据上采样,在 2 : 1 时采用 6 阶滤波器,该滤波器也应用于半像素插值. 可将同时满足 BLMM 和 BLRM 的编码宏块称为基本层预测模式宏块(BLMRM).

为了分析增强层宏块运动估计的特点,利用 JSVM6. 8. 2 对空域增强层中采用不同层间预测的宏块进行统计. 实验选取具有中等运动特性的 Paris 序列,设置了两个空域层,在各层编码中图像组(GOP)为 16,时域分 4 层. 空域基本层序列为 QCIF 格式,增强层为 CIF 格式,其他为 JSVM 默认设置. 提取编码序列的一个 GOP 进行分析,实验结果如图 2 所示. 横坐标为编码图像的图像编号,纵坐标为各编码模式所占的比. 为了便于分析,图 2 只统计了基于层间预测技术的各宏块编码模式的比例.

根据时间域分解,第 8 帧属于时域分辨率的第 2 层,第 4 和 12 帧属于第 3 层,其余为第 4 层. 由图 2 可知,BLMM 宏块随着编码帧所在时域层的增加而减少,这是因为随着时域层的增加,参考帧的距离越近,编码宏块基于同层预测运动估计得到的跳跃模式(SKIP)就越多,所以采用 BLMM 的宏块很少. 而除了在第 8 帧处,BLRM 和 BLMRM 曲线基本重合并且在 0 点附近,

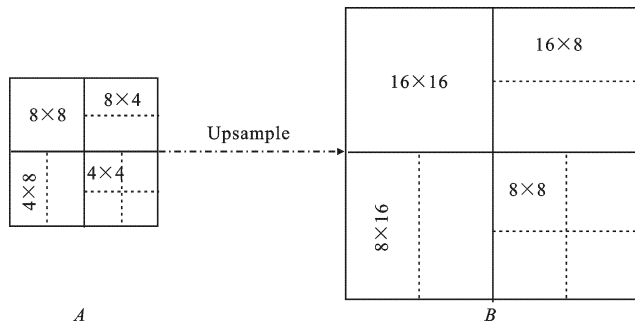


图 1 层间预测示意图

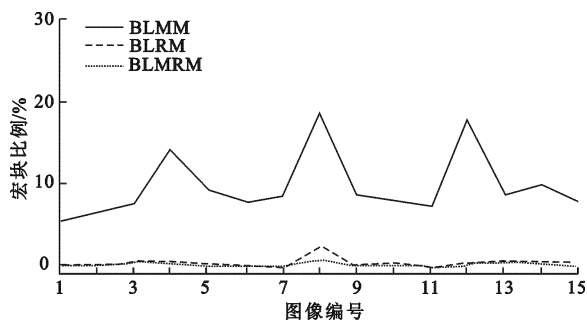


图 2 不同层间预测的概率曲线

说明采用基于残差预测 BLRM 的宏块很少并且其运动信息大多数采用 BLMM 模式,由此可知大量的宏块无需作基于残差预测的运动估计. 通过进一步实验,可发现第 8 帧中采用 BLMRM 的宏块模式,除了层间预测模式外有 90% 以上都是 8×8 的模式,这是因为层间预测模式一般都大于 8×8 ,而参考帧距离较远,采用小的块分割模式可以取得更好的编码效果.

2 空域增强层的快速运动估计算法

2.1 基于 MPEG-4 下采样滤波的 SAD 预测

由于增强层中采用残差预测 BLRM 模式的宏块很少,大量的宏块无需进行基于残差预测的运动估计. 为了减少编码时间,引入了增强层 SAD 预测提前终止策略,来减少层次 B 帧运动估计时对层间残差预测的遍历. 基本层编码序列由增强层序列下采样获取,如果下采样比例为 $2 : 1$,SVC 采用 MPEG-4 下采样滤波器,取偶数点生成 $2 : 1$ 的下采样序列. 该 13 阶滤波系数为 $D = \{2, 0, -4, -3, 5, 19, 26, 19, 5, -3, -4, 0, 2\} / 64$,先进行行方向滤波,然后是列方向滤波. 为了说明下采样滤波的特点,下面分析以行方向来说明,对于列方向与之相似. 图 3 是下采样滤波示意图.

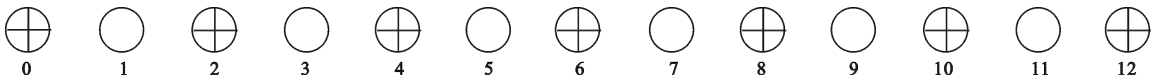


图 3 下采样示意图

设当前滤波点是 x_{2k} , 滤波系数 $d_i \in D$, 则滤波后的值为 $x'_{2k} = \sum_{i=-6}^{+6} (x_{2k-i} \cdot d_{6-i})$, x_{2k} 点对应的滤波系数是 $d_6 = 26/64$, 其余各个点对应的滤波系数与 d_6 的位置关系和该点与 x_{2k} 点相对位置一致. 设有 3 个点序列 X_1, X_2 和 X_3 , 经过行滤波输后为 X'_1, X'_2 和 X'_3 , 由 X'_2 和 X'_3 生成 X'_1 的预测值, 可得到 X'_1 与预测值的绝对残差和 SAD 为 $\sum_{n=N_1}^{N_2} |X'_{1,n} - (X'_{2,n} + X'_{3,n}) / 2|$. 根据下采样滤波的特点, 得到如下不等式

$$\sum_{n=N_1}^{N_2} \left| X'_{1,n} - \frac{1}{2}(X'_{2,n} + X'_{3,n}) \right| = \sum_{k=N_1}^{N_2} \left| \sum_{i=-6}^{+6} x_{1,2k-i} \cdot d_{6-i} - \sum_{i=-6}^{+6} \frac{1}{2}(x_{2,2k-i} + x_{3,2k-i}) \cdot d_{6-i} \right| = \sum_{k=N_1}^{N_2} \left| \sum_{i=-6}^{+6} d_{6-i} \cdot \left(x_{1,2k-i} - \frac{1}{2}(x_{2,2k-i} + x_{3,2k-i}) \right) \right| \leq \sum_{k=N_1}^{N_2} \sum_{i=-6}^{+6} |d_{6-i}| \cdot \left| x_{1,2k-i} - \frac{1}{2}(x_{2,2k-i} + x_{3,2k-i}) \right|. \quad (1)$$

将式(1)展开合并同类项, 当 $4 \leq N_2 - N_1 \leq 8$ 时, 可得到下面的关系式

$$\sum_{k=N_1}^{N_2} \sum_{i=-6}^{+6} |d_{6-i}| \cdot \left| x_{1,2k-i} - \frac{1}{2}(x_{2,2k-i} + x_{3,2k-i}) \right| = \left| x_{1,2N_1} - \frac{1}{2}(x_{2,2N_1} + x_{3,2N_1}) \right| \cdot \sum_{i=0}^{+3} |d_{6+2i}| + \left| x_{1,2N_1+1} - \frac{1}{2}(x_{2,2N_1+1} + x_{3,2N_1+1}) \right| \cdot \sum_{i=0}^{+3} |d_{5+2i}| + \left| x_{1,2(N_1+k)} - \frac{1}{2}(x_{2,2(N_1+k)} + x_{3,2(N_1+k)}) \right| \cdot \sum_{i=-k}^{+3} |d_{6+2i}| + \left| x_{1,2(N_1+k)} - \frac{1}{2}(x_{2,2(N_1+k)+1} + x_{3,2(N_1+k)+1}) \right| \cdot \sum_{i=-k}^{+3} |d_{5+2i}| + \dots + \left| x_{1,2N_2} - \frac{1}{2}(x_{2,2N_2} + x_{3,2N_2}) \right| \cdot \sum_{i=0}^{+3} |d_{6+2i}| + \left| x_{1,2N_2+1} - \frac{1}{2}(x_{2,2N_2+1} + x_{3,2N_2+1}) \right| \cdot \sum_{i=0}^{+2} |d_{5+2i}| + |R_{es}|, |R_{es}| 是余项之和, 大于 0.$$

计算各项 $|d_i|$ 之和, 可以得不等式

$$\sum_{n=N_1}^{N_2} \left| X'_{1,n} - \frac{1}{2}(X'_{2,n} + X'_{3,n}) \right| \leq 0.58 \cdot \left| x_{1,2N_1} - \frac{1}{2}(x_{2,2N_1} + x_{3,2N_1}) \right| + 0.64 \left| x_{1,2N_1+1} - \frac{1}{2}(x_{2,2N_1+1} + x_{3,2N_1+1}) \right| + \dots + 0.58 \cdot \left| x_{1,2N_2} - \frac{1}{2}(x_{2,2N_2} + x_{3,2N_2}) \right| + 0.64 \cdot \left| x_{1,2N_2+1} - \frac{1}{2}(x_{2,2N_2+1} + x_{3,2N_2+1}) \right| + |R_{es}| =$$

$$\alpha \cdot \sum_{j=0}^{2(N_2-N_1)+1} \left| x_{1,2N_1+j} - \frac{1}{2}(x_{2,2N_1+j} + x_{3,2N_1+j}) \right| + |R_{es}|, \quad (2)$$

即原式满足

$$\sum_{n=N_1}^{N_2} \left| X'_{1,n} - \frac{1}{2}(X'_{2,n} + X'_{3,n}) \right| \leq \alpha \cdot \sum_{j=0}^{2(N_2-N_1)+1} \left| x_{1,2N_1+j} - \frac{1}{2}(x_{2,2N_1+j} + x_{3,2N_1+j}) \right| + |R_{es}|, \quad (3)$$

当 $N = 4$ 时, $\alpha = 0.63$, $N_2 - N_1 = 8$ 时 $\alpha = 0.75$.

$\sum_{n=N_1}^{N_2} \left| X_{1,n} - \frac{1}{2}(X_{2,n} + X_{3,n}) \right|$ 是点列 X_1 与其预测值之间的 SAD, 该预测值由 X_2 和 X_3 取均值生成. 而 $\sum_{j=0}^{2(N_2-N_1)+1} \left| x_{1,2N_1+j} - \frac{1}{2}(x_{2,2N_1+j} + x_{3,2N_1+j}) \right|$ 为其对应的滤波前像素点列的 SAD. 可发现经过一次行滤波后像素点列的 SAD 满足不等式(3).

因为列方向的滤波与行方向一致, 同理可以得到经过两个方向滤波后的不等式.

$$\sum_{N_1}^{N_2} \sum_{K_1}^{K_2} \left| X''_{1,n,k} - \frac{1}{2}(X''_{2,n,k} + X''_{3,n,k}) \right| \leq \alpha \cdot \beta \cdot \sum_{j=0}^J \sum_{i=0}^I \left| x_{1,2N_1+j,2K_1+i} - \frac{1}{2}(x_{2,2N_1+j,2K_1+i} + x_{3,2N_1+j,2K_1+i}) \right| + |R_{es2}|, \quad (4)$$

$J = 2 \cdot (N_2 - N_1) + 1, \quad I = 2 \cdot (K_2 - K_1) + 1,$

其中 $|R_{es2}|$ 是余项之和, 大于 0 且是相对小值. 当 $N_2 - N_1 = 4$ 时 $\alpha = 0.63$, $N_2 - N_1 = 8$ 时 $\alpha = 0.75$. 当 $K_2 - K_1 = 4$ 时 $\beta = 0.63$, $K_2 - K_1 = 8$ 时 $\beta = 0.75$.

设一个像素点列和其预测值在滤波前的差值绝对值之和为 S_{up} , 在经过行列两个方向的下采样滤波后对应差值绝对值之和为 SAD(可标记为 S), 由不等式(4)可以得到

$$S \leq \alpha \cdot \beta \cdot S_{up} + O(S_{up}) \approx \alpha \cdot \beta \cdot S_{up}, \quad (5)$$

其中 $O(S_{up})$ 是 S_{up} 的相对小值, S_{up} 覆盖的像素点区域为 S 的 4 倍.

因为基本层编码帧由增强层通过下采样滤波产生, 如果增强层 B 帧中编码块在运动估计后具有层间预测矢量, 即该编码块的预测块与基本层编码块的预测块相对应, 则该编码块运动估计时得到的整像素 SAD 与基本层对应块 SAD 必定满足不等式(5). 设基本层编码块在运动估计时得到的预测失真为 SAD(标记为 S), 其运动矢量经过上采样后为 \mathbf{P}_{MVP} . 对应的增强层编码块运动估计得到的运动矢量为 \mathbf{M}_V , 失真为 S_{up} , 如果 \mathbf{M}_V 与 \mathbf{P}_{MVP} 相等或者相差很小, 则根据不等式(5), 可以得到如下不等式.

$$S \leq \alpha \cdot \beta \cdot S_{up} \approx \alpha \cdot \beta \cdot S_{up}, \quad \text{即 } S_{up} \geq \mu S, \quad \mu = \frac{1}{\alpha \cdot \beta}. \quad (6)$$

由此, 可得到增强层编码块在无层间残差预测运动估计 SAD 的下门限 $\mu \cdot S$. 根据 α 和 β 的不同取值, 可得到不同块分割对应的下门限 T : $T_{8 \times 8} = 1.78S$, $T_{8 \times 4} = T_{4 \times 8} = 2.12S$, $T_{4 \times 4} = 2.51S$.

对于增强层中的 BLRM 宏块, 其编码块在无层间残差预测运动估计的整像素点 SAD 主要集中在一定的区域内. 因为如果增强层的编码块无层间残差预测运动估计的整像素点 SAD 小于对应的门限, 说明经过运动补偿后的残差相对较小, 无需再进行层间残差预测编码. 对于 SKIP 模式, 因为不需要编码残差和运动矢量等信息, 所以如果其预测失真小于 $T_{8 \times 8}$, 则说明采用 SKIP 的编码效果肯定优于 BLRM 编码, 也不需要再进行层间残差预测编码.

2.2 算法描述

根据上述分析, 可得出: (1) 对于参考帧距离小于 8 的编码帧, 如果增强层宏块在 PMBP 模式下, 进行无残差预测运动估计时所有编码块的整像素点 SAD 小于门限, 其最终采用层间残差预测的概率很小, 为此可以避免基于层间残差预测的运动估计. 而对于采用层间残差预测的宏块其运动信息的编码基本上采用 BLMM 模式, 所以可以直接利用层间预测模式和层间预测矢量进行运动估计, 从而进一步降低了运动估计的复杂度. (2) 对于参考帧距离大于等于 8 的编码帧, 在进行基于残差的运动估计时, 需要遍历 PMBP 模式和 8×8 宏块模式及其子宏块模式. 详细算法如下:

Step 1 确定增强层当前编码宏块对应的基本层 8×8 块, 以及该 8×8 所在的宏块. 将基本层的相应 $8 \times$

8 块的运动矢量进行上采样, 得到层间预测矢量 P_{MVP} 和编码模式 PMBP. 根据 PMBP 的块分割, 由基本层 8×8 块的各个子块 SAD, 按照 2.1 算法计算 TSAD.

Step 2 进行无残差预测的运动估计, 遍历所有的编码模式得到最佳运动矢量和编码模式. 由 2.1 的描述, 设定以下判决条件: (1) 基本层对应块的残差为零. (2) 在 SKIP 模式下预测失真小于 $TS_{8 \times 8}$. (3) 基本层对应块的残差不全为零, 在 PMBP 模式下所有子块 SAD 小于对应门限. 如果编码时只要有一个条件满足, 进入 Step4, 否则进入 Step 3.

Step 3 进行基于残差预测运动估计: 如果参考帧距离小于 8, 宏块编码模式取 PMBP 和运动矢量预测值取 P_{MVP} ; 否则还需要遍历 8×8 及其子模式, 得到基于残差预测的最佳运动矢量和宏块模式.

Step 4 如果进行了 Step 3, 比较 Step 2 和 Step 3 的编码宏块率失真, 得到当前编码宏块的最佳运动矢量和参考帧以及编码模式, 否则由 Step 2 得到编码宏块的最佳编码信息.

上述算法主要针对最低空域增强层来描述, 对于存在多个空域增强层的编码器, 如果某增强层宏块对应的基本层编码块所在的宏块采用了 BLRM 模式, 则该宏块的运动估计采用原 JSVM 算法, 不进行 SAD 预估和判决.

3 试验与分析

利用 JSVM6. 8. 2 对算法进行验证. 实验计算机平台为 Intel P4 3.0 GHz CPU, 512M 内存. 实验采用具有不同运动测试序列, 设置两个空域层, 基本层序列为 QCIF 格式, 增强层为 CIF 格式. 时域分级设 4 级, GOP 设置为 16, 编码帧频都为 30, 共编码 96 帧. 其他的编码参数设置如下: 采用 CAVLC 熵编码, 运用无误码的率失真模式判别, 参考帧数为 1, 基本层 QP 为 40. 各序列增强层运动估计时间和编码图像 PSNR 的比较结果如表 1.

表 1 算法的性能比较

Sequence	增强层 QP	PSNR/dB		ME Time/(毫秒·帧 ⁻¹)		
		JSVM	文中算法	JSVM	文中算法	提升/%
Akiyo	20	46.40	46.38	20726	13424	35.23
	30	41.84	41.84	17884	10988	38.56
Paris	20	43.17	43.16	21109	13872	34.28
	30	36.90	39.89	19938	12870	35.45
Foreman	20	42.66	42.65	23842	15173	36.36
	30	38.14	38.13	22648	14245	37.10

从表 1 可以看出, 与原 JSVM 相比, 文中算法 PSNR 的变化很小, 该变化来源于提前判决导致的误判断. 而算法使增强层的每帧运动估计时间平均减少 35%, 这是因为增强层 B 帧中采用基于残差预测的宏块很少, 文中算法的 SAD 提前终止判决可以保证满足判决条件(2)和(3)的编码宏块, 无需进行基于残差预测的运动估计, 从而节约了大量的编码时间. 在参考帧为 1 的情况下, 如果基本层对应块的残差不全为零, 原 JSVM 算法需要进行不同模式、不同预测矢量下 16 次的运动估计, 文中算法在最差情况下只需要进行 12 次运动估计, 减少了运动估计和编码时间, 从而使编码速度得到了明显的提高.

4 结束语

分析了 SVC 空域增强层运动估计的特点, 提出了一种增强层快速运动估计算法. 该算法利用基本层运动估计整像素 SAD 对增强层基于残差预测的运动估计进行提前判决, 避免不必要的基于残差预测的运动估计; 并且对于必须进行残差预测运动估计的编码宏块, 直接利用层间预测运动信息进行运动估计, 舍弃利用层内相邻块预测矢量的运动估计编码. 实验结果表明文中算法显著地降低了增强层运动估计的运算复杂度,