

H. 264 SVC 快速模式选择算法

黄爱爱, 田翔, 周凡, 陈耀武

(浙江大学 数字技术及仪器研究所, 310027 杭州, huangaiai@zju.edu.cn)

摘要: 为了减小可伸缩视频编码(SVC)的编码复杂度,提出了一种 H. 264 SVC 快速模式选择算法. 通过分析 SVC 各编码模式的使用频率和条件概率情况,对出现概率高的模式进行优先搜索,当满足局部最优和率失真代价单调性条件时,提早结束模式搜索. 同时,利用层间模式相关性,除去不适用的模式. 缩小候选模式搜索范围. 实验结果表明,与模式全搜索方法相比,本文算法在图像质量和码率损失极小的前提下,编码速度显著提高.

关键词: 可伸缩视频编码;模式选择;提早退出;层间模式相关性

中图分类号: TN919.81 **文献标志码:** A **文章编号:** 0367-6234(2012)02-0129-05

Fast mode decision algorithm for H. 264 scalable video coding

HUANG Ai-ai, TIAN Xiang, ZHOU Fan, CHEN Yao-wu

(Institute of Advanced Digital Technology and Instrumentation, Zhejiang University, 310027 Hangzhou, China, huangaiai@zju.edu.cn)

Abstract: To reduce the computational complexity of scalable video coding (SVC), a novel fast mode decision algorithm for H. 264 scalable video coding is proposed. First, according to the analysis of mode occurrence frequency and conditional probability, the frequently occurred modes were examined at priority. If both the local-optimal condition and the monotony of rate-distortion cost condition were satisfied, the mode search process was terminated. Moreover, the candidate modes were selected based on the inter-layer mode correlation and the not suitable modes were eliminated. Comparing with the full mode search method, the proposed method accelerates the coding speed remarkably with negligible loss of quality and bit-rate.

Key words: scalable video coding; mode decision; early-termination; inter-layer mode correlation

可伸缩视频编码(Scalable video coding, SVC),作为 H. 264/AVC 的可伸缩扩展,是目前解决网络异构性和终端多样性问题的有力工具. 和以往可伸缩内容相比,SVC 在很多环节上采用了新的技术,比如可分级 P/B 帧结构、层间运动和残差预测、单一解码重建环路等,这些新技术的应用大大提高了 SVC 的编码效率,与联播(Simulcast)相比,在相同编码质量下,SVC 至少可以节省 10% 的码率^[1]. SVC 不仅使用了 H. 264/AVC

中原有的 7 种帧间预测模式(16×16 、 16×8 、 8×16 、 8×8 、 8×4 、 4×8 、 4×4)和 2 种帧内预测模式(intra4 $\times$ 4、intra16 $\times$ 16),还使用了 2 种基于基本层信息预测的新模式(BLSKIP、intraBL). 为了提高层间编码效率,SVC 引入层间预测工具,从残差、运动和纹理 3 方面来去除层间信息冗余. 多模式的运用和层间预测工具的使用极大地提高了编码效率和率失真性能. 然而,传统的模式选择方法是对所有模式在所有可能的层间预测选项下进行全搜索,获得最优模式,其运算复杂度极大. 随着编码层数的扩充,其运算复杂度也将成倍增加. 因此,模式选择的快速算法是目前 SVC 研究中的重要内容.

目前,已有一些针对 SVC 模式快速选择的研

收稿日期: 2012-08-16.

基金项目: 国家自然科学基金重大科研仪器设备研制专项(40927001);浙江省重大科技专项(2009C11035);中央高校基本科研业务费专项资金资助项目.

作者简介: 黄爱爱(1982—),女,博士研究生;
陈耀武(1963—),男,教授,博士生导师.

究^[2-5]. 文献[2]利用 8×8 和 $\text{intra}4 \times 4$ 的搜索结果来选择 $\text{intra}/\text{inter}$ 类型,然后根据基本层宏块模式来筛选候选模式. 文献[3]设计了宏块候选模式、基本层宏块模式以及量化参数 (Q_P) 之间的对照表. 文献[4]则提出从基本层和空间相邻宏块中预测候选模式,通过与率失真代价门限值比较来确定是否提早退出模式搜索. 文献[5]利用基本层对应位置的相邻宏块模式信息来预测候选模式. 然而,上述算法都有一定的局限性. 文献[2]需要强制搜索 8×8 等所有小块模式,复杂度仍然很高;文献[3]中候选模式都是通过查表得到,不能根据实际编码情况进行调整;文献[5]需要进行多次阈值比较,编码效果依赖于阈值设定.

本文提出了一种新的快速模式选择算法,该算法通过判断局部最优和率失真代价单调性条件,提早结束模式搜索,并且利用 SVC 层间模式相关性,缩小候选模式搜索范围. 实验表明,本文算法在信噪比和比特率性能下降很小的情况下,极大地节省了运算复杂度.

1 SVC 模式特性分析

尽管 SVC 中编码模式很多,但通过对不同复杂度视频序列的仿真,发现各种模式的使用频率并不均等,且有一定的规律性. 测试 4 个典型序列 (bus、football、foreman 和 mobile), 2 组量化参数,其中基本层量化参数 (Q_{PB}) 分别取 $Q_{PB} = 20$ 和 $Q_{PB} = 30$,增强层量化参数 (Q_{PE}) 根据可伸缩类型设置. 首先,统计增强层上各个模式的使用频率,图 1 给出了测试序列在量化参数 $Q_{PB} = 30$ 时的实验结果,分别包括粗粒度质量可伸缩 (CGS) 和空间可伸缩编码,其中 $8 \times 8\text{sub}$ 包括 8×8 、 8×4 、 4×8 、 4×4 共 4 种模式. 由图可知,BLSKIP 和 SKIP 出现频率最高,占了总数的 50% 以上,当 Q_{PB} 与 Q_{PE} 差距增大时,其使用频率也随之增加;intra 模式使用频率较小,平均小于 10%;其他模式的使用频率随图像内容和量化参数而变化. 其次,统计增强层宏块在不同基本层宏块模式 ($\text{mode}_{\text{base}}$) 下的条件概率,表 1 列出了空间可伸缩下的统计结果. 结果表明,增强层和基本层在宏块模式选择上存在很强相关性,表现为:1) 当 $\text{mode}_{\text{base}}$ 为 inter 类型时,增强层选择 BLSKIP 的可能性很高. 2) 如果 $\text{mode}_{\text{base}}$ 为小块模式,增强层宏块选择大块模式的可能性就很小. 3) 如果 $\text{mode}_{\text{base}}$ 为 16×8 (8×16),增强层宏块选择 8×16 (16×8) 的几率很小. 4) 如果 $\text{mode}_{\text{base}}$ 为 intra(inter),增强层宏块选择 inter(intra) 的可能性就很小.

虽然全搜索的方式能得到最佳模式,却需要非常复杂的计算. 如果能在模式搜索之前先对模式进行筛选,优先搜索最有可能的模式,或者除去不可能的模式,就可以减少不必要的计算,大大提高编码速度.

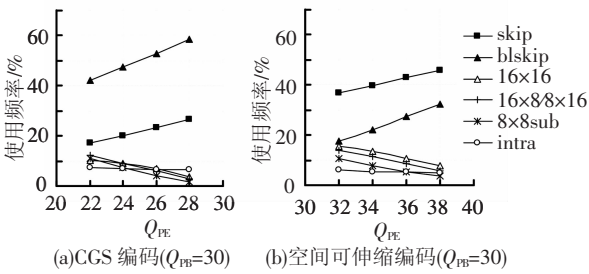


图 1 增强层宏块模式使用频率统计

表 1 空间可伸缩增强层宏块的模式条件概率统计

增强层 宏块 模式	基本层宏块模式条件概率					
	SKIP	16 × 16	16 × 8	8 × 16	8 × 8 sub	intra
SKIP	0. 65	0. 36	0. 22	0. 22	0. 11	0
BLSKIP	0. 14	0. 30	0. 30	0. 31	0. 27	—
16 × 16	0. 12	0. 15	0. 17	0. 16	0. 13	0. 02
16 × 8	0. 03	0. 05	0. 10	0. 07	0. 10	0. 02
8 × 16	0. 02	0. 05	0. 07	0. 11	0. 11	0. 01
8 × 8sub	0. 04	0. 08	0. 13	0. 12	0. 26	0. 02
intra	0	0. 01	0. 01	0. 01	0. 02	0. 93

2 算法介绍

2.1 提早退出算法

根据第 1 节模式统计可知,BLSKIP 和 SKIP 在编码模式中使用频率很高,与其他模式相比,由于这 2 种模式不需要运动估计运算,其运算复杂度也比较低. 如果能尽早检测出这 2 种模式可以省去计算复杂度相当高的运动估计过程,极大地加快模式决策.

对于提早退出判断方法,文献[4]提出了阈值方法,利用空间相邻宏块和基本层宏块的率失真代价计算得到阈值 (RD_{thresh}),如果候选模式的率失真代价小于 RD_{thresh} ,就结束本次模式搜索. 该方法当 RD_{thresh} 与当前宏块的率失真代价存在较大偏差时,容易造成误判. 文献[6-7]分别提出了 P 帧和 B 帧下 SKIP 模式的提早判断方法,当 $16 \times 16/\text{Direct}$ 的运动估计结果满足特定条件时,判定 SKIP 为最佳模式. 由于该方法只考虑了单层编码情况,在多层编码情况下该方法存在较高误判,其中 50% 以上的误判发生在最佳模式为 BLSKIP 的情况下. 为此,本文提出了基于局部最优和率失真代价单调性的提早退出判断方法,来准确预测 SKIP 和 BLSKIP 是否为最优模式.

文中,全体模式集合记为 Φ ,优先搜索的模式

集合记为 Φ_{pri} , 则 $\Phi_{pri} = \{BLSKIP, SKIP\}$, 同理, $\overline{\Phi}_{pri} = \Phi - \Phi_{pri}$. 设 m^* 、 m_{pri}^* 、 $\overline{m}_{pri}^*$ 分别为 Φ 、 Φ_{pri} 、 $\overline{\Phi}_{pri}$ 中具有较小率失真代价的模式, RD_{m^*} 、 $RD_{m_{pri}^*}$ 、 $RD_{\overline{m}_{pri}^*}$ 分别为这 3 个模式的率失真代价. 由率失真优化理论^[8]可知, 宏块最优模式 m^* 是 m_{pri}^* 和 $\overline{m}_{pri}^*$ 中率失真代价值较小者. 因为计算 $\overline{m}_{pri}^*$ 也需要运动估计和率失真优化运算, 为了减少这部分的运算量, 可以从已有的编码信息得到 $\overline{m}_{pri}^*$ 的预测值, 记为 $m_{compare}$, 这样保证了 m_{pri}^* 满足局部最优条件.

从块大小分析, SKIP 属于大块模式, BLSKIP 的块大小则依赖于基本层对应块. 统计发现, 在空间可伸缩下约有 90% 的 BLSKIP 使用的是 16×16 块大小, 所以 BLSKIP 也可以看作是大块模式. 如果检测出当前宏块是大块模式, 那么最佳模式是 m_{pri}^* 的可能性就很大. 根据文献[9], 可以利用模式率失真代价的单调性来判断大块/小块模式.

结合以上分析, 本节先以 $m_{compare\ 1} = 16 \times 16$ 、 $m_{compare\ 2} = 8 \times 8$ 为例来阐述提前退出算法的基本思路: 如果 m_{pri}^* 满足以下条件, 则判定 $m^* = m_{pri}^*$. $RD_{m_{pri}^*} < RD_{16 \times 16}$, $RD_{m_{pri}^*} < RD_{8 \times 8}$, $RD_{16 \times 16} < RD_{8 \times 8}$.

(1)

图 2 给出了阈值方法^[4]、SKIP 判断方法^[6-7]和本文方法在预测准确率上的比较结果. 由图可知, 文献[4]方法在不同量化参数下预测准确率存在明显波动, 文献[6-7]的方法由于没有考虑多层编码情况造成预测准确率较低, 而本文方法预测准确率最高, 并且波动较小. 实验还进一步比较了率失真单调性条件对预测准确率的影响, 在空间可伸缩下, 率失真单调性条件的约束作用比较明显, 进一步提高了预测准确率.

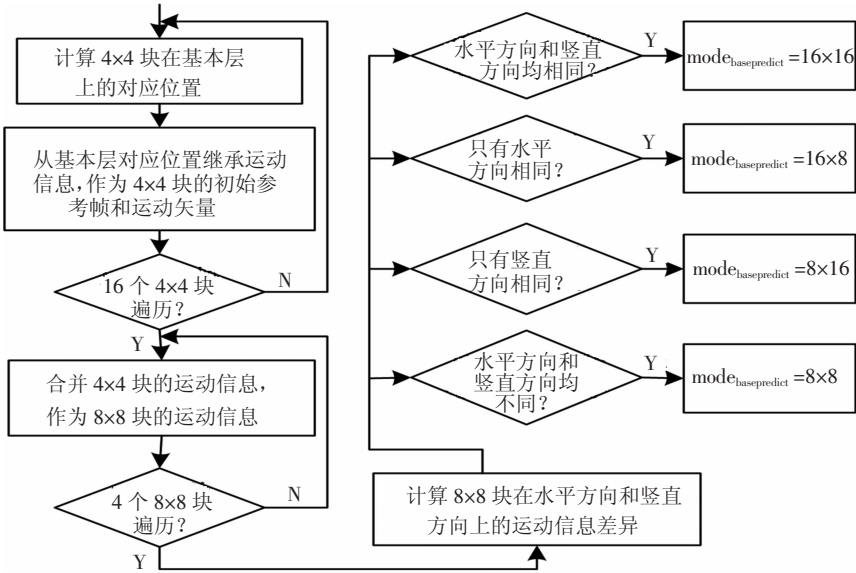


图3 从基本层信息预测增强层宏块大小的流程

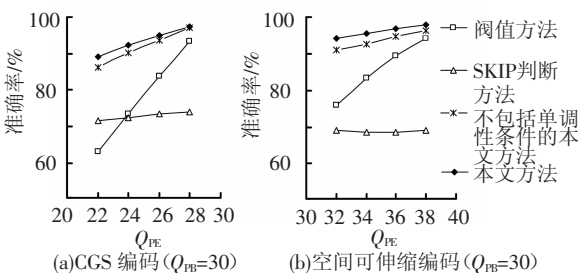


图2 BLSKIP和SKIP模式预测准确率结果比较

从前面分析可知, $m_{compare}$ 的选择是提前退出算法的关键, 一方面用来与 m_{pri}^* 进行比较(以判断局部最优), 另一方面需要包括大块模式和小块模式, 以判断率失真代价单调性. 本文利用增强层宏块和基本层宏块在模式选择上的相关性来计算 $m_{compare}$, 即

$$\begin{cases} m_{compare\ 1} = mode_{basepredict}, \\ m_{compare\ 2} = 16 \times 16, \text{ 若 } mode_{basepredict} = 8 \times 8, \\ m_{compare\ 2} = 8 \times 8, \text{ 若 } mode_{basepredict} \neq 8 \times 8. \end{cases} \quad (2)$$

式中: $mode_{basepredict}$ 为利用基本层宏块的运动矢量信息预测得到的增强层宏块模式^[10], 如图 3 所示.

式(1)中列出的率失真单调性判据只适用于大块模式时的情况, 根据块大小, 提前退出判断条件修正为: 如果 $mode_{basepredict}$ 为 8×8 , 且满足以下条件, 则提前退出.

$$RD_{m_{pri}^*} < RD_{min}, RD_{m_{compare\ 2}} < RD_{m_{compare\ 1}}. \quad (3)$$

如果 $mode_{basepredict}$ 不是 8×8 , 且满足以下条件, 则提前退出.

$$RD_{m_{pri}^*} < RD_{min}, RD_{m_{compare\ 2}} > RD_{m_{compare\ 1}}. \quad (4)$$

式中: RD_{min} 为 $m_{compare\ 1}$ 和 $m_{compare\ 2}$ 中率失真代价值较小值.

2.2 快速模式选择

根据增强层和基本层在模式选择上的相关性,候选模式对应关系如表 2 所示. 与文献[3]相比,根据不同可伸缩选项的模式特性,候选模式对应关系有所不同. 例如,当基本层宏块 $mode_{base} = 8 \times 8_{sub}$ 时,空间可伸缩下仍然需要搜索 $16 \times 16, 16 \times 8, 8 \times 16$,而 CGS 下则不用. 根据表 2, Φ_{pri} 中的 SKIP 模式当 $mode_{base} = 8 \times 8_{sub}$ 或 $16 \times 8/8 \times 16$ (CGS) 时,不需要搜索,从而可以进一步降低运算复杂度.

表 2 候选模式查找表

增强层	基本层宏块模式(空间/CGS)				
候选模式	SKIP	16×16	16×8	8×16	$8 \times 8_{sub}$
SKIP	○/○	○/○	○/×	○/×	×
BLSKIP	○/○	○/○	○/○	○/○	○/○
16×16	○/○	○/○	○/×	○/×	○/×
16×8	×	○/○	○/○	×	○/×
8×16	×	○/○	×	○/○	○/×
$8 \times 8_{sub}$	×	○/×	○/○	○/○	○/○

注:○表示包括,×表示不包括.

图 4 为本文提出的快速模式选择算法整体流程(虚线部分表示提前退出模块),首先进行提前退出决策,当满足提前退出条件时,结束本次模式搜索;当不满足提前退出条件时,根据表 2 查找候选模式,利用提前退出模块的计算结果对候选模式进行调整,然后进行模式搜索获得宏块最佳模式. 考虑到处于不同时间层位置的视频帧对整个 GOP 编码质量的影响程度,提早退出策略只在时间增强层上使用.

图 4 中 intra 模式搜索策略不同于 inter,具体为:当基本层宏块为 intra 编码时,增强层进行传统的模式全搜索;当基本层宏块为 inter 编码时,根据 8×8 运动矢量在水平方向和竖直方向的运动差异来决定是否搜索 intra 模式,即满足以下关系.

$$D_{HM} = \sum_{i=x,y} |mv_{0,i} - mv_{1,i}| + |mv_{2,i} - mv_{3,i}|, \tag{5}$$

$$D_{VM} = \sum_{i=x,y} |mv_{0,i} - mv_{2,i}| + |mv_{1,i} - mv_{3,i}|. \tag{6}$$

式中: D_{HM} 表示水平方向运动差异, D_{VM} 表示竖直方向运动差异; mv_0, mv_1, mv_2, mv_3 为 4 个 8×8 块的运动矢量,按照从左往右,从上至下的顺序; $mv_{i,x}, mv_{i,y}$ 分别为运动矢量 mv_i 的水平分量和竖直分量.

如果满足 $D_{HM} > T$ 或者 $D_{VM} > T$,则进行 intra 模式搜索,实验中 $T = 50$. 此外,在运动估计方面,实验统计发现 CGS 下,使用层间残差预测的

运动估计和不使用层间残差预测的运动估计,产生的结果很接近,所以 CGS 下只需要进行一次运动估计,从而进一步降低运算复杂度.

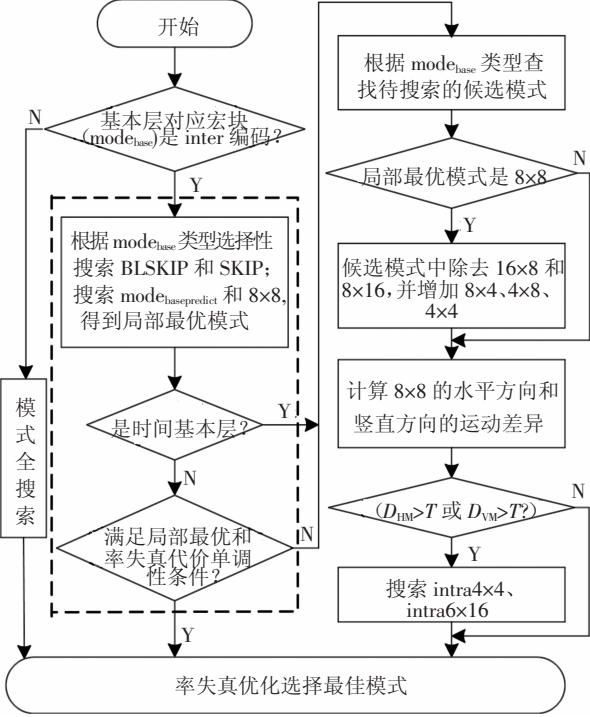


图 4 增强层宏块快速模式选择流程

3 实验结果与分析

本文采用 SVC 参考模型 JSVM9.18^[11] 作为测试平台. 测试配置软件:JSVM9.18;运动估计参数:运动搜索范围为 $\pm 32, 1/4$ 像素精度,1 个参考帧;RDO 正常;GOP 大小为 8. 测试平台: Intel 双核处理器 E4500, CPU 主频 2.20 GHz,内存 0.98 GB;熵编码模式:CAVLC;时域编码结构:可分级 B 帧.

如表 3 所示. 实验使用了 8 个典型测试序列: city, crew, harbour, soccer, bus, football, foreman, mobile,涵盖了全局运动和局部运动,平滑纹理和复杂纹理,以及镜头的平移、缩放等变化. 实验中,采用以下指标来比较本文算法与模式全搜索算法的性能:1)亮度信号的峰值信噪比(R_{SN});2)编码码率(Δ_B);3)编码所花费的时间(T_S).

表 3、4 分别列出了本文算法与 JSVM 模式全搜索方法在 2 种可伸缩选项下的编码性能结果比较. 实验结果表明,在不同量化参数下,编码时间比模式全搜索方法平均减少了 50%,其中在 CGS 下, R_{SN} 平均下降 0.04 dB,码率平均上升 0.09%;在空间可伸缩下, R_{SN} 平均下降 0.05 dB,码率平均上升 0.92%. 可见,提出的快速模式选择算法可以适用于各种类型的视频. 当编码模式中 SKIP 和 BLSKIP 的使用频率增加时,编码速度的提升

效果也变好,这和提早退出算法的应用有着直接的关系.

表 3 CGS 测试数据

序列	(Q_{PB}, Q_{PE})	R_{SN}/dB	$\Delta_B/\%$	$T_S/\%$
city	(20,14)	-0.02	-0.01	51.94
	(30,24)	-0.03	0.17	41.68
crew	(20,14)	-0.05	-0.30	51.55
	(30,24)	-0.08	-0.37	45.17
harbour	(20,14)	-0.01	-0.06	62.12
	(30,24)	-0.02	-0.10	52.60
soccer	(20,14)	-0.03	-0.07	49.98
	(30,24)	-0.06	-0.02	45.30
bus	(20,14)	-0.02	-0.07	57.16
	(30,24)	-0.04	-0.03	47.27
football	(20,14)	-0.03	-0.10	44.95
	(30,24)	-0.07	-0.21	43.71
foreman	(20,14)	-0.05	-0.25	49.88
	(30,24)	-0.06	0.05	41.63
mobile	(20,14)	-0.01	-0.01	59.09
	(30,24)	-0.04	-0.03	49.85

表 4 空间可伸缩测试数据

序列	(Q_{PB}, Q_{PE})	R_{SN}/dB	$\Delta_B/\%$	$T_S/\%$
city	(20,26)	-0.03	2.11	53.70
	(30,36)	-0.01	1.65	57.95
crew	(20,26)	-0.07	0.31	50.33
	(30,36)	-0.06	-0.07	54.82
harbour	(20,26)	-0.05	0.80	53.18
	(30,36)	-0.05	1.28	56.02
soccer	(20,26)	-0.03	0.91	50.97
	(30,36)	-0.03	0.50	54.95
bus	(20,26)	-0.05	1.24	49.45
	(30,36)	-0.06	1.45	52.34
football	(20,26)	-0.07	0.05	45.47
	(30,36)	-0.06	-0.25	51.24
foreman	(20,26)	-0.02	1.47	53.34
	(30,36)	-0.04	0.89	57.56
mobile	(20,26)	-0.06	0.74	53.72
	(30,36)	-0.03	1.58	56.29

4 结 论

1)在深入研究 SVC 模式分布特性的基础上,提出了快速模式选择算法.

2)利用 SVC 模式分布特性对出现概率高的模式进行优先搜索,当满足局部最优和率失真代

价单调性条件时,提早结束模式搜索;利用 SVC 层间模式相关性,缩小候选模式搜索范围.

3)实验结果表明,H. 264 SVC 快速模式选择与模式全搜索相比,在性能损失极小的前提下,编码速度显著提高.

参考文献:

[1] SCHWARZ H, MARPE D, WIEGAND T. Overview of the scalable video coding extension of the H. 264/AVC standard[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2007, 17(9): 1103 – 1120.

[2] LI H, LI Z G, WEN C. Fast mode decision algorithm for inter-frame coding in fully scalable video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(7): 889 – 895.

[3] LIN H C, PENG W H, HANG H M, *et al.* Layer-adaptive mode decision and motion search for scalable video coding with combined coarse granular scalability (CGS) and temporal scalability[C]// Proceedings of International Conference on Image Processing (ICIP). San Antonio: IEEE, 2007: II289 – II292.

[4] REN J, KEHTARNAVAZ N. Fast adaptive termination mode selection for H. 264 scalable video coding[J]. Journal of Real-Time Image Processing, 2009, 4(1): 13 – 21.

[5] KIM S T, KONDA K R, PARK C S, *et al.* Fast mode decision algorithm for inter-layer coding in scalable video coding[J]. IEEE Transactions on Consumer Electronics, 2009, 55(3): 1572 – 1580.

[6] JEON B, LEE J. Fast mode decision for H. 264[C]// Doc JVT-J033. Hawaii: Joint Video Team, 2003.

[7] LEE J, CHOI I, CHOI W, *et al.* Fast mode decision for B slice[C]// Doc JVT-K021. Munich: Joint Video Team, 2004.

[8] SULLIVAN G J, WIEGAND T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Magazine, 1998, 15(6): 74 – 90.

[9] YIN P, TOURAPIS H Y C, TOURAPIS A M, *et al.* Fast mode decision and motion estimation for JVT/H. 264 [C]// Proceedings of 2003 International Conference on Image Processing (ICIP). Barcelona: IEEE, 2003: 853 – 856.

[10] WIEGAND T, SULLIVAN G J, REICHEL J, *et al.* Joint draft 11 of SVC amendment[C]//Doc JVT-X201. Geneva: Joint Video Team, 2007.

[11] Joint Video Team, H. 264/SVC reference software (JS-VM 9.18) and manual[EB/OL]. [2009-06-19]. CVS sever: garcon. ient. rwth-aachen. de.

(编辑 魏希柱)