

# 符合 H. 264 标准的一种基于内容的 变尺寸宏块分割方法

陈倩<sup>1,2)</sup> 高新波<sup>1)</sup>

<sup>1)</sup>(西安电子科技大学工程学院, 西安 710071) <sup>2)</sup>(上海交通大学图像通信与信息处理研究所, 上海 200240)

**摘要** 提出了一种基于视频内容预分析的变尺寸宏块分割(CBVSBM)的实现方案,变尺寸宏块分割是 H. 264 编码器端提供的重要编码工具。该方案根据每帧视频的内容,自适应地在自顶向下的变尺寸运动估计(TDVSBM)和自底向上的变尺寸运动估计(BUVSBM)两种变尺寸宏块分割算法(VSBM)间进行切换,从而结合了两种算法的优点。为了实现两种 VSBM 算法间的自适应选择,提出了一种简单而有效的基于 STGS 图的预判准则。理论分析和实验结果表明,该算法的运算复杂度低于 TDVSBM 和 BUVSBM,而在视频质量上总是优于 TDVSBM 且与 BUVSBM 几乎相当。

**关键词** 变尺寸宏块分割 自适应选择 H. 264/AVC STGS 图

**中图分类号:** TN919.81 **文献标识码:** A **文章编号:** 1006-8961(2007)03-0420-07

## A Content-based Variable-size Block Partition Method for H. 264

CHEN Qian<sup>1,2)</sup>, GAO Xin-bo<sup>1)</sup>

<sup>1)</sup>(School of Electronic Engineering, Xidian University, Xi'an 710071)

<sup>2)</sup>(Institute of Image Communication & Information Processing, Shanghai Jiao Tong University, Shanghai 200240)

**Abstract** As one of the effective coding tools in the newly-published H. 264/AVC coding standard, variable-size block motion estimation is of great importance in performance evaluation of the encoder. A content-based variable-size block partition method is presented for H. 264. The proposed algorithm is based on pre-analysis of video content to select between TDVSBM and BUVSBM that best suits the video in unit of the frame. To enable the adaptive selection, a method based on the STGS-image is proposed to develop an efficient and computationally simple criterion. Experimental results show that the proposed content-based variable-size block motion estimation(CBVSBM) algorithm reduces the computational complexity in contrast to both TDVSBM and BUVSBM, but with similar visual quality.

**Keywords** variable-size block ME, adaptive selection, H. 264/AVC, STGS-image

## 1 引言

H. 264/AVC 是由 ISO/IEC 的 MPEG 和 ITU-T 的 VCEG 共同建立的联合视频编码工作组 (Joint Video Team, 简称 JVT) 制定的新一代面向未来的 IP 和无线环境下的视频编码压缩标准,其在视频的压缩效率以及主、客观视频质量上都比目前所有的视

频编码标准有显著的提高。H. 264/AVC 采用了传统的基于宏块的预测和变换混合编码方案,沿用了以前编码标准中的优秀技术,同时它还提供了一系列新的编码工具。

从运算复杂度来看,解码器的复杂度与以前的标准相当,而编码器由于使用了新的编码工具,在提高编码性能的同时大大增加了其复杂度。其中编码器在帧间预测中支持树形结构的 7 种不同尺寸的宏

基金项目:国家自然科学基金项目(60202004);教育部重点项目(104173)

收稿日期:2005-06-19;改回日期:2005-11-18

第一作者简介:陈倩(1983 ~ ),女。2005 年于西安电子科技大学获学士学位,目前在上海交通大学通信与信息工程专业攻读硕士学位。主要研究方向为视频编码、无线视频传输技术。E-mail: qianchen@sjtu.edu.cn

块分割模式(图 1),如果使用穷尽搜索,虽然从理论上总可以得到最优的模式选择,但处理时间将随着允许使用的模式数目而线性增加。同时,测试结果表明对每个宏块使用 7 种不同尺寸的分割模式将比只使用固定尺寸的  $16 \times 16$  节省近 15% 的码率<sup>[1]</sup>。因此,如何快速获得最佳宏块分割模式已成为目前研究的热点。Chan 等人<sup>[2]</sup>提出了一种自顶而下的变尺寸运动估计方法(TDVSBM);Rhee 等人<sup>[3]</sup>提出了一种自底向上的变尺寸运动估计方法(BUVSBM)。

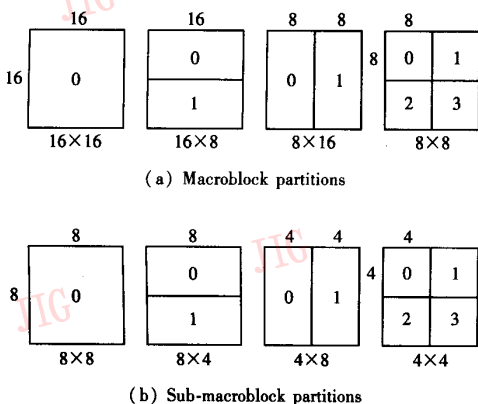


图 1 H. 264 中变尺寸大小的宏块分割模式

Fig. 1 Variable-size macroblock partition modes in H. 264

两种方法均属于四叉树类的决策方法,它们的主要区别在于:TDVSBM 算法根据预先设定的阈值对大块进行分裂操作,而 BUVSBM 算法则对小块进行合并,并在合并的过程中使用动态规划技术决定最佳块尺寸。由于 BUVSBM 算法比 TDVSBM 算法更系统化,因此其视频质量往往要优于后者<sup>[4]</sup>。

然而无论是 TDVSBM 算法还是 BUVSBM 算法,在决定运动估计的最佳块尺寸时都没有考虑视频内容的变化。实际上,如果将视频质量和运算复杂度作为衡量算法优劣的标准,通过后面的实验可以发现,只有对特定的视频序列,BUVSBM 算法的性能才明显优于 TDVSBM 算法。这就启迪我们:如果能预先对视频帧进行内容分析,自适应地从 TDVSBM 算法和 BUVSBM 算法中选择较优的,将会取得比使用单一的任何一种方法更好的效果。在自适应选择中,最关键的问题是找到一种预判准则,它既能准确的识别该帧所具有的关键特征,以决定选用 TDVSBM 算法或 BUVSBM 算法,同时又不过多的增加运算复杂度。为此,提出一种基于 STGS 图

(spatio-temporal gradient scale image)的自适应预判准则,实验结果表明,该准则的运算开销相对于 TDVSBM 算法和 BUVSBM 算法非常小(小于 VSBM 算法的 1%),又能有效地识别帧的特征。将本文算法与 TDVSBM 算法和 BUVSBM 算法进行了广泛的对比,实验结果表明,本文算法在视频质量较 BUVSBM 算法略有下降(小于 1dB)的条件下,运算复杂度较后两者均低。

## 2 视频内容对两种 VSBM 算法性能影响

通过对不同内容的视频序列进行宏块分割的实验结果,分析 TDVSBM 和 BUVSBM 算法性能与视频内容之间的关系。

2.1 两种经典的 VSBM 算法与 H. 264 标准相结合  
将经典的 VSBM 算法与 H. 264 相结合时要明确以下两点:

(1) 进行运动估计的块大小只能是 H. 264 中允许的尺寸,按树形结构依次是  $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8$  和  $4 \times 4$ ;

(2) 运动估计中使用 RD 优化准则,即每一块的最佳运动矢量应使下面的代价函数极小化:

$$J(m, \lambda_{\text{motion}}) = SAD(s, c(m)) + \lambda_{\text{motion}} \cdot R(m - p) \quad (1)$$

其中,  $m = (m_x, m_y)^T$  是真实的运动矢量;  $p = (p_x, p_y)^T$  是预测矢量;  $\lambda_{\text{motion}}$  是拉格朗日乘子;  $R(m - p)$  是对运动矢量差值编码的比特数,可以很容易地通过查表的方式得到<sup>[5]</sup>;绝对误差和 SAD 的定义如下:

$$SAD(s, c(m)) = \sum_{x=1}^M \sum_{y=1}^N |s(x, y) - c(x - m_x, y - m_y)| \quad (2)$$

其中,  $s$  是原始的视频信号;  $c$  是经过运动估计得到的参考信号;  $M, N$  的取值可为  $16, 8, 4$ 。

图 2 示出了 TDVSBM 算法与 H. 264 相结合的具体步骤:以前一帧作参考,对每一个  $16 \times 16$  大小的宏块使用某种搜索方法根据式(1)进行运动估计,如果匹配块对应的 SAD 大于预先设定的某一阈值,则将该宏块依次分裂为  $16 \times 8, 8 \times 16, 8 \times 8, 8 \times 4, 4 \times 8, 4 \times 4$  的块重复该过程,直到各尺寸的子块对应的 SAD 之和小于阈值或该宏块已被分裂为 16 个  $4 \times 4$  的子块为止。

需要指出的是,如果该宏块已被分裂为 16 个  $4 \times 4$  的子块,检测此时对应的 SAD 之和是否是对该

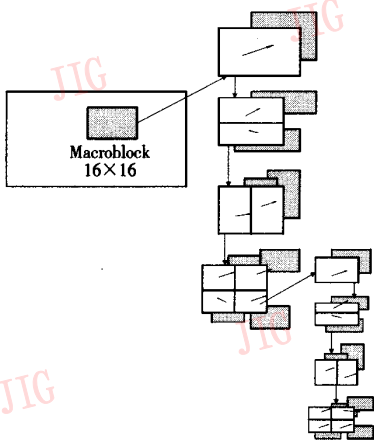


图 2 TDVSBM 与 H.264 相结合的算法图解

Fig.2 TDVSBM combined with H.264

宏块的所有分裂形式中最小的;若不是,将与最小 SAD 值相对应的分裂形式作为该宏块的最终分割模式,以避免小块运动估计的 SAD 之和可能大于相应大块运动估计得到的 SAD 的情况<sup>[6]</sup>。

经典的 BUVSBM 算法与 H.264 相结合时,将每个宏块分成 16 个  $4 \times 4$  的块,并对每个块以前一帧作参考使用某种搜索方法进行运动估计,根据得到的运动矢量情况由  $4 \times 4$  大小的块开始逐次向上合并,合并的具体方法参见文献[3]。

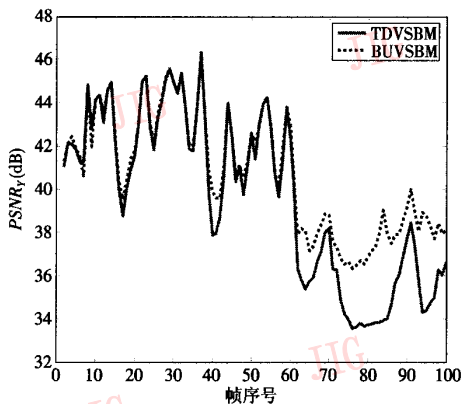
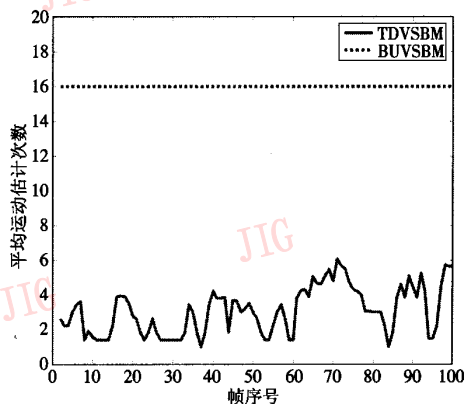
## 2.2 两种 VSBM 应用于不同测试序列的实验结果

实验中,将两种 VSBM 算法分别应用于标准测试序列 Miss\_am, football, foreman, 并对比两种算法得到的视频质量(以 Y 分量的峰值信噪比  $PSNR_Y$  表示)和运算复杂度(假设各尺寸的块进行运动估计

的搜索范围完全相同,对每个宏块进行运动估计的次数可以等效为对该宏块的处理时间,并以每帧中宏块的平均运动估计次数表示运算复杂度),观察视频内容对两种 VSBM 算法性能的影响。

图 3 给出了 TDVSBM 和 BUVSBM 算法对 Miss\_am 视频质量的影响和相应的平均运动估计次数曲线。显然两者的  $PSNR_Y$  性能相仿,但 TDVSBM 算法的平均运动估计次数要明显少于 BUVSBM 算法。因此,TDVSBM 算法的整体性能比 BUVSBM 算法优;图 4 是 football 序列的实验结果,可以看出,使用 TDVSBM 算法较 BUVSBM 算法的视频质量差,同时运算复杂度也明显高于后者,即 BUVSBM 算法比 TDVSBM 算法性能要优;而对于 foreman 序列,由图 5 的结果可以发现,使用 TDVSBM 算法和 BUVSBM 算法的性能随不同的帧而不同,相对前两个序列来说变化情况比较复杂。

从视频内容分析, Miss\_am 序列中静止背景所占的宏块数目较多,此类宏块经一次运动估计后大多小于预定的 SAD 阈值,因此使用 TDVSBM 算法在保证质量的同时全帧的平均运动估计次数较少;而 football 序列中剧烈和复杂运动占有很大比例,若仍使用 TDVSBM 算法,大部分宏块都不能在一次运动估计后满足 SAD 阈值要求,必须经多次分裂(实验中发现相当数目的宏块都被分裂到最小的  $4 \times 4$  尺寸,此过程需要 41 次运动估计),相应的平均运动估计次数也会升高;而使用 BUVSBM 算法,运动估计次数不随帧内容的运动程度而改变,始终是 16 次/宏块,因此复杂度明显低于 TDVSBM 算法;foreman 序列的运动情况随帧的波动比较大,使得

(a) Miaa\_am 中  $PSNR_Y$  对比

(b) Miaa\_am 中平均运动估计次数对比

图 3 Miss\_am 中两种 VSBM 算法的  $PSNR_Y$  和平均运动估计次数对比Fig.3 The contrasts in  $PSNR_Y$  and average ME times between TDVSBM and BUVSBM in Miss\_am

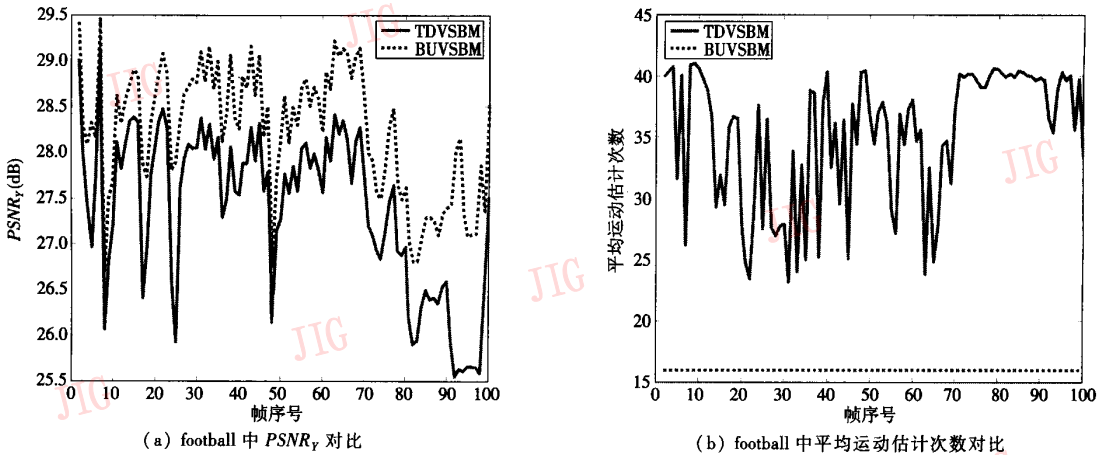


图 4 Football 中两种 VSBM 算法的 PSNR<sub>Y</sub> 和平均运动估计次数对比

Fig. 4 The contrasts in PSNR<sub>Y</sub> and average ME times between TDVSBM and BUVSBM in football

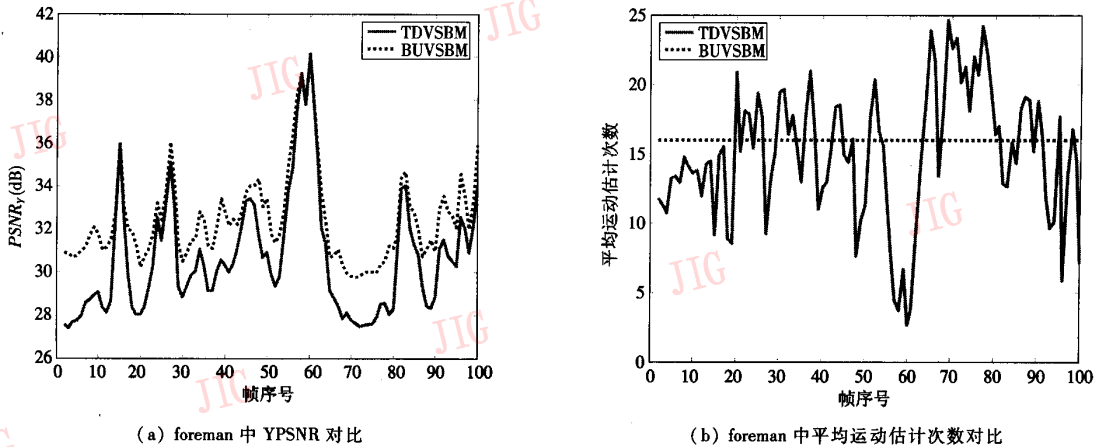


图 5 Foreman 中两种 VSBM 算法的 PSNR<sub>Y</sub> 和平均运动估计次数的对比

Fig. 5 The contrasts in PSNR<sub>Y</sub> and average ME times between TDVSBM and BUVSBM in foreman

TDVSBM 和 BUVSBM 算法的性能交替占优。而在实际的视频中,运动波动情况往往比 foreman 序列更复杂,因此以帧为单位进行 VSBM 算法的选择是有其实际意义的。

### 3 基于内容预分析的变尺寸宏块分割

#### 3.1 STGS 图

STGS 图可以作为视频内容预分析的非常有效的工具,具有较低的运算复杂度,其定义及理论分析参见文献[7]。表 1 示出了 3 个 QCIF 序列中各帧的 STGS 图计算时间相对于其 TDVSBM 和 BUVSBM 算法计算时间的百分比,由表 1 可以看出各帧 STGS 图的运算开销非常小(小于 VSBM 算法的 1%)。

#### 3.2 基于 STGS 图的自适应预判准则

第 2 节中的实验结果显示两种 VSBM 算法对于不同序列在 PSNR<sub>Y</sub> 和平均运动估计次数两个指标上的差异,但还应明确每一帧具体选择哪种 VSBM 算法。为此提出以下算法选择准则:

$$\eta = \lambda \cdot \frac{S_{BU} - S_{TD}}{S_{BU}} + (1 - \lambda) \cdot \frac{T - 16}{16} \quad (5)$$

当  $\eta > 0$ , 该帧选用 BUVSBM 算法;反之当  $\eta < 0$ , 选用 TDVSBM 算法。其中  $S_{BU}$  是使用 BUVSBM 算法得到的 PSNR<sub>Y</sub>;  $S_{TD}$  是使用 TDVSBM 算法得到的 PSNR<sub>Y</sub>;  $T$  是使用 TDVSBM 算法得到的每帧平均运动估计次数;  $\lambda$  是权重因子,它反映视频质量和运算复杂度在选择过程中所占的影响力:  $\lambda \rightarrow 0$ , 运算复杂度的影响力越大;  $\lambda \rightarrow 1$ , 视频质量的影响力越大。

表 1 coastguard,trevor,foreman 各帧 STGS 图相对 TDVSBM 和 BUVSBM 计算时间百分比

Tab.1 Time proportion to TDVSBM and BUVSBM of STGS-image in sequences of coastguard, trevor, and foreman

单位:%

| coastguard |        |        | trevor |        |        | foreman |        |        |
|------------|--------|--------|--------|--------|--------|---------|--------|--------|
| 帧序号        | TDVSBM | BUVSBM | 帧序号    | TDVSBM | BUVSBM | 帧序号     | TDVSBM | BUVSBM |
| 71~72      | 0.23   | 0.41   | 31~32  | 0.81   | 0.41   | 1~2     | 1.31   | 1      |
| 72~73      | 0.29   | 0.36   | 32~33  | 0.91   | 0.47   | 2~3     | 0.5    | 0.35   |
| 73~74      | 0.18   | 0.36   | 33~34  | 0.92   | 0.53   | 3~4     | 1      | 0.66   |
| 74~75      | 0.16   | 0.37   | 34~35  | 0.61   | 0.36   | 4~5     | 0.52   | 0.43   |
| 75~76      | 0.19   | 0.35   | 35~36  | 0.57   | 0.35   | 5~6     | 0.42   | 0.34   |
| 76~77      | 0.15   | 0.38   | 36~37  | 0.60   | 0.34   | 6~7     | 0.48   | 0.38   |
| 77~78      | 0.18   | 0.35   | 37~38  | 0.66   | 0.4    | 7~8     | 0.41   | 0.38   |
| 78~79      | 0.24   | 0.54   | 38~39  | 0.62   | 0.34   | 8~9     | 0.41   | 0.36   |
| 79~80      | 0.16   | 0.39   | 39~40  | 0.51   | 0.35   | 9~10    | 0.45   | 0.37   |

由式(5)经过统计,对 VSBM 算法的选择可以分为以下 3 类:(1)整个序列全部选择 TDVSBM 算法。它对应于具有静止背景的视频,典型代表序列有 Miss\_am,trevor;图 6(a)是 Miss\_am 前 20 帧的 VSBM 算法选择随  $\lambda$  的变化情况,可见在  $\lambda < 0.9$  时,该序列总是选择 TDVSBM 算法;(2)整个序列全部选择 BUVSBM 算法。它对应于具有明显全局运动的视频,典型序列代表是 mobile;图 6(b)是 mobile 前 20 帧的 VSBM 算法选择随  $\lambda$  的变化情况,可见无论  $\lambda$  怎样变化,该序列总是选择 BUVSBM 算法;(3)TDVSBM 算法和 BUVSBM 算法混合使用。它对应于具有复杂运动的视频,典型代表序列是 foreman;图 6(c)是 foreman 前 20 帧的 VSBM 算法选择随  $\lambda$  的变化情况,可以看出  $\lambda$  的变化对各帧的影响不一。

基于 STGS 图的自适应预判准则实际上就是判定每一帧属于上述 3 类运动情况的哪一种。表 2 给出了分属于 3 类的典型序列 Miss\_am, mobile, foreman 前 100 帧对应的 STGS 图中白、黑、灰像素点的平均数目。

表 2 不同视频序列中白、黑、灰像素点平均数目

Tab.2 Average pixel numbers of white, black and grey in sequences Miss\_am, mobile and foreman

| 序列名称          | h-STGS 图中不同像素点平均数目 |        |        | v-STGS 图中不同像素点平均数目 |        |        |
|---------------|--------------------|--------|--------|--------------------|--------|--------|
|               | 白                  | 黑      | 灰      | 白                  | 黑      | 灰      |
| Miss_am(QCIF) | 8 601              | 4 168  | 12 575 | 8 476              | 4 151  | 12 716 |
| mobile(SIF)   | 60 853             | 11 913 | 11 714 | 45 069             | 27 583 | 11 828 |
| foreman(QCIF) | 11 046             | 7 732  | 6 566  | 10 415             | 8 310  | 6 619  |

第 1 类具有静止背景的视频对应的 STGS 图的明显特征是灰色像素点数目即静止部分在整帧中所占的比例较大;第 2 类具有明显全局运动的视频对应的某一类表示运动的像素点在某一方向上(mobile 中是白色像素,水平方向)所占的比例要明显高于其他两种像素;第 3 类具有复杂运动的视频中表示运动的两种像素点(黑色和白色)和表示静止的灰色像素点的数目在两个方向上均相对平衡。

据此规律可以给出下面的预判准则:假设当前帧  $f(t)$  的尺寸为  $M \times N$ ,  $N_{grey}$  为 h-STGS 图和 v-STGS

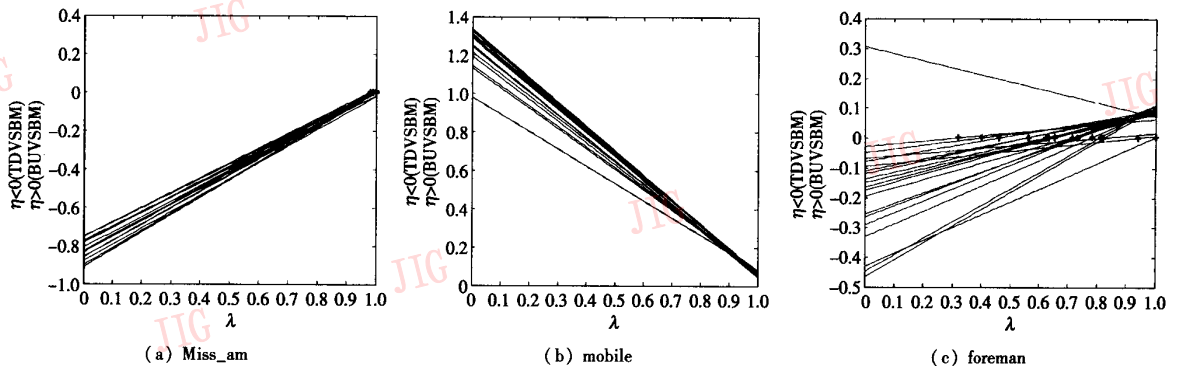


图 6 不同视频序列中的  $\eta \sim \lambda$  的关系

Fig. 6  $\eta \sim \lambda$  in sequences Miss\_am, mobile and foreman

图中灰色像素点数目较大值,  $N_{1-h}$  为 h-STGS 图中黑色或白色像素点中较大一种的数值,  $N_{2-h}$  为较小一种的数值,  $N_{1-v}$ ,  $N_{2-v}$  在 v-STGS 图中表示与  $N_{1-h}$ ,  $N_{2-h}$  相同的含义。

(1) 如果  $\frac{N_{grey}}{M \times N} > T_1$ ,  $f(t)$  属于第 1 类;

(2) 定义运动复杂度

$$d = \frac{1}{\frac{N_1}{N_2} - 1}$$

对于 h-STGS 图  $N_1 = N_{1-h}$ ,  $N_2 = N_{2-h}$ ; 对于 v-STGS 图  $N_1 = N_{1-v}$ ,  $N_2 = N_{2-v}$ ; 取  $d$  为 h-STGS 图和 v-STGS 图所得  $d$  中较小的一个。如果  $d < T_2$ ,  $f(t)$  属于第 2 类; 否则  $f(t)$  属于第 3 类;

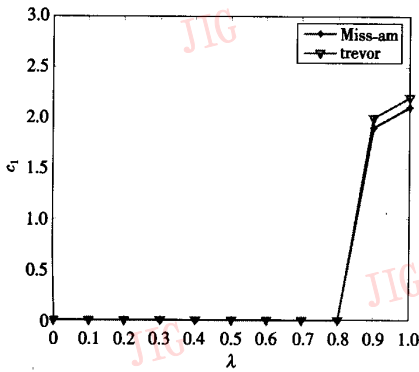
(3) 若已判定  $f(t)$  属于第 2 类,  $f(t)$  选用 BUVSBM 算法;

(4) 若已判定  $f(t)$  属于第 1 类, 则如果  $\frac{N_{grey}}{M \times N} > c_1 \cdot T_{static}$ , 选用 TDVSBM 算法, 否则选用 BUVSBM 算法;

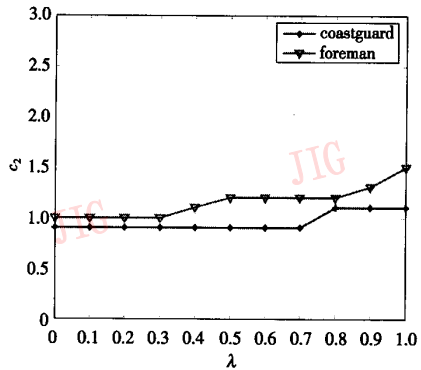
如果已判定  $f(t)$  属于第 3 类, 则如果  $\frac{N_{grey}}{M \times N} > c_2 \cdot T_{complex}$ , 选用 TDVSBM 算法, 否则选用 BUVSBM 算法。

3.3 参数  $\lambda$  的选择对预判准则的影响

由式(5)可知,  $\lambda$  影响着当前帧对 VSBM 算法的选择, 而在基于 STGS 图的预判准则中, 第 1 类和第 3 类视频帧对 VSBM 算法的选择分别取决于阈值  $c_1 \cdot T_{static}$  和  $c_2 \cdot T_{complex}$ , 因此该阈值应该随着  $\lambda$  而变化。这里, 为了讨论方便, 设  $T_{static}$  和  $T_{complex}$  为定值, 只需讨论  $c_1 \sim \lambda$ ,  $c_2 \sim \lambda$  之间的关系。图 7(a) 为第 1 类视频序列 Miss\_am, trevor 得到的  $c_1 \sim \lambda$  关系曲线, 图 7(b) 为第 3 类视频序列 foreman, coastguard 得到的  $c_2 \sim \lambda$  关系曲线。



(a) Miss\_am, trevor 的  $c_1 \sim \lambda$  关系曲线



(b) foreman, coastguard 的  $c_2 \sim \lambda$  关系曲线

图 7 第 1 类视频序列中  $c_1 \sim \lambda$ , 第 3 类视频序列中  $c_2 \sim \lambda$  关系曲线

Fig. 7  $c_1 \sim \lambda$  in Sequences of Class 1 and  $c_2 \sim \lambda$  in Sequences of Class 3

可以看到, 第 1 类视频序列对应的  $c_1 \sim \lambda$  关系曲线, 第 3 类视频序列对应的  $c_2 \sim \lambda$  关系曲线均具有相似的形式; 可分别用一个通式表达:

$$c_1 = 0 \cdot [\varepsilon(\lambda) - \varepsilon(\lambda - t_1)] + (a_2\lambda^2 + a_1\lambda + a_0) \cdot [\varepsilon(\lambda - t_1) - \varepsilon(\lambda - 1)] \quad (6)$$

$$c_2 = 1 \cdot [\varepsilon(\lambda) - \varepsilon(\lambda - t_2)] + (a_2\lambda^2 + a_1\lambda + a_0) \cdot [\varepsilon(\lambda - t_2) - \varepsilon(\lambda - 1)] \quad (7)$$

其中,  $a_0, a_1, a_2$  是二次多项式拟合系数, 由实验给出;  $t_1, t_2$  分别是  $c_1 \sim \lambda, c_2 \sim \lambda$  曲线中的拐点, 亦由实验确定。

### 4 实验结果与分析

实验中选取 miss\_am, foreman, coastguard, table\_

tennis 序列的前 100 帧, 将本文算法与 TDVSBM 算法和 BUVSBM 算法相对比, 其中  $T_1 = 0.4, T_2 = 1, T_{static} = 0.4, T_{complex} = 0.25$ 。得到的经平均后的  $PSNR_Y$  和平均运动估计次数分别如表 3 和表 4 所示。

由结果可以看到, 基于视频内容预分析的变尺寸宏块分割 (CBVSBM) 算法在运算复杂度上基本总是低于 TDVSBM 算法和 BUVSBM 算法, 同时视频质量总是等于或优于 TDVSBM 算法; 在  $\lambda < 1$  时, 视频质量略低于 BUVSBM 算法, 但程度随着  $\lambda$  的取值不同, 视频质量在  $-1 \sim -0.02$  dB 之间变化。由于增加了一个自由度  $\lambda$ , 只要  $\lambda$  选择合适, 将可以在视频质量和运算复杂度之间取得最佳的折衷效果。

表 3 不同  $\lambda$  取值下 3 种算法在 4 组视频序列上得到的  $PSNR_{\gamma}$  对比结果

Tab.3 Comparative results of  $PSNR_{\gamma}$  on four sequences varied with  $\lambda$

|        |             | 单位: dB  |         |            |              |
|--------|-------------|---------|---------|------------|--------------|
|        |             | Miss_am | foreman | coastguard | table_tennis |
| TDVSBM |             | 39.6512 | 30.6717 | 28.8298    | 31.1054      |
| BUVSBM |             | 40.7373 | 32.4702 | 31.0972    | 31.9382      |
| CBVSBM | $\lambda$ 0 | 39.6512 | 31.4330 | 30.6840    | 31.6894      |
|        | 0.1         | 39.6512 | 31.4330 | 30.7504    | 31.6894      |
|        | 0.2         | 39.6512 | 31.4330 | 30.7850    | 31.6894      |
|        | 0.3         | 39.6512 | 31.4429 | 30.8499    | 31.6894      |
|        | 0.4         | 39.6512 | 31.5370 | 30.9219    | 31.6894      |
|        | 0.5         | 39.6512 | 31.6192 | 30.9845    | 31.6894      |
|        | 0.6         | 39.6512 | 31.6799 | 31.0802    | 31.6894      |
|        | 0.7         | 39.6512 | 31.8595 | 31.0802    | 31.6894      |
|        | 0.8         | 39.6512 | 32.1618 | 31.0972    | 31.6894      |
|        | 0.9         | 39.9440 | 32.3800 | 31.0972    | 31.6985      |
| 1.0    | 40.7542     | 32.4702 | 31.0972 | 31.9383    |              |

表 4 不同  $\lambda$  取值下 3 种算法在 4 组视频序列上得到的平均运动估计次数对比结果

Tab.4 Comparative results of average ME times on four sequences varied with  $\lambda$

|        |             | 单位: ME 次数/宏块 |         |            |              |
|--------|-------------|--------------|---------|------------|--------------|
|        |             | Miss_am      | foreman | coastguard | table_tennis |
| TDVSBM |             | 3.0910       | 14.9696 | 23.4716    | 20.4246      |
| BUVSBM |             | 16           | 16      | 16         | 16           |
| CBVSBM | $\lambda$ 0 | 3.0910       | 13.5632 | 15.8693    | 12.6808      |
|        | 0.1         | 3.0910       | 13.5632 | 15.8703    | 12.6808      |
|        | 0.2         | 3.0910       | 13.5632 | 15.8737    | 12.6808      |
|        | 0.3         | 3.0910       | 13.5647 | 15.8841    | 12.6808      |
|        | 0.4         | 3.0910       | 13.5913 | 15.9049    | 12.6808      |
|        | 0.5         | 3.0910       | 13.6239 | 15.9274    | 12.6808      |
|        | 0.6         | 3.0910       | 13.6613 | 15.9805    | 12.6808      |
|        | 0.7         | 3.0910       | 13.8237 | 15.9805    | 12.6808      |
|        | 0.8         | 3.0910       | 14.2763 | 16         | 12.6808      |
|        | 0.9         | 4.0808       | 14.8464 | 16         | 12.7057      |
| 1.0    | 12.8409     | 15.6475      | 16      | 16.2525    |              |

## 5 结论

提出了一种基于视频内容预分析的变尺寸宏块分割算法,它能够根据视频帧的运动特征,在 TDVSBM 算法和 BUVSBM 算法中自适应地选择较优的一种。同时,由于给出了一种简单而有效的基于 STGS 图的自适应预判准则,在节省了运算复杂度的同时视频质量下降很小。并且该算法还给出了

1 维自由度  $\lambda$ , 允许算法在视频质量和运算复杂度之间选择合适的折衷,提高了算法的灵活性。

本文算法还可以从以下两个方向进行改进: (1) 该算法只是自适应地在经典的 TDVSBM 算法和 BUVSBM 算法进行切换,而实际上还有很多改进的 VSBM<sup>[8,9]</sup> 算法都可以与该算法相结合。可以预见,由于各类改进算法都不同程度的比经典的 VSBM 算法性能有所提高,因此它们的自适应组合性能也应相应的有所提高。(2) 文中的自适应预判准则所涉及的阈值  $T_1, T_2, T_{static}$  和  $T_{complex}$  等目前还只是实验得到的经验值,今后的工作还包括为这些阈值的选取提供理论指导。

## 参考文献 (References)

- 1 Chang Andy, Oscar C Au, Yeung Y M. A novel approach to fast multi-block motion estimation for H. 264 video coding [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo (ICME'03) [C], Baltimore, Maryland, USA, 2003, I: 105 ~ 108.
- 2 Chan W, Yu Y, Constantinides A. Variable size block matching motion compensation with applications to video coding [J]. IEE Proceedings on Communications, Speech and Vision, 1990, 137(4): 205 ~ 212.
- 3 Rhee I, Marlin G, Muthukrishnan S, et al. Quadtree-structured variable-size block-matching motion estimation with minimal error [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2000, 10(1): 42 ~ 50.
- 4 Chang Chin-chen, Chen Lin-li, Chen Tung-shou. An improvement of bottom-up variable-sized block matching technique for video compression [J]. IEEE Transactions on Consumer Electronics, 1998, 44(4): 1234 ~ 1242.
- 5 Richardson Iain E G. H. 264 and MPEG4 video compression video coding for next-generation multimedia [M]. Chichester, West Sussex, England: John Wiley & Sons Ltd publishing, 2004.
- 6 Chen Mei-juan, Chu Ming-chung, Pan Chih-wei. Efficient motion estimation algorithm for reduced frame-rate video transcoder [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(2): 269 ~ 275.
- 7 Wei Qi, Zhang Hong-jiang, Zhong Yu-zhuo. A pre-analysis method for robust global motion estimation [A]. In: Proceedings of the 6th IEEE International Conference on Image Processing (ICIP'99) [C], Kobe, Japan, 1999: 625 ~ 628.
- 8 Kuang Tu Y, Yang J, Shen Y, et al. Fast variable-size block motion estimation using merging procedure with an adaptive threshold [A]. In: Proceedings of IEEE International Conference on Multimedia and Expo (ICME'03) [C], Baltimore, Maryland, USA, 2003, II: 789 ~ 792.
- 9 Chang Andy, Peter H W Wong, Yeung Y M. et al. Fast multi-block selection for H. 264 video coding [A]. In: Proceedings of International Symposium on Circuits and Systems (ISSCAS'04) [C], Vancouver, Canada, 2004, III: 817 ~ 820.