

一种高精度的压缩域视频目标分割算法

唐志峰 王诗俊 杨树元
(中国科学院声学研究所 北京 100080)

摘要: 该文提出了一种工作于 MPEG 压缩域的快速视频目标分割算法。该算法以从 MPEG1/2 码流中部分解码提取的特征为输入, 提取 P 帧中的运动目标。针对一般的压缩域算法目标边界精度不高的特点, 算法采用 I 帧和 P 帧中每个块的直流 DCT 系数和 3 个交流 DCT 系数, 以及运动补偿信息, 重建出 P 帧的原图像 1/16 大小的子图像, 采用快速平均移聚类得到具有较高边界精度的亮度一致的区域; 针对运动向量的噪声容易造成错误检测的缺点, 算法结合聚类分析结果和运动块的分布, 采用基于马尔可夫随机场的统计标号方法对目标和背景区域进行分类, 得到每个 P 帧的目标掩模。该算法可以得到 4×4 子块的边界精度, 对于 CIF 格式的码流, 在 Pentium IV 2GHz 平台上可以达到每秒 40 帧的处理速度。

关键词: 视频目标分割; 压缩域; 快速平均移聚类; 马尔可夫场

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2007)12-2965-05

A High Precision Compressed Domain Approach for Video Object Segmentation

Tang Zhi-feng Wang Shi-jun Yang Shu-yuan
(Institute of Acoustics, Chinese Academy of Sciences, Beijing 100080, China)

Abstract: A fast video object segmentation method working in MPEG compressed domain is presented in this paper. Moving object masks in P frames are extracted by exploiting features obtained by partial decoding. To increase object boundary precision, for each P frame, a 1/16 sub image is constructed using DC and three AC coefficients, and motion compensation information, then a fast mean shift clustering algorithm is used to divide the image into regions with coherence luminance and obtain high precision region boundaries. For reducing the influence of motion vector noise, a MRF-based statistical labeling method is exploited to classify regions into two classes: moving object and background. The proposed algorithm can get a boundary precision of 4×4 sub-block with a high processing speed. For CIF video streams, the algorithm can run at a speed of 40 frames per second in a Pentium IV 2GHz platform.

Key words: Video object segmentation; compressed domain; fast mean shift clustering; Markov Random Field (MRF)

1 引言

视频目标分割是基于内容的视频图像分析和基于目标的视频压缩算法中的关键技术。对于已有的采用视频编码算法压缩的视频数据, 像素域的目标分割算法^[1-3]需要对视频码流进行全解码, 得到每帧图像, 并估计每个像素的运动向量, 因而运算复杂度较高, 不能满足实时提取视频目标的需要。压缩域的算法^[4-12]只需对码流进行部分解码, 或者利用 DCT 变换系数, 节省完全解码的时间的同时, 使算法在相对较小的子图像上进行, 或者利用码流中包含的运动向量, 省略运动估计过程, 从而达到快速目标分割的目的。

已有的压缩域的算法只能达到 8×8 块的边界精度, 通过采用像素域的边界细化的方法可以得到像素精度的边

界^[7, 8], 但是由于边界细化是在压缩域分割的基础上进行的, 最终的分割结果取决于压缩域分割的准确度。另一方面, 视频码流中的运动向量是用于运动补偿目的的, 一般采用块匹配的方法获得, 因而与实际的运动向量存在较大误差, 容易对视频分割的结果造成影响。

本文提出的算法利用 4 个 DCT 系数(1 个 DC 系数, 3 个 AC 系数)和运动补偿信息构建相当于原图 1/16 大小的子图像, 结合快速平均移聚类^[13, 14]得到 4×4 子块的区域边界, 从而提高边界精度; 在前景和背景分离阶段, 采用基于马尔可夫场的统计标号方法将所有的区域分成前景和背景两类, 去除运动向量噪声对分割结果的影响。

2 算法设计

在 MPEG1/2 标准中, 16×16 大小的宏块是运动补偿的最小单位, 依靠运动信息区分运动目标与背景只能得到 16

×16 宏块级的精度；并且宏块的运动向量的目的是为了块的预测，与真实的运动向量相比，存在匹配错误和量化误差。因此，依据宏块的运动向量只能得到运动目标的大致分布。视频目标的时空分割同时利用了颜色和运动特征，可以得到运动目标较精细的边界，并减弱错误的运动向量的影响。像素域的时空联合分割算法利用了每个像素的颜色特征，可以得到像素级的边界精度。由于 MPEG1/2 中存在以 8×8 的块为单位的 DCT 变换系数，而其中直流系数表示块的颜色平均值，利用从直流系数和运动向量构建的直流图像表示原图像的颜色特征进行分割，可以得到 8×8 块的分割精度。本文中提出的算法将 8×8 的块等分成 4 个子块，同时利用了直流系数和少数的交流系数，估计每个 4×4 子块的颜色平均值，使从颜色空间的分割达到 4×4 子块的精度。本文中采用颜色空间的亮度分量的空间分割，结合宏块级的运动块分布，采用基于马尔可夫场的统计标号方法，得到子块级的目标分割精度。

算法的流程如图 1 所示。主要步骤包括子图像构建、空间聚类、运动块标记和背景分离，以下是各个步骤的详细描述。

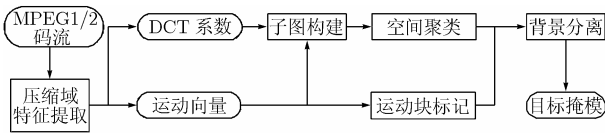


图 1 算法流程图

2.1 子图的构建

MPEG1/2 中以 8×8 的块为单位进行二维 DCT 变换：

$$F(u, v) = \frac{1}{4} e(u) e(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) C_4^{(2x+1)u} C_4^{(2y+1)v},$$

$$u, v = 0, \dots, 7 \tag{1}$$

其中 $e(\tau) = \begin{cases} 1/\sqrt{2}, & \tau = 0 \\ 1, & \text{其它} \end{cases}$ $C_{2N}^r = \cos(\frac{\pi r}{2N})$

将每个 8×8 的块分成四个 4×4 大小的子块，可以利用部分 DCT 系数近似仅算每个子块的均值^[15]。对于帧内编码的块，如果仅采用 $F(0,0)$, $F(0,1)$, $F(1,0)$ 和 $F(1,1)$ 3 个系数，每个子块的均值可以表示为

$$S_{mn} = \frac{1}{4} \sum_{u=0}^1 \sum_{v=0}^1 e(u) e(v) F(u, v) C_4^{(2m+1)u} C_4^{(2n+1)v},$$

$$m, n = 0, 1 \tag{2}$$

对于帧间预测编码的块，码流中的 DCT 系数为运动补偿后残差部分的 DCT 系数，不能直接采用上述公式计算子块的均值。将当前的一个帧间编码的子块的均值表示为 S^{cur} ，可以采用类似于计算直流图像的方法^[16]近似计算：

$$S^{cur} = S^{res} + S^{ref} \tag{3}$$

其中 S^{res} 是残差部分的均值， S^{ref} 是参考位置的均值：

$$S^{ref} = \frac{hw}{16} S_0^P + \frac{h(4-w)}{16} S_1^P + \frac{(4-h)w}{16} S_2^P + \frac{(4-h)(4-w)}{16} S_3^P \tag{4}$$

其中 S_i^P , $i = 0, 1, 2, 3$, 是当前块在参考帧中的匹配块覆盖的子块的均值，如图 2 所示。MPEG2 中采用了半像素精度的运动向量，计算中仅使用其整数部分，对实验结果没有明显的影响。

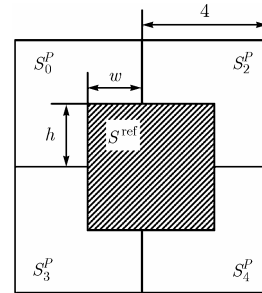


图 2 子块均值估计

2.2 空间分割

空间分割模块采用快速平均移聚类将子图像分割成亮度相似的区域。平均移算法是一种用来估计概率密度梯度的非参数化方法，它不需要概率分布的类型和簇的个数作为先验知识，且具有很好的鲁棒性。平均移算法通过估计特征空间中的局部概率密度，将每个点引申至局部概率密度最大值点，把共享同一个局部密度最大值点的特征点归于同一个簇，达到聚类的目的。

对每个 P 帧，通过前述子图像的构建方法，得到亮度分量的子图像。对子图像中的每个的像素建立一个特征向量：

$$\mathbf{x}(m, n) : (L, m, n) \tag{5}$$

其中 L 为像素的 Y 分量在 LUV 颜色空间的亮度值，采用 LUV 颜色空间便于使用欧几里德距离度量两种颜色的差别^[17]； m 和 n 为像素的水平和垂直坐标。

特征向量 \mathbf{x} 的平均移向量为

$$\text{MeanShift}(\mathbf{x}) = \frac{1}{n_{\mathbf{x}}} \sum_{\mathbf{x} \in S_H(\mathbf{x})} \mathbf{x}_i - \mathbf{x} \tag{6}$$

其中 $S_H(\mathbf{x})$ 为以点 \mathbf{x} 为中心，亮度分量和位置分量的半径分别为 h_l 和 h_s 的超球体， $n_{\mathbf{x}}$ 为落入 $S_H(\mathbf{x})$ 中的数据点的个数； h_l 和 h_s 分别为亮度分量和位置分量的带宽参数。

平均移向量正比于归一化的概率密度梯度，总是指向概率密度上升最快的方向，因此从特征空间中的某个特征点开始，通过计算当前点的平均移向量，将当前点根据平均移向量移动到下一个点，再计算平均移向量继续移动，当平均移向量的模小于某个阈值后，算法收敛，该点停留在局部概率密度最大的位置，收敛点为该簇的模(mode)。设当前点为 \mathbf{x} ，则下一个位置为

$$\text{Move}(\mathbf{x}) = \mathbf{x} + \text{MeanShift}(\mathbf{x}) = \frac{1}{n_{\mathbf{x}}} \sum_{\mathbf{x} \in S_H(\mathbf{x})} \mathbf{x}_i \tag{7}$$

平均移聚类算法的步骤如下:

(1) 选择亮度空间和位置空间的带宽参数 h_l 和 h_s ;

(2) 对于图像中的每一个像素的特征向量 \mathbf{x}_j , 使用平均移式 **错误! 未找到引用源**. 移动, 直到算法收敛, 得到收敛点 $\mathbf{z}_j : (L_j, m_j, n_j)$;

(3) 区域标号: 对属于同一个收敛点的相邻的块标以相同的区域编号, 将该收敛点作为这个区域的模;

(4) 合并亮度特征的模的距离小于 h_l 的相邻区域;

(5) 为了减少下面背景分离步骤的运算量, 将面积较小的区域合并到相邻的区域。

算法主要的运算量集中在第(2)步寻找每个特征点的收敛点的过程, 由于有很多点的收敛路径部分重合, 并且在特征空间中距离较近的点趋向于具有相同的收敛点, 可以通过下面两种途径加快算法^[14]: 移动的过程中记录所有在收敛路径周围半径为 h_l/c 和 h_s/c 的区域内的点, 对这些点赋予相同的收敛点; 当某个点移动到一个已标记过收敛点的位置时, 直接以当前点的收敛点作为该点的收敛点。常数 c 为速度控制因子。实验中采用 $c = 2$ 时仍能获得较好的边界精度, 而运算量减少至未优化算法的约七分之一。

假设经过平均移聚类, 得到 N 个亮度一致的区域:

$R = \{R_1, \dots, R_N\}$; 各个区域的亮度特征的模表示为: $L^m = \{L_1^m, \dots, L_N^m\}$ 。

2.3 运动块标记

通过估计视频中相邻两帧的全局运动, 找出与全局运动不一致的区域, 可以得到运动目标的大致分布。MPEG1/2 码流中存在以宏块为基础的运动向量, 可以用来估计全局运动。假设当前 P 帧与其参考帧(I 帧或 P 帧)之间的全局运动可以由仿射运动参数为 $\mathbf{a} : (a_0, a_1, a_2, a_3, a_4, a_5)$ 的两维仿射运动模型表示, 坐标为 $s : (x, y)$ 的宏块的运动向量为

$$\mathbf{v}_s(\mathbf{a}) = \begin{pmatrix} a_0 + a_1x + a_2y \\ a_3 + a_4x + a_5y \end{pmatrix} \quad (8)$$

将从码流中提取的宏块的运动向量记为 \mathbf{v}_s , 全局运动估计可以表示为下面的最小化问题:

$$\hat{\mathbf{a}} = \arg \min_{\mathbf{a}} \sum_s \rho(\|\mathbf{V}_s - \mathbf{V}_s(\mathbf{a})\|) \quad (9)$$

其中 $\rho(x)$ 是某种鲁棒的 M 估计子。使用鲁棒估计子取代最小平方估计子可以有效去除场景内目标的运动和运动矢量中的噪声引入的外点(outlier)对全局运动估计造成的影响。本文算法采用了 Geman-McClure 算子^[18]: $\rho(x) = x^2 / (x^2 + \sigma^2)$, 其中 σ 为尺度常数。由于式 **错误! 未找到引用源**. 所示的最小化问题可能是非凸的, 为了使结果趋近全局最小值, 先采用较大的尺度计算出的运动参数作为初始值, 然后减小尺度参数重新计算, 迭代至某个较小的尺度终止计算。

在获得了全局运动的仿射运动参数后, 采用提取的运动向量 \mathbf{v}_s 和由运动参数计算得到的运动向量 $\mathbf{v}_s(\hat{\mathbf{a}})$ 之间的欧几

里得距离 $d(\mathbf{v}_s, \mathbf{v}_s(\hat{\mathbf{a}}))$ 衡量宏块与全局运动的一致性, 如果 $d(\mathbf{v}_s, \mathbf{v}_s(\hat{\mathbf{a}}))$ 大于某个门限 T_r , 则认为坐标为 s 的宏块包含的 4 个块(坐标记为 s')为潜在的运动块, 即运动检测掩模为:

$$M(s') = \begin{cases} 1, & d(\mathbf{v}_s, \mathbf{v}_s(\hat{\mathbf{a}})) > T_r \\ 0, & \text{其它} \end{cases} \quad (10)$$

其中门限 T_r 由实验确定。

2.4 背景分离

背景分离模块将聚类所得的区域分成背景和目标两类。一般地, 属于运动目标的区域中包含较多的运动块, 但由于运动向量中存在的噪声, 仅仅依靠运动块的分布, 会产生许多错误的分类。通过基于马尔可夫随机场的统计标号方法, 可以结合观测量(运动块的分布)和上下文信息(相邻区域), 得到统计意义上最优的分类。

将每个区域的标号视为可取两种值的离散随机变量: FL (代表前景, 即运动目标)和 BL (代表背景), 各个随机变量之间的相邻关系由根据各个区域间的邻接关系建立的区域邻接图表示。将各个随机变量对应的观测值表示为 $O = \{O_1, \dots, O_N\}$, 其中第 i 个区域的观测量为 $O_i = \{L_i^m, M_i, N_i\}$, 其中 M_i 表示该区域内运动块的分布信息, N_i 表示该区域包含的块的个数。随机场的一种标号方式 $\omega = \{\omega_1, \dots, \omega_N\}$ 称为随机场的一个配置(configuration)。已知观测值和随机场中各个随机变量之间的相邻关系, 可以使用最大后验概率准则得到最优的配置:

$$\hat{\omega} = \arg \min_{\omega} U(\omega | O) \quad (11)$$

其中 $U(\omega | O)$ 为最大后验概率能量函数:

$$U(\omega | O) = \sum_{i=1}^N V_i^D(\omega_i, O_i) + \frac{1}{2} \sum_{i=1}^N \sum_{j \in G_i} V_{ij}^R(\omega_i, \omega_j, O_i, O_j) \quad (12)$$

第 1 项表示数据项的能量函数:

$$V_i^D(\omega_i, O_i) = \begin{cases} N_i \cdot (1 - \text{CM}_i)^2, & \omega_i = \text{FL} \\ N_i \cdot \text{CM}_i^2, & \omega_i = \text{BL} \end{cases} \quad (13)$$

其中 CM_i 表示第 i 个区域中运动块的比例:

$$\text{CM}_i = \sum_{s' \in R_i} M(s') / N_i \quad (14)$$

其中 $M(s')$ 在全局运动估计部分获得。区域的面积越大, 观测量越可靠, 采用区域中块的个数对观测量加权。对于较小的区域, 不同的标号引入的能量函数的差别主要由规整化项决定。

第 2 项表示规整化项的能量函数:

$$V_{ij}^R(\omega_i, \omega_j, O_i, O_j) = \begin{cases} \lambda \cdot \text{CL}_{ij} \cdot (1 - \text{sig}_{\gamma, T}(|L_i^m - L_j^m|)), & \omega_i = \omega_j \\ \lambda \cdot \text{CL}_{ij} \cdot \text{sig}_{\gamma, T}(|L_i^m - L_j^m|), & \omega_i \neq \omega_j \end{cases} \quad (15)$$

其中 sigmoid 函数: $\text{sig}_{\gamma, T}(x) = 1 / (1 + \exp(-\gamma(x - T)))$, 将区域间的亮度特征的差归一化, 其中 T 为 sigmoid 函数的中

心, 参数 γ 控制 sigmoid 函数的形状。采用 sigmoid 函数可以避免阶梯函数对转折点的敏感。 CL_{ij} 为区域 i 与区域 j 共享边界的长度占区域 i 的周长的比例(不包括属于图像的边界的部分), 对相邻区域间的邻接关系依据共享边界的长度进行加权。 λ 调整规整化项相对于数据项的权重。通过规整化项, 使得亮度相似的相邻区域具有相同的标号。

式 (11) 的最优化问题采用 HCF(Highest Confidence First)^[19] 算法计算。HCF 算法是一种用于马尔可夫随机场的确定性的组合优化算法, 可以借助优先级队列实现快速算法。HCF 算法中除了前景 (FL) 和背景 (BL) 两种标号外, 引入一种辅助标号: 未确定的标号 (UL), 表示某个点还没有确定的标号。原始的 HCF 算法将各个点的标号均初始化为 UL, 为了加快算法的收敛速度, 将运动与全局运动一致, 即 $d(v_s, v_s(\hat{a}))$ 接近于零的区域的标号初始化为 BL, 其余区域初始化为 UL。

3 实验结果

测试序列为 MPEG 的标准测试序列 Akiyo 和 Coast guard(图像尺寸为 CIF: 352×288), 均采用 IBBPBBPBBPBBPBBPBBPBB 的图像组格式压缩成逐行编码的 MPEG2 的码流。图 3 和图 4 分别为两个序列中一个 P 帧的实验结果。其中 (a) 为完全解码的图像; (b) 为构建的子图像; (c) 为聚类的结果, 不同的灰度表示各个区域亮度特征的模; (d) 中黑色部分为与全局运动不一致的块; (e) 为空间分割部分采用直流图像作为亮度特征时分割的结果; (f)-(j) 为背景分离过程的中间结果(灰色表示不确定的标号, 白色表示背景, 黑色表示运动目标区域), 其中 (f) 为初始的标号, (j) 为最终的标号, 图 3 中 (g)-(j) 分别为 2, 5, 9, 11 次迭代的结果, 图 4 中 (g)-(j) 分别为 7, 12, 15, 17 次迭代的结果。

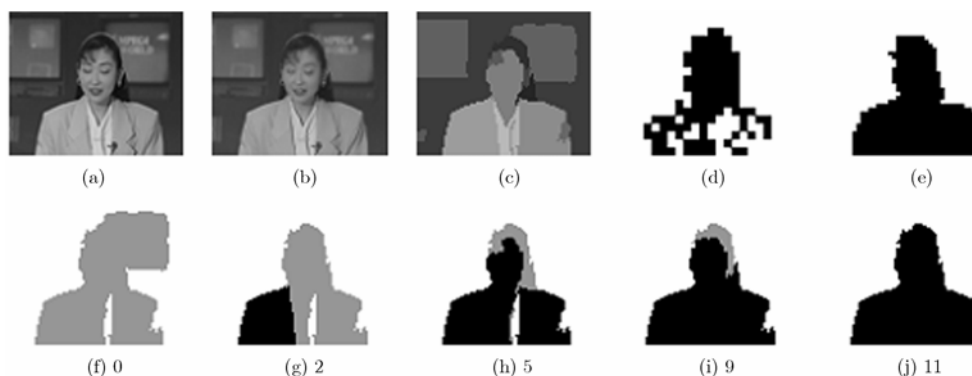


图 3 Akiyo 序列的实验结果

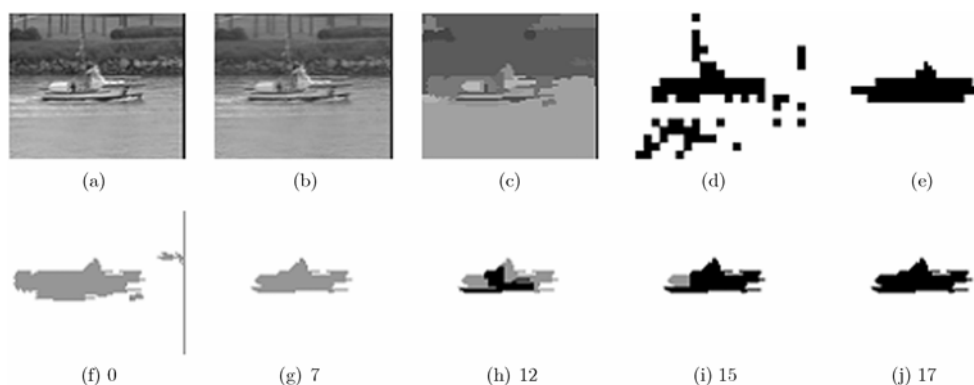


图 4 Coast guard 序列的实验结果

从图中可以看出, 子图像上的平均移聚类可以很好地消除运动适量中噪声的影响, 生成亮度和运动特性一致的区域; 而背景分离过程则可以填补由运动向量的噪声引入的空洞, 并去除运动向量噪声引起的错误检测。比较两图中的 (g) 和 (j), 可以看到, 利用子块的平均值作为亮度特征, 可以有效地提高分割的精度。

表 1 所示为算法各个主要步骤的运行时间, 运行平台为

Pentium 4 2.00GHz, 512M 内存的系统。其中平均移聚类过程的运算速度在速度控制因子确定后, 运算复杂度主要由平均移过程的搜索半径的大小决定, 采用的带宽参数越小, 距离搜索的半径就越小, 运行的速度越快, 实验中采用的带宽参数分别为: $h_t = 6$, $h_s = 3$; 背景分离过程的运算时间与聚类得到的区域的个数相关, 算法的运行时间与聚类过程

得到的区域的个数相关, 实验中聚类过程得到约 20 个区域。处理每个 P 帧的时间约为 25ms。

表 1 主要步骤的运算时间(ms)

主要步骤	运算时间
压缩域信息提取	3.5~4.3
子图像重建	0.2~0.3
全局运动估计	2.4~3.3
快速平均移聚类	11~16
背景分离	0.2~0.3

图 5 和图 6 为本文算法与文献[7]的算法分割效果的比较, 其中(b)为本文算法的结果。可以看出本文的算法可以得到更精细的分割边界, 为进一步像素域的处理提供了更好的基础。

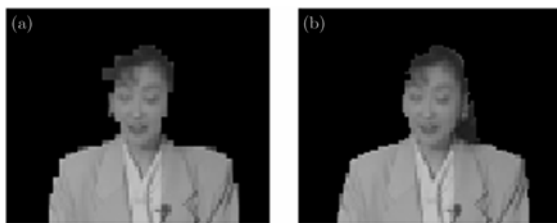


图 5 Akiyo 序列的分割结果比较

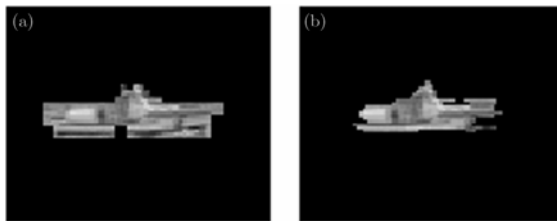


图 6 Coast guard 序列的分割结果比较

4 结束语

本文提出了一种高精度的压缩域视频目标分割算法。该算法采用部分从压缩的码流中提取的 DCT 系数和运动补偿信息重建的子图像代替直流图像, 结合快速平均移聚类提高目标边界的精度; 利用基于马尔可夫场的统计标号方法去除运动向量噪声对分割结果的影响。实验结果表明, 该算法能够在保持实时处理的条件下取得较好的分割结果。在该算法的基础上进行边界细化, 得到像素域的分割精度, 是下一步将要进行的工作。

参考文献

- [1] Megret R, and DeMenthon D. A survey of spatio-temporal grouping techniques. Technical report: University of Maryland, College Park; 2002.
- [2] Tsaig Y, and Averbuch A. Automatic segmentation of moving objects in video sequences: a region labeling approach. *IEEE Trans. on Circuits and Systems for Video Technology*, 2002, 12(7): 597-612.
- [3] Xiao J, and Shah M. Motion layer extraction in the presence of occlusion using graph cuts. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2005, 27(10): 1644-1659.
- [4] Wang R, Zhang H J, and Zhang Y Q. A confidence measure

- based moving object extraction system built for compressed domain. *IEEE International Symposium on Circuits and Systems*, Geneva, Switzerland, 2000, 5: 21-24.
- [5] Porikli F M. Real-time video object segmentation for MPEG-encoded video sequences. *SPIE Conference on Real-Time Imaging VIII*, 2004, 5297: 195-203.
- [6] Zeng W, Du J, and Gao W, et al. Robust moving object segmentation on H.264/AVC compressed video using the block-based MRF model. *Real-Time Imaging*, 2005, 11(4): 290-299.
- [7] Mezaris V, Kompatsiaris I, and Boulgouris N V, et al. Real-time compressed-domain spatiotemporal segmentation and ontologies for video indexing and retrieval. *IEEE Trans. on Circuits and Systems for Video Technology*, 2004, 14(5): 606-621.
- [8] Babu R V, Ramakrishnan K R, and Srinivasan S H. Video object segmentation: a compressed domain approach. *IEEE Trans. on Circuits and Systems for Video Technology*, 2004, 14(4): 462-474.
- [9] How-Lung E, and Kai-Kuang M. Spatiotemporal segmentation of moving video objects over MPEG compressed domain. *IEEE International Conference on Multimedia and Expo*, New York, 2000, 3: 1531-1534.
- [10] Benzougar A, and Bouthemey P, Fablet R. MRF-based moving object detection from MPEG coded video. *IEEE International Conference on Image Processing*, Thessaloniki, Greece, 2001, 3: 402-405.
- [11] Babu R V, and Ramakrishnan K R. Compressed domain motion segmentation for video object extraction. *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Florida, USA, 2002, 4: 3788-3791.
- [12] Sukmarg O, and Rao K R. Fast object detection and segmentation in MPEG compressed domain. *IEEE Region 10 Technical Conference*, Malaysia, 2000, 3: 364-368.
- [13] Comaniciu D and Meer P. Mean shift analysis and applications. *IEEE International Conference on Computer Vision*, Kerkyra, Greece, 1999, 2: 1197-1203.
- [14] Christoudias C M, Georgescu B, and Meer P. Synergism in low level vision. *International Conference on Pattern Recognition*, Quebec City, Canada, 2002, 4: 150-155.
- [15] Chang H S, and Kang K. A compressed domain scheme for classifying block edge patterns. *IEEE Trans. on Image Processing*, 2005, 14(2): 145-151.
- [16] Yeo B L, and Liu B. Rapid scene analysis on compressed video. *IEEE Trans. on Circuits and Systems for Video Technology*, 1995, 5(6): 533-544.
- [17] Comaniciu D, and Meer P. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2002, 24(5): 1-18.
- [18] Black M J, and Anandan P. The robust estimation of multiple motions: parametric and piecewise-smooth flow fields. *Computer Vision and Image Understanding*, 1996, 63(1): 75-104.
- [19] Li S Z. Markov Random Field Modeling in Image Analysis. Tokyo: Springer-Verlag, 2001: 235-237.

唐志峰: 男, 1979 年生, 博士生, 研究方向为计算机视觉、视频编码,
王诗俊: 男, 1978 年生, 博士生, 研究方向为图像处理、视频编码,
杨树元: 男, 1942 年生, 博士生导师, 研究方向为数字信号处理、阵列处理、图像处理与识别。