

采用特征描述的 H. 264 压缩视频运动分割

陆宇¹⁾ 张兆杨^{1,2)} 刘志¹⁾ 韩忠民¹⁾

¹⁾(上海大学通信学院, 上海 200072)

²⁾(上海大学新型显示技术及应用集成教育部重点实验室, 上海 200072)

摘要 H. 264 视频压缩标准以其优良的压缩效率和编码灵活性得到了广泛的应用。提出了一种基于 H. 264 压缩域的运动对象分割方法, 首先从压缩视频中提取运动场, 采用加权中值滤波方法滤除运动场的噪声矢量, 再运用后向估计的方法重建预测运动场并进行运动场的累积, 然后基于幅度、散度和旋度 3 个运动特征, 采用改进的统计区域合并方法将运动对象分割出来。实验结果表明, 该方法可有效地从 H. 264 压缩视频中提取运动对象且分割质量较好。

关键词 H. 264 压缩域 运动特征 分割

中图法分类号: TN941.1 文献标识码: A 文章编号: 1006-8961(2008)10-1995-04

Motion Segmentation Using Characteristic Description for H. 264 Compressed Video

LU Yu¹⁾, ZHANG Zhao-yang^{1,2)}, LIU Zhi¹⁾, HAN Zhong-min¹⁾

¹⁾(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

²⁾(Key Laboratory of Advanced Display and System Application, Ministry of Education, Shanghai University, Shanghai 200072)

Abstract The H. 264 video compression standard is extensively applied thanks to its excellent compression efficiency and coding flexibility. A moving object segmentation approach in H. 264 compressed domain is proposed in this paper. The motion fields are first extracted from the compressed video, in which the noise vectors are removed by weighted median filter. Then the predicted motion fields reconstructed by backward estimation are used to accumulate the motion field. After that, the modified statistical region merging is exploited to segment the moving object based on three motion characteristics magnitude, divergence and curl. Experimental results demonstrate that our approach can efficiently extract the moving objects from H. 264 compressed and as the segmentation quality is good.

Keywords H. 264 compressed domain, motion characteristic, segmentation

1 引言

视频序列中的运动对象常常是引人关注的, 因此视频运动对象分割被广泛地应用于基于内容的视频编码、视频检索和智能视频监控等。基于运动的分割方法主要有 3 种: 第 1 种是基于变化检测的方

法^[1], 先求出当前帧和参考帧的帧差图像, 然后运用阈值分割技术将帧差图像分为静止区域和变化区域, 从而把运动对象分割出来。背景也可以代替参考帧与当前帧进行差分, 对背景的建模可采用高斯混合模型或是非参数的核密度估计等方法, 这种方法容易受到光照变化引入的噪声影响; 第 2 种是同时进行运动估计和运动分割的方法^[2,3], 该方法主

基金项目: 国家自然科学基金项目(60572127, 60602012); 上海市重点学科基金项目(T0102)

收稿日期: 2008-07-11; 改回日期: 2008-07-31

第一作者简介: 陆宇(1977~), 男, 上海大学信息与通信工程专业博士研究生。研究方向为数字视频处理和计算机视觉等。

E-mail: pcrshylu@gmail.com

要采用 MAP 估计和 Bayesian 准则构建包含观察场模型、运动场模型和分割场模型,然后通过优化算法求出代价函数的最优解,这种并行处理的方法运算量比较大;第 3 种是基于运动场的方法,其无需进行预先的运动估计,直接从压缩视频流中提取运动场来分割,提高了分割效率。Xu 等人^[4]采用最大熵模型方法,Zeng 等人^[5]采用马尔可夫模型方法,Liu 等人^[6]则采用遍历二叉树的方法。随着 H. 264 视频压缩标准以其优良的压缩效率和编码灵活性得到广泛应用,近年来基于压缩域的视频对象分割引人关注。

本文 H. 264 压缩域对象分割方法,首先采用加权中值滤波的方法滤除噪声矢量来获得可靠的运动场,然后运用后向估计的方法重建预测运动场并进行运动场的累积。最后基于幅度、散度和旋度 3 个运动特征,采用改进的统计区域合并方法将具有一致性运动特征的区域合并,将运动对象提取出来。

2 加权中值滤波

将 H. 264 压缩视频中不同尺寸的宏块归一化为统一的 4×4 大小的块,然后采用加权中值滤波方法处理归一化的运动场:

$$V_c = \text{med}_{N \times N} \{w_i V_i\} \quad (1)$$

式中, V_c 是滤波后 $N \times N$ 窗口中心的运动矢量, $\text{med}(\cdot)$ 为集合 $\{w_i V_i\}$ 的中值函数, V_i 是 $N \times N$ 窗口内的运动矢量,加权系数 w_i 用于度量窗口内各个运动矢量与统计平均矢量的相关度,由下式计算:

$$w_i = \frac{|\mathbf{V}_i \cdot \mathbf{V}_m|}{\|\mathbf{V}_i\| \|\mathbf{V}_m\|} \quad (2)$$

式中, \mathbf{V}_m 为平均运动矢量。

$$\mathbf{V}_m = \frac{1}{N \times N} \sum_{i=1}^{N \times N} \mathbf{V}_i \quad (3)$$

3 运动场累积

为了得到更为稠密和显著的运动场,需要对运动场进行累积。这里首先采用后向估计的方法重建预测运动场,如图 1 所示。

图中利用 t 时刻的运动场 MF_t 和 $t+1$ 时刻的运动场 MF_{t+1} 来重建预测运动场 $PMF_{t+1,t}$ 。以求取图 1 中 $PMF_{t+1,t}$ 被圆圈标识的块的运动矢量为例,该块被 A、B、C 3 个块所覆盖,这 3 个块是由它们在 MF_t 中的位置以其在 MF_{t+1} 中对应块的运动

矢量为距离经移动而得到的,将这些运动矢量加权平均就得到 $PMF_{t+1,t}$ 中由圆圈标识的块的运动矢量。因此 $PMF_{t+1,t}$ 中每个块的运动矢量由下式求出:

$$PMV_i = \sum_{j=1, S_j \in B_i}^n MV_j \times \frac{S_j}{\sum_j S_j} \quad (4)$$

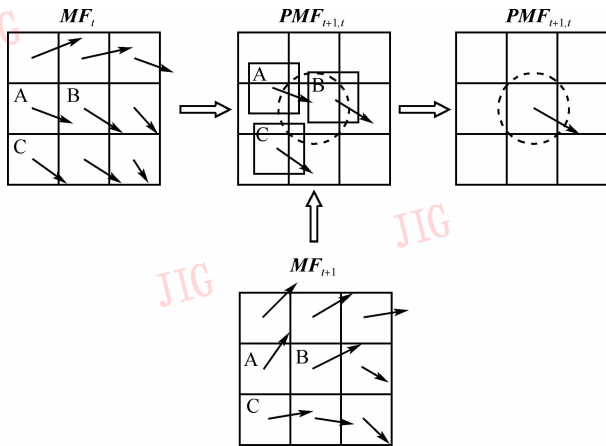


图 1 预测运动场的重建

Fig. 1 Reconstruction of predicted motion field

PMV_i 为 $PMF_{t+1,t}$ 中块 B_i 的运动矢量, n 为覆盖在块 B_i 上面的块数, S_j 为这些块覆盖在块 B_i 上的面积, MV_j 为 MF_t 中对应块的运动矢量。预测运动场重建之后,便可作运动场累积:

$$AMF_t = MF_t + PMF_{t+1,t} + PMF_{t+2,t} \quad (5)$$

式中, AMF_t 为 t 时刻的累积运动场, MF_t 为 t 时刻的运动场, $PMF_{t+1,t}$ 和 $PMF_{t+2,t}$ 为重建的预测运动场。

4 运动对象分割

采用幅度、散度和旋度来描述运动特征。幅度描述了强度特征;散度描述了伸缩度特征;旋度描述了旋转度特征。这 3 个运动特征分别用 M 、 D 、 C 表示:

$$\begin{aligned} M &= \sqrt{V_x^2 + V_y^2} \\ D &= \frac{\partial V_x}{\partial x} + \frac{\partial V_y}{\partial y} \\ C &= \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \end{aligned} \quad (6)$$

式中, $\{V_x, V_y\}$ 是运动矢量的两个分量, $\left\{ \frac{\partial V_x}{\partial x}, \frac{\partial V_x}{\partial y}, \frac{\partial V_y}{\partial x}, \frac{\partial V_y}{\partial y} \right\}$ 由 Sobel 算子计算。

然后基于文献[7]采用改进的统计区域合并方法,将具有 $\{M, D, C\}$ 相似特征的运动对象分割出来,步骤如下:

(1) 计算每个运动矢量 $\{M, D, C\}$ 的值,并将其映射到彩色空间。

(2) 从运动场的左上方第1个运动矢量块开始依次扫描,记录下每个运动矢量块与其四邻域相邻的运动矢量块的 $\{M, D, C\}$ 差值,称为特征差。

(3) 将所有特征差按从小到大的次序进行排列,以具有最小特征差的那一对运动矢量块作为区域合并的开始点。如果两个区域满足以下条件就进行合并:

$$|\overline{R_a^p} - \overline{R_b^p}| \leq \delta(R_a) + \delta(R_b), \forall p \in \{M, D, C\} \quad (7)$$

式中, $\overline{R_a^p}$ 和 $\overline{R_b^p}$ 分别是区域 a 和区域 b 的 $\{M, D, C\}$ 的平均值,边界函数 $\delta(\cdot)$ 为

$$\delta(\mathbf{R}) =$$

$$L \sqrt{\frac{1}{2Q|\mathbf{R}|} (\min(L, |\mathbf{R}|) \log(1 + |\mathbf{R}|) + 2\log_6 |\mathbf{F}|)} \quad (8)$$

式中, Q 为可调节的分割品质因数, L 取值为灰度阶256, $|\mathbf{R}|$ 为区域 \mathbf{R} 的大小, $|\mathbf{F}|$ 为整个运动场的大小。

(4) 将对应于最小运动矢量的区域标记为背景区域,其余区域标记为各个对象区域。

(5) 先作开运算再作闭运算对分割结果进行形

态学滤波。

5 实验结果及评价

采用两个典型的测试序列 Hall 和 Erik 来验证所提出的算法,两个序列均为 352×288 的 CIF 格式。Hall 序列中的人物交替出现而且较小,分割难度较大;而 Erik 序列中人物的头部和身体存在不对称运动。采用 JM8.6 版本的 H.264 编码器,设置为 Baseline Profile, IPPP, I 帧间隔 20 帧,2 个参考帧,搜索范围 $[-32, 32]$ 。

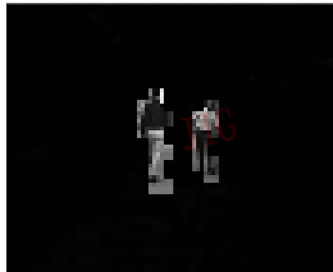
图 2 是 Hall 序列的分割结果,第 1 行为第 88 帧,第 2 行为第 155 帧。与文献[4]分割结果比较,本文分割结果第 88 帧无多余的噪声,第 155 帧中左边人物的双腿被完整地分割出来。

图 3 是 Erik 序列的分割结果,第 1 行为第 16 帧分割结果,第 2 行为第 31 帧分割结果。由图 3 可以看出,文献[5]分割结果的人物内部存在较多的空洞,而且人物的边缘粗糙。而本文分割结果的人物内部均匀而且轮廓平滑。文献[5]方法的 Precision 指标和 Recall 指标起伏较大,显示出分割质量的不稳定。而本文方法这两个指标起伏较小, Precision 指标均在 0.9 以上, Recall 指标均在 0.95 以上,分割效果优于文献[5]结果。

第88帧



第155帧



(a) 原始图像

(b) 文献[4]分割结果

(c) 本文方法分割结果

图 2 Hall 序列分割结果

Fig. 2 Segmentation results of Hall

第16帧



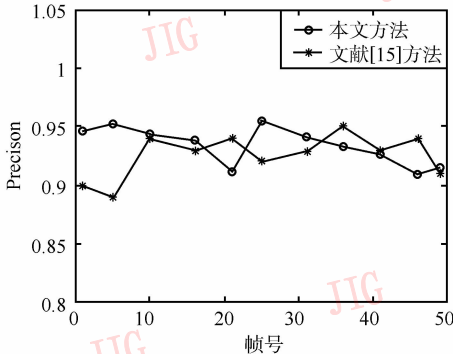
第31帧



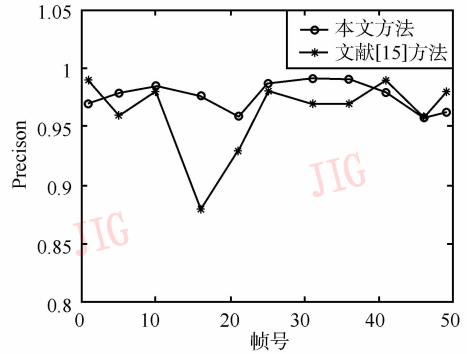
(a) 原始图像

(b) 文献[15]方法分割结果

(c) 本文分割结果



(d) 本文方法和文献[15]方法的Precision评价



(e) 本文方法和文献[15]方法的Precision评价

图 3 Erik 序列分割结果

Fig. 3 Segmentation results of Erik

6 结 论

提出了一种基于幅度、散度和旋度 3 个运动特征作 H. 264 压缩域运动对象分割的方法,实验结果表明该方法能够获得较好的分割质量。由于本文方法仅使用了运动场进行分割,它也可应用于 MPEG 压缩域的运动对象分割。

参考文献 (References)

- Radke R J, Andra S, Al-Kofahi O, et al. Image change detection algorithm; A systematic survey [J]. IEEE Transactions on Image Processing, 2005, 14(3): 294 ~ 307.
- Wang Y, Loe K F, Tan T, et al. Spatiotemporal video segmentation based on graphical models [J]. IEEE Transactions on Image Processing, 2005, 14(7): 937 ~ 947.

- Shen H F, Zhang L P, Huang B, et al. A MAP approach for joint motion estimation, segmentation, and super resolution [J]. IEEE Transactions on Image Processing, 2007, 16(2): 479 ~ 490.
- Xu J F, Liu Z, Zhang Z Y. Compressed domain moving object detection based on entropy model [J]. Journal of Image and Graphics, 2007, 12(10): 1815 ~ 1818. [徐剑峰, 刘志, 张兆杨. 基于熵模型的压缩域运动对象检测[J]. 中国图象图形学报, 2007, 12(10): 1815 ~ 1818.]
- Zeng W, Du J, Gao W, et al. Robust moving object segmentation on H. 264/AVC compressed video using the block-based MRF model [J]. Real-time Imaging, 2004, 11(4): 290 ~ 299.
- Liu Z, Zhang Z Y, Shen L Q. Moving object segmentation in the H. 264 compressed domain [J]. Optical Engineering, 2007, 46(1): 0170031 ~ 0170035.
- Nock R, Nielsen F. Statistical region merging [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(11): 1452 ~ 1458.