

基于变化检测的视频对象提取及 后继帧的对象跟踪

种衍文¹ 江 柳² 沈未名¹

(1 武汉大学测绘遥感信息工程国家重点实验室, 武汉市珞喻路 129 号, 430079)
(2 武汉精伦电子股份有限公司, 武汉市东湖新技术开发区, 430074)

摘要: 提出一种多尺度帧间边界变化检测方法, 将当前图像划分为变化区域和非变化区域, 变化区域内采用基于 Hausdorff 距离跟踪器找到对象在后继帧的最佳匹配位置; 然后利用 Snake 模型拟合该位置上的非刚性形变, 得到对象真实边缘; 最后采用一种基于距离变换的最短路径法使开环闭合。

关键词: 视频对象分割; 对象跟踪; 活动轮廓; 距离变换
中图分类号: TP751

视频编码技术历经了近 30 年的发展, 成功实现了从波形编码到模型编码的转变, 更通用、更强大的第二代编码方式成为多媒体技术的主要发展趋势, 其中 MPEG4 就是一个典型的代表。作为 MPEG4 编码的一个关键环节——视频分割已成为目前研究的热点和难点。

视频分割根据人工参与程度分为半自动分割^[1,2]和全自动分割。在不要求实时性, 但是对视频对象边界精度要求较高的应用场合, 可以采用人工交互的方式确定分割对象, 提高视频分割的精度。全自动分割没有人工参与, 完全根据图像序列中表现出来的运动特征和当前待分割图像中的空间信息实现运动对象的提取。用于视频对象全自动分割的技术主要有运动分割法^[3,4]、时空联合法^[5,6]、变化检测法^[7,8]。

1 首帧对象提取

1.1 变化检测

变化检测通常利用帧间变化差将当前图像划分为变化区域和非变化区域。基于非参数模型的变化检测主要从像素的统计特征着手, 根据其在后续帧中的变化情况判断是否属于变化区域。为此需要选择存储器的时间深度 T , 以便估算像素在 T 内的变化情况, 这势必会导致较大的延时,

而且算法不具备良好的抗噪性。

Kim^[7]等人通过判定帧差图像的边缘像素与当前图像的边缘像素的最小距离是否小于给定阈值来实现前景运动对象的检测, 其中阈值的确定是关键因素。韩军等^[8]在假定噪声服从高斯分布的基础上, 通过互帧差的方差 4 次统计量与某个门限的比较判断出变化区域, 背景假设在一定程度上限制了检测性能, 同时检测开销比较大。Kuo^[9]等提出一种基于灰度帧差的块检测方法, 能够检测出物体的显著变化和微小变化, 并且运算开销相对减小, 但是该检测方法对于图像边界处的抗噪性不够理想。

本文在此基础上提出一种基于多尺度块的帧间边界差的变化检测法, 能够加强抗噪性, 且计算开销小, 具有实时应用的可能。本文将块变化的度量函数 $M_{V_{p \times q}}$ 以及平均变化的度量函数 $M_{V_{avg}}$ 分别定义为:

$$M_{V_{p \times q}} = \sum_{i=0}^{m,j=0}^n \left\{ [G_f(i, j) - G_f'(i, j)] - \sum_{i=0}^{m,j=0}^n \frac{[G_f(i, j) - G_f'(i, j)]}{(m+1) \times (n+1)} \right\}^2 \quad (1)$$

$$M_{V_{avg}} = \frac{\sum_{\text{num}} M_{V_{p \times q}}}{N} \quad (2)$$

其中, m, n 是块的大小; G_f, G_f' 分别表示对当前帧

和下一帧采用 Canny 算子进行边缘检测的结果;

\sum_{num} 表示对所有大小为 $p \times q$ 的块进行求和; N 是大小为 $p \times q$ 的块的个数。当 $M_{V_{p \times q}} > M_{V_{\text{avg}}}$ 时, 则该块标识为变化区域; 当 $M_{V_{p \times q}} < M_{V_{\text{avg}}}$ 时, 则该块标识为非变化区域。多尺度块的检测过程如下。

- 1) 采用 Canny 算子计算相邻两帧的边缘图, 并将边缘图进行帧差计算。
- 2) 选定块的大小, 令初始尺度为 1, 将图像分割成彼此不相邻的块。
- 3) 根据式 (1)、式 (2) 分别计算 $M_{V_{p \times q}}$ 和 $M_{V_{\text{avg}}}$ 。
- 4) 如果 $M_{V_{p \times q}} > M_{V_{\text{avg}}}$, 则该块通过检测, 标识为变化区域; 否则标识为未变化区域。
- 5) 扫描每个块, 如果该块周围相邻 8 个块均落在非变化区域, 则将该块标识为未变化块, 该过程称为孤立块消除。
- 6) 尺度增倍, 判断是否满足尺度终止条件。如果满足, 则转到步骤 7); 否则将初始被检块进行块内划分, 得到更小的子块, 返回步骤 3)。
- 7) 利用被检块的灰度信息, 计算该块的灰度帧差的变化度量, 如果该值大于阈值 TH , 则该块被标识为变化区域; 否则被标识为未变化区域。
- 8) 进行孤立块消除。

该算法中尺度终止条件为尺度和阈值均为 4。阈值与最终尺度块的大小成正比。阈值用来判定被检块的灰度变化情况, 如果被检块在背景区域, 变化度量比较小; 如果该块在变化区域, 其值相对较大。通过结合被检块的灰度帧差信息, 基本上能够消除背景误检块。

分别采用 Kuo 和本文提出的方法对测试序列 Claire 中第 3、4 帧进行变化检测(图 1(a)、1(b)所示), 其中尺度参数设置为 1, 初始块大小为 8×8 , 检测结果分别如图 1(c)、图 1(d)所示。从图中可以看出, 图 1(a) 上端处的误检噪声块在图 1(b) 中没有出现。因此本文提出的检测方法更具鲁棒性, 而且能够检测到前景对象的细微变化, 具有更高的灵敏度。

1.2 提取对象的真实边缘

真实边缘提取包括内部填充、掩模边缘提取、Snake 模型^[10]提取真实边缘三步。内部填充采用传统的填充技术^[7]对检测结果进行填充, 并结合形态学操作对填充结果进行处理从而得到对象掩模。采用 Canny 算子得到的掩模边缘将作为 Snake 模型的初始轮廓, 最后采用 Metropolis 抽样算法^[11]求解 Snake 模型, 边缘提取结果如图 2 所示。

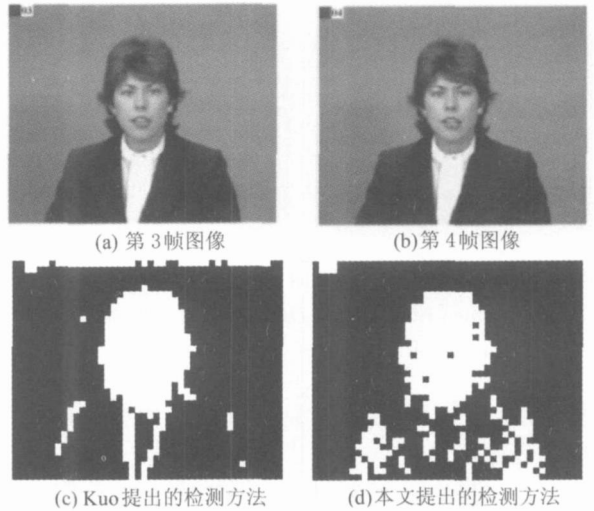


图 1 检测结果对比

Fig. 1 Comparison of Detection Result

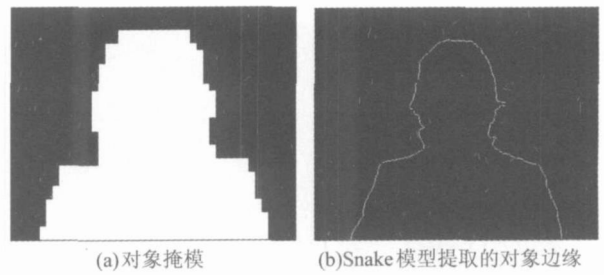


图 2 对象真实边缘提取

Fig. 2 Object Edge Extraction

1.3 开环闭合和 VOP 提取

采用 Snake 模型得到的结果通常是不闭合的, 需要进行开环闭合处理, 再通过填充过程提取视频对象, 从而保证了视频对象的灰度信息不会丢失。

本文采用寻找相邻两个断点 $p(x, y)$ 、 $p'(x', y')$ (这里假定 $x' > x$, $y' > y$) 之间的最短路径的方法使开环闭合。该方法是在对象模型的距离变换图上的, 距离图像上的每个像素值都代表了该像素到最近模型点的距离。

采用该算法对图 2(b) 所示的对象轮廓模型进行闭合处理, 并进行填充实现 VOP 提取, 实验结果分别如图 3 和图 4 所示。

2 后继帧的对象跟踪

根据对象的边缘图像进行匹配, 在计算上是可行的, 而且匹配的效果对亮度变化不敏感, 这也是算法中选择二值模型表达对象的原因。本文选用 Hausdorff 距离作为匹配准则跟踪后继帧中的

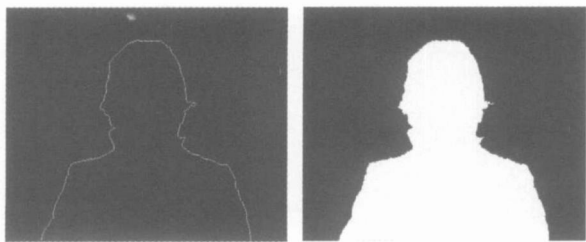


图3 开环闭合的结果
Fig.3 Open Loop Closing Result

图4 视频对象平面提取结果
Fig.4 Video Object Plane Extraction

对象,当对象被部分遮挡或有旋转、变形时,该方法具有鲁棒性。

对于传统定义的 Hausdorff 距离,假设其他所有点都匹配良好,只要存在某个远离中心的模型点或边缘像素,就会导致 Hausdorff 距离较大。为了不受到上述问题的困扰,本文采用广义

Hausdorff 距离^[8],使用 Chamfer5-7-11 模板对边缘图像和模型进行距离变化,在此基础上计算所有位移的 Hausdorff 距离,并取其最小值对应的位移实现模型跟踪。

后继帧中对象的形变分为刚性形变和非刚性形变两种,而实际上非刚性形变的估计是比较困难的。基于 Hausdorff 距离的跟踪器只是在后继帧中找到了对象的最佳匹配位置,但是对象的形状可能在该位置上发生了变化。为了容纳这种变化,本文使用 Snake 模型进一步匹配运动对象的非刚性形变。最后采用开环闭合法对 Snake 模型的解进行处理,得到闭合的对象轮廓。以上结果如图 5 所示,其中图 5(a) 是采用 Hausdorff 距离跟踪得到的对象轮廓,图 5(b) 为 Snake 模型拟合的非刚性形变结果,并进行轮廓闭合处理,图 5(c) 为提取的 VOP。



图5 后继帧的对象跟踪

Fig.5 Object Tracking in Video Sequence

3 结 语

本文的工作重点是从具有静止背景的视频序列中提取前景运动物体,通过结合目标在时域和空域中表现出的各种特征实现分割。

视频对象分割是综合人工智能和模式识别、图像理解、计算机视觉相关学科的一个研究方向,其目的是通过模拟人类视觉系统实现运动对象和背景的分隔。这个过程不仅需要利用对象的运动特征,联合对象的空间信息也是充分必要的。然而对象分割一直以来是机器视觉和图像处理领域中的一个难点,因而语义对象的自动分割尚待进一步深入研究,例如如何结合三维深度信息提取有遮挡的交叠对象或静止对象;如何将运动分割信息与空间分割信息更好地融合等。此外,实时性分割算法也在研究中。

参 考 文 献

- [1] Roberto C. Video Segmentation Based on Multiple Features for Interactive and Automatic Multimedia Application[D]. Trieste: University in Trieste, Italia, 1998
- [2] Hua Zhong, Liu Wenjin. Interactive Tracker: a Semi-Automatic Video Object Tracking and Segmentation System[C]. International Conference on Multimedia Expo, Tokyo, 2001
- [3] Yang Chunke, Shunichiro O. A New Gradient-Based Optical Flow Method and Its Application to Motion Segmentation[C]. 26th Annual Conference of the IEEE, Nagoya, Japan, 2000
- [4] Michael M C, Murat T, Ibrahim S. Simultaneous Motion Estimation and Segmentation[J]. IEEE Trans on Image Processing, 1997(6): 1326-1333
- [5] Moscheni F. Spatio-Temporal Segmentation and Object Tracking: an Application to Second Genera-

- tion Video Coding[D]. Lausanne: Swiss Federal Institute of Technology, 1997
- [6] Moscheni F, Bhattacharjee S, Mural K. Spatio-Temporal Segmentation Based on Region Merging [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 1998, 20(9): 897-915
- [7] Changick K, Hwang J N. Fast and Automatic Video Object Segmentation and Tracking for Content-Based Applications[J]. IEEE Trans on Circuits and Systems for Video Technology, 2002, 12(2): 122-129
- [8] 韩军, 熊璋, 孙文彦, 等. 自动分割及跟踪视频运动对象的一种实现方法[J]. 中国图像图形学报, 2001, 6(8): 732-738
- [9] Kuo C M, Hsieh C H. A New Mesh-Based Temporal-Spatial Segmentation for Image Sequence [C]. Computer Software and Applications Conference, Taipei, 2000
- [10] Kass M, Witkin A, Terzopoulos D, et al. Snake: Active Contour Models[J]. International Journal of Computer Vision, 1987, 1(4): 321-331
- [11] 张蓉, 彭宏. 一种快速的模拟退火算法及其在数据聚类中的应用[J]. 计算机工程与应用, 2001(15): 85-87

第一作者简介: 种衍文, 副研究员, 现主要从事图像与视频编解码、视频通信、计算机视觉等方面的研究。

E-mail: chongyw@lmars.whu.edu.cn

Video Objects Extraction and Tracking in the Video Sequences

CHONG Yanwen¹ JIANG Liu² SHEN Weiming¹

(1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, 129 Luoyu Road, Wuhan 430079, China)

(2 Wuhan Routon Electronic Co. Ltd, Donghu New Technology Development Zone, Wuhan 430074, China)

Abstract: A method for multi-scale change detection is proposed based on inter-frame edge difference. The result shows that the algorithm is robust and well sensitive to change. Snake model is used to extract the object contour. Because of blindness of Snake model, Hausdorff tracker is introduced prior to snake tracker. Snake model is used to track non-rigid deformation in object translation position indicated by Hausdorff tracker. The result of snake model is not always closed, a method of the shortest path based on distance transform is proposed. The segmentation algorithm avoids many of the complex problems associated with optical flow estimation and gray-based segmentation in the space. Experiments also demonstrate this algorithm can handle many traditional types of sequences.

Key words: video object segment; object tracking; Snake model; distance transform

About the first author: CHONG Yanwen, associate researcher fellow, Ph.D. He is engaged in research on realtime video processing and multimedia communication.

E-mail: chongyw@lmars.whu.edu.cn