工学博士学位论文

视频运动对象分割算法研究

曾 炜

哈尔滨工业大学

2005 年 2 月

国内图书分类号：TP391.41
国际图书分类号：681.39

工学博士学位论文

视频运动对象分割算法研究

博士研究生：　曾 炜
导　　师：　高 文　教授
申 请 学 位：　工学博士
学科、专业：　计算机应用技术
所 在 单 位：　计算机科学与工程系
答 辩 日 期：　2005 年 2 月
授予学位单位：　哈尔滨工业大学

Classified Index: TP393.41
U.D.C.: 681.39

Dissertation for the Doctoral Degree in Engineering

A STUDY OF MOVING OBJECT SEGMENTATION FROM VIDEO

Candidate： Zeng Wei
Supervisor： Prof. Gao Wen
Academic Degree Applied for： Doctor of Engineering
Specialty： Computer Application
Date of Defence： February, 2005
Degree-Conferring-Institution： Harbin Institute of Technology

摘 要

视频对象分割是计算机视觉的一个基本问题，得到了广泛而深入的研究。视频对象分割具有广阔的应用前景，如视频编码、视频检索、视频监控、多媒体创作、图像理解、模式识别等。对象分割技术目的是在时间轴上将视频图像分割成一系列相互关联的空间区域。这种基于区域的图像表示方法为视频的灵活处理提供了可能，例如，在视频编码国际标准 MPEG-4 中，采用的就是基于对象的编码方式，不仅提高了压缩效率，而且提供了更加灵活的视频操作方式。又如最新制定的多媒体描述国际标准 MPEG-7，本身就要求支持基于对象方式的内容组织与检索。

本文重点研究面向多媒体检索领域的视频对象分割方法。在时空对象分割算法方面，提出一种基于区域的时空 MRF 模型对象分割方法和一种基于区域的层次 MRF 模型对象分割方法。在基于变化检测的对象分割算法方面，提出一种基于块的运动区域检测方法。在压缩视频对象分割算法方面，提出一种基于变化检测的 MPEG 压缩视频对象分割方法和一种基于块 MRF 模型的 H.264/AVC 压缩视频对象分割方法。具体的研究内容如下：

提出一种基于区域的时空 MRF 模型对象分割方法。该方法在区域 MRF 模型对象分割算法的基础上，集成连续三帧的分割结果，考查了对象区域标记在时间轴上的连续性，集成了区域的时间和空间信息以分割视频运动对象。在算法实现中，为了有效确定区域的运动关系，提出了一种遮挡区域检测方法，有效去除了由于遮挡而形成的运动噪声信息。同时，为了解决区域的过分割问题，提出了一种新的区域合并策略，结合颜色和区域尺寸信息来有效去除小区域。

提出一种基于层次 MRF 模型的对象分割方法。在不同区域分割尺度上，提出的算法利用区域分割的多尺度特性，采用不同形式的 MRF 模型分别完成对象分割的不同任务，并通过各层 MRF 模型的层间相关，形成精确的对象分割。层次 MRF 模型能够有效刻画对象分割的层次关系，是一种在空间层次上渐进式优化对象分割过程的技术。

提出一种基于块的运动区域检测方法。该方法将区域分割结果集成到变化检测中，使分割的运动对象具有精确的边界。提出的算法集成了多种局部线索，形成一个简单而有效的运动区域检测策略规则。为了有效计算局部块的运动特征，提出了一种基于块的运动重估方案，利用相邻块运动估计的结

果作为当前块运动估计的初值，使得重估后的块运动场更加平滑。

提出一种基于变化检测的 MPEG 压缩视频对象分割方法。该方法利用 DC 图重建技术获取压缩视频帧的图像信息，并利用高阶统计技术检测运动的宏块。为了有效确定背景噪声，提出了一种自适应阈值计算方法来自动检测背景区域，并由此估算当前帧的背景噪声。当压缩视频码率较低，或出现大部分运动估计失效的宏块时，该算法能稳定地检测出运动对象。

提出一种基于块 MRF 模型的 H.264/AVC 压缩视频对象分割方法。该方法利用压缩视频帧的运动矢量场来分割运动对象。为了有效遏制噪声，算法利用运动对象标记的跟踪过程来提高分割的精度。此外，针对变尺度块的运动补偿，提出的 MRF 模型利用了块的尺寸信息来提高最终分割结果的致密性，获得了满意的分割效果。

关键词 视频分析；运动对象分割；MRF 模型；压缩视频对象分割

Abstract

Video object (VO) segmentation is a fundamental problem in computer vision, and has been studied widely. VO segmentation techniques have many applications, such as video coding, video retrieval, video surveillance, multimedia production, image understanding, pattern recognition, etc. The goal of object segmentation is to partition the video frames into a series of spatially correlated regions along the time axis. Such separated image representation can provide flexible operations for videos. For example, in the new international video coding standard MPEG-4, video frames are encoded in the object-based way. This type of object-based video coding has not only improved the video compression, but also provided the object-based accessing functionality. Another example is the MPEG-7 standard that is the new international video description standard. In MPEG-7, the content organization and retrieval are required in the object-based way.

In this thesis, the research focuses on the problem of moving VO segmentation for the multimedia processing applications. Two spatiao-temporal VO segmentation algorithms are proposed. One is the region-based spatio-temporal Markov Random Field (STMRF) model, and the other is the region-based hierarchical MRF (HMRF) model. On the other hand, the block-based moving region detection approach is proposed to extract moving VOs accurately. To efficiently extract VOs on compressed video, the change-based VO extraction approach on MPEG compressed video is proposed. For H.264/AVC compressed video, the block-based MRF model is proposed. The detailed description of the above techniques is as follows:

The region-based VO segmentation algorithm using the STMRF model is proposed. This approach integrates three temporary segmentation results obtained from the region-based MRF (RMRF) labeling approach. The proposed method employs the temporal continuity of regions along the time axis, and provides a solution to robust VO segmentation based on temporal and spatial information of regions. In the implementation, an occluded region detection approach to efficiently remove the motion classification errors posed by the occluded

background is presented. On the other hand, a new region merging strategy to deal with the over-segmentation problem of the watershed algorithm is presented, too. The region merging process combines the region color and size information to remove the small noise regions.

The region-based VO segmentation algorithm using the HMRF model is proposed. Based on multi-resolution region segmentation, the approach apply different RMRF models to segment VO in the related region layer. The HMRF model employs the correlation of the region layers, and achieves accurate VO segmentation. The proposed model can efficiently describe the layered character of VO segmentation, and is a gradual optimization of the VO segmentation process along the region resolution.

The block-based moving region detection approach is proposed. The method combines the region segmentation and the change-based VO segmentation approach. Therefore, the obtained VOs have accurate boundaries. The proposed method combines several local clues, and provides a simple but efficient moving region detection results. In order to obtain reliable motion information, a block-based motion re-estimation scheme is presented. The idea underlying the motion re-estimation scheme is using the estimated motion parameters of the adjacent blocks as the initialization of motion estimation for the current block. This method can provide a good initialization in block motion estimation, and obtain smooth and stable block motion field.

The change-based VO segmentation approach on MPEG compressed video is proposed. The approach calculates the frame difference based on the DC image reconstruction from the bitstream directly. The changed macroblocks can be detected through the high order testing technique. In order to estimate the background noise, an adaptive thresholding technique to separate background automatically is presented. As a consequence, the proposed algorithm can determine the threshold of change detection adaptively. The approach is suitable for the video that has the majority of the intra-coded macroblocks or the invalid motion vectors.

The VO segmentation algorithm using the block-based MRF model on H.264/AVC compressed video is proposed. The algorithm extracts VOs from the sparse motion vectors of blocks in the bitstream. To deal with the problem of

noise motion vector field for compressed video, the proposed approach employs the moving object label tracking procedure. To deal with the problem of variable size motion compensation in H.264/AVC, the approach integrates the block size information in the uniformed MRF model. The proposed algorithm obtains more compact VO segmentation finally.

Keywords video analysis; moving object segmentation; MRF model; object segmentation on compressed video