



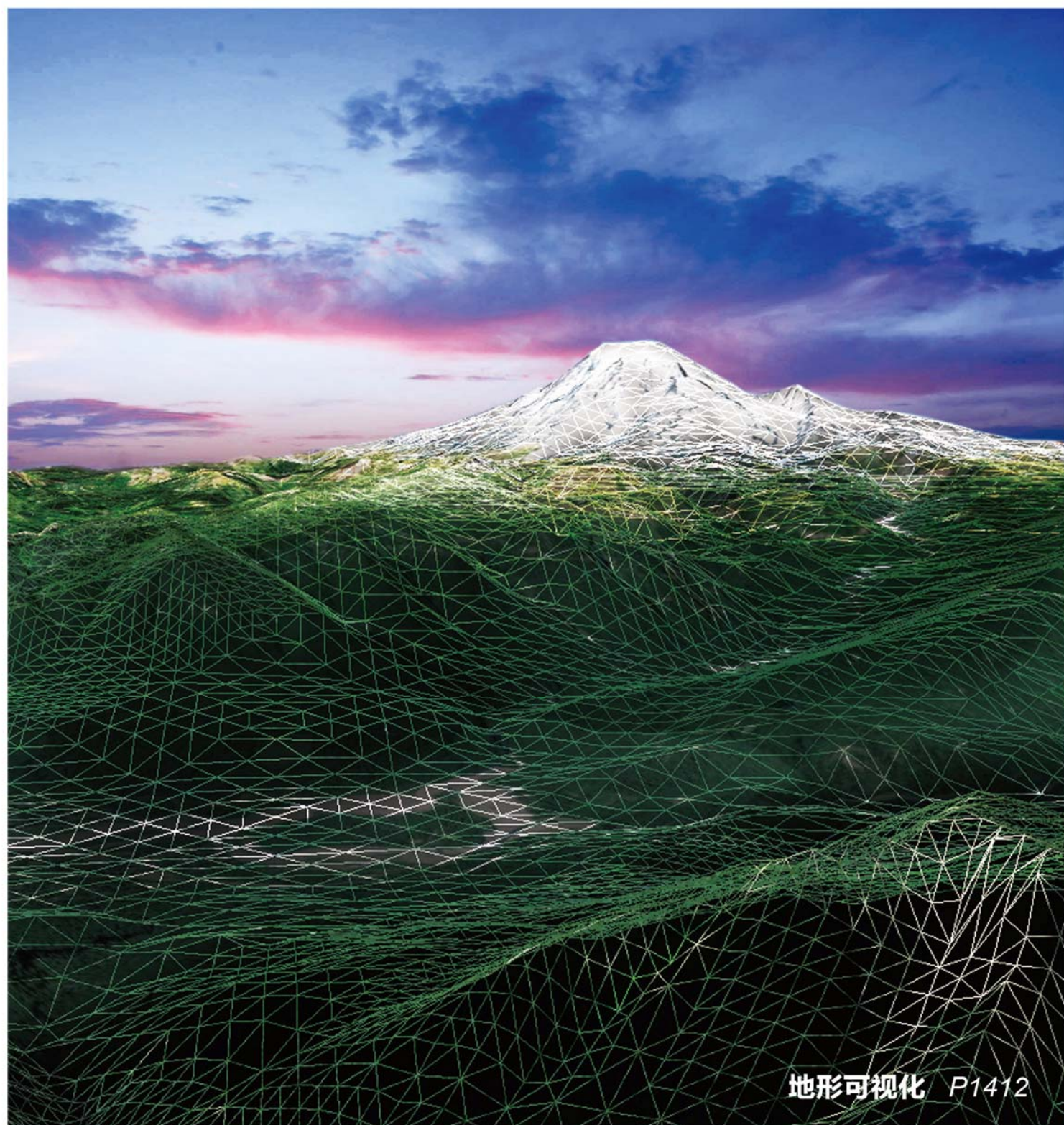
JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2015
10
VOL.20

ISSN1006-8961
CN11-3758/TB



地形可视化 P1412



第20卷第10期（总第234期）

2015年10月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文
编辑出版 《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen
Editor, Publisher Editorial and Publishing Board of Journal of
Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

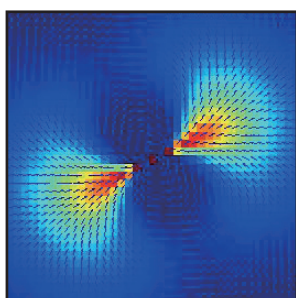
CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 50.00元



JPEG图像双重压缩偏移量估计的篡改区域自动检测定位(第1304页)



基于表情传输的交互式照片编辑(第1366页)



结合张量投票和Snakes模型的SAR图像道路提取(第1403页)

图像处理和编码

基于模式特征的H.264/AVC可逆视频水印

李淑芝, 张翔, 邓小鸿, 吴晓燕1285

二阶总广义变分小波修复模型

许建楼, 郝岩, 张翼1297

JPEG图像双重压缩偏移量估计的篡改区域自动检测定位

赵洁, 郭继昌, 张艳, 张众维1304

图像分析和识别

基于非对称局部梯度编码的人脸表情识别

胡敏, 程轶红, 王晓华, 任福继, 许良凤, 黄晓音1313

采用积分图块间距离检测图像边缘

贾迪, 孟琢, 孙劲光, 李思慧, 赵明远1322

自适应属性加权2维FCM分割算法

侯晓凡, 吴成茂1331

利用层间相关性的岩心CT图像半自动分割方法

徐永进, 滕奇志, 吴晓红, 卿粲波1340

图像理解和计算机视觉

在线特征选取的多示例学习目标跟踪

周志宇, 彭小龙, 吴迪冲, 朱泽飞1346

融合外观和运动特征的在线目标分割

张雷, 李成龙, 汤进, 高思晗1358

基于RGB-D深度相机的室内场景重建

梅峰, 刘京, 李淳芃, 王兆其1366

计算机图形学

带平台伸缩函数的参数曲线变形

张莉, 余慧芳, 檀结庆1374

视景建模中树木纹理图像的随机变形网格方法

施冠羽, 欧阳清1384

基于表情传输的交互式照片编辑

刘锦云, 彭宏京1390

遥感图像处理

结合张量投票和Snakes模型的SAR图像道路提取

符喜优, 张风丽, 王国军, 邵芸1403

GPU Tessellation全球地形可视化方法

李尚林, 郑利平, 张迎凯, 李琳1412

CONTENTS

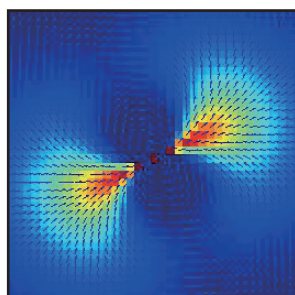
JOURNAL OF IMAGE AND GRAPHICS



Automatic detection and localization of image forgery regions based on offset estimation of double JPEG compression (P1304)



Interactive photo editing based on expression transmission (P1366)



Road extraction from SAR images using tensor voting and Snakes model(P1403)

Image Processing and Coding

- Reversible video watermarking algorithm for H.264/AVC based on mode feature
Li Shuzhi, Zhang Xiang, Deng Xiaohong, Wu Xiaoyan 1285
- Second order total generalized variational model for wavelet inpainting
Xu Jianlou, Hao Yan, Zhang Ji 1297
- Automatic detection and localization of image forgery regions based on offset estimation of double JPEG compression
Zhao Jie, Guo Jichang, Zhang Yan, Zhang Zhongwei 1304

Image Analysis and Recognition

- Facial expression recognition based on asymmetric region local gradient coding
Hu Min, Cheng Yihong, Wang Xiaohua, Ren Fuji, Xu Liangfeng, Huang Xiaoyin 1313
- Edge detection method of block distance combining with summed area table
Jia Di, Meng Lu, Sun Jinguang, Li Sihui, Zhao Mingyuan 1322
- Adaptive weighted two-dimensional histogram FCM segmentation algorithm
Hou Xiaofan, Wu Chengmao 1331
- Semi-segmentation of rock CT images using the correlation of adjacent frames
Xu Yongjin, Teng Qizhi, Wu Xiaohong, Qing Linbo 1340

Image Understanding and Computer Vision

- Object tracking based on multiple instance learning with online feature selection
Zhou Zhiyu, Peng Xiaolong, Wu Dichong, Zhu Zefei 1346
- Online object segmentation via fusing appearance and motion features
Zhang Lei, Li Chenglong, Tang Jin, Gao Sihan 1358
- Improved RGB-D camera based indoor scene reconstruction
Mei Feng, Liu Jing, Li Chunpeng, Wang Zhaoqi 1366

Computer Graphics

- Deformation of parametric curves based on platform extension function
Zhang Li, Yu Huifang, Tan Jieqing 1374
- Method of random transmogrification mesh for tree texture images in scene modeling
Shi Guanyu, Ouyang Qing 1384
- Interactive photo editing based on expression transmission
Liu Jinyun, Peng Hongjin 1390

Remote Sensing Image Processing

- Road extraction from SAR images using tensor voting and Snakes model
Fu Xiyu, Zhang Fengli, Wang Guojun, Shao Yun 1403
- GPU Tessellation method of global terrain visualization
Li Shanglin, Zheng Liping, Zhang Yingkai, Li Lin 1412

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2015)10-1358-08

论文引用格式: Zhang L, Li C L, Tang J, Gao S H. Online object segmentation via fusing appearance and motion features[J]. Journal of Image and Graphics, 2015, 20(10): 1358-1365. [张雷, 李成龙, 汤进, 高思哈. 融合外观和运动特征的在线目标分割[J]. 中国图象图形学报, 2015, 20(10): 1358-1365.][DOI:10.11834/jig.20151009]

融合外观和运动特征的在线目标分割

张雷¹, 李成龙¹, 汤进^{1,2}, 高思哈¹

1. 安徽大学计算机科学与技术学院, 合肥 230601; 2. 安徽省工业图像处理与分析重点实验室, 合肥 230039

摘要: 目的 视频中的目标分割是计算机视觉领域的一个重要课题, 有着极大的研究和应用价值。为此提出一种融合外观和运动特征的在线自动式目标分割方法。方法 首先, 融合外观和运动特征进行目标点估计, 结合上一帧的外观模型估计出当前帧的外观模型。其次, 以超像素为节点构建马尔可夫随机场模型, 结合外观模型和位置先验把分割问题转化为能量最小化问题, 并通过 Graph Cut 进行优化求解。结果 最后, 在两个数据集上与 5 种标准方法进行了对比分析, 同时评估了本文方法的组成成分。本文算法在精度上至少比其他的目标分割算法提升了 44.8%, 且具有较高的分割效率。结论 本文通过融合外观与运动特征实现在线的目标分割, 取得较好的分割结果, 且该方法在复杂场景中也具有较好的鲁棒性。

关键词: 特征融合; MRF 模型; 在线分割; 自动分割

Online object segmentation via fusing appearance and motion features

Zhang Lei¹, Li Chenglong¹, Tang Jin^{1,2}, Gao Sihan¹

1. School of Computer Science and Technology, Anhui University, Hefei 230601, China;

2. Key Lab of Industrial Image Processing & Analysis of Anhui Province, Hefei 230039, China

Abstract: **Objective** Object segmentation in video is an important subject in computer vision and has gained various research and application values. The online automatic object segmentation method is proposed in this paper; this method fuses appearance and motion features. **Method** First, the object points were roughly estimated by employing appearance and motion boundaries. Then, we utilized these estimated object points to refine the appearance model (GMM) of the previous frame as current appearance model. Second, a Markov random field (MRF) model was constructed by taking the superpixels as nodes, and integrating the appearance model and the location prior. Therefore, the object segmentation can be converted to an energy minimization problem, which is optimized by graph cut in this paper. **Result** After extensive experiments which included comparison analysis of five approaches and component analysis of the proposed approach on two datasets, the proposed approach improved accuracy of segmentation by at least 44.8% than other approaches, and achieved higher efficiency of segmentation. **Conclusion** The proposed algorithm achieved online automatic object segmentation by fusing appearance and motion features, and obtained good segmentation performance. Furthermore, this algorithm was also robust in several complicated scenes.

收稿日期: 2015-04-15; 修回日期: 2015-06-23

基金项目: 安徽省科技攻关项目(1301b042002); 安徽省电力公司科技项目(521200130M0U)

第一作者简介: 张雷(1990—), 现为安徽大学计算机应用技术专业硕士研究生, 主要从事视频分析方面的研究工作。E-mail: zhanglei_edu@foxmail.com

通信作者: 汤进, 教授, E-mail: ahhftang@qq.com

Supported by: Scientific and Technological Project in Anhui Province(1301b042002)

Key words: feature fusion; Markov random field model; online segmentation; automatic segmentation

0 引言

视频中的目标分割是把视频中的感兴趣目标从视频中分割出来,并保持空间上的准确性和时间上的一致性。它是计算机视觉中的一个基础的底层问题,分割的目标可用于运动识别、3维重建、视频检索和视频摘要等高层的视觉任务^[1-7]。但目标分割是一个比较难的课题,如视频中的背景复杂、噪声干扰、光照变化等因素极易导致分割失败或精度不高。因此,研究目标分割是非常有必要的。

根据目标初始化方式的不同,目标分割可以分为有监督的目标分割和无监督的目标分割。前者是通过人为方式在前一帧或前几帧标注出目标和背景^[7],或者通过大量标注样本进行离线训练,建立目标和背景模型,而后者则一般利用目标的自身属性来粗略地区分于背景^[4-6],然后建立目标和背景模型并进行迭代优化。无监督的目标分割无需人工干预,有着更广泛的应用,因此本文研究的是无监督的目标分割。

Lee 等人^[4]首先对所有视频帧进行区域划分^[8],并依据外观与运动特征对区域打分,然后对这些区域进行谱聚类,再用得分对这些类进行排序,并认为得分高的类为目标概率大,最后对目标与背景分别建立高斯混合模型,并增加局部形状匹配位

置先验,最后以所有像素点作为节点构建时空马尔可夫随机场(MRF)模型进行目标分割。在此基础上,Zhang 等人^[5]构造了一个分层的有向无环图(DAG)来选取最终的目标区域,并利用光流对目标区域进行扩张,这种方法分割的结果较为准确。但 Lee 和 Zhang 等人的分割方法都依赖于选取的目标区域,只适合单目标的分割情况,且时间复杂度较高。Papazoglou 等人^[6]通过目标的运动特征来初始化目标,然后结合位置先验以及时空平滑约束建立马尔可夫随机场模型快速分割出目标。由于该方法只采用了目标运动特征,在静止或者运动较小的目标分割中往往会失败。值得注意的是,这些视频分割方法都利用了视频的全局信息,这样就导致这些方法都不能处理任意长的视频,具有较大的局限性。

针对以上视频分割中存在的问题,提出一种融合外观特征和运动特征的在线目标分割方法。该方法利用目标的外观和运动特征对目标进行初始化,估计出目标和背景的外观模型,即高斯混合模型(GMM),并结合目标位置先验建立 MRF 模型,进行目标的精确分割,算法流程图如图 1 所示。为了方便描述,图 1 中假设目标和背景的 GMM 均有 3 个单高斯成分组成,分别用红色和黑色表示。本文实验验证了该方法的有效性,分割精度优于现有的其他目标分割方法。

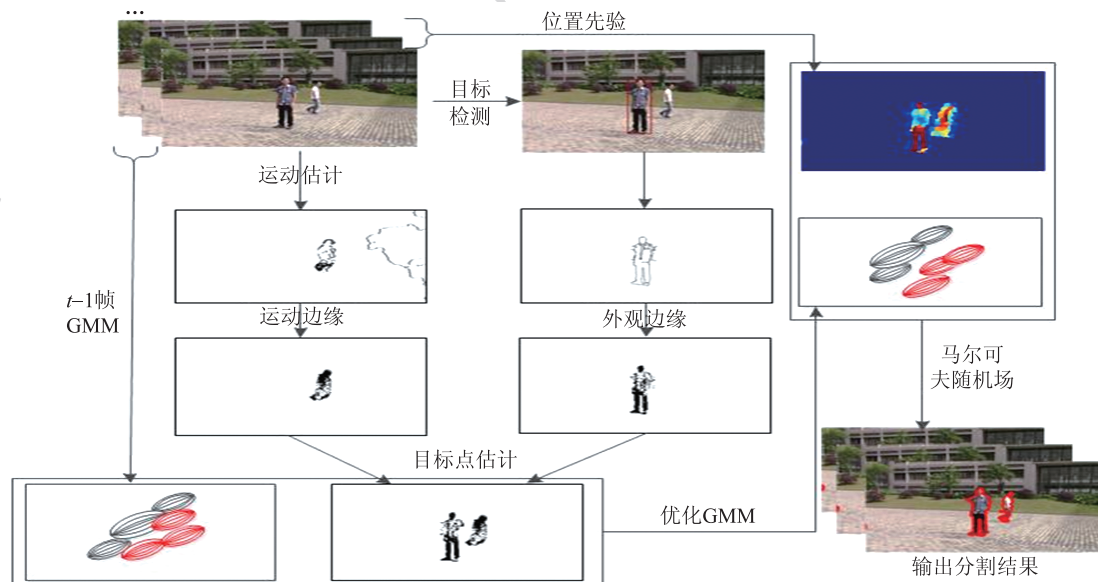


图1 融合外观和运动特征的在线目标分割算法流程图

Fig.1 Flow chart of online object segmentation with fusing appearance and motion features

1 基于外观和运动特征的在线目标分割

提出融合目标的外观特征和运动特征的在线目标分割方法,该方法先对视频帧进行目标检测,然后利用目标的外观和运动特征对目标点进行估计,得到的目标点在目标静止和运动两种情况下均具有较高的准确性,由此保证了在线分割视频目标的精度。在得到初始化的目标后,将剩余的部分作为背景,对目标和背景分别建立高斯混合模型,并结合由上一帧目标传递到该帧的位置先验建立 MRF 模型,对目标进行精确的分割。

1.1 融合外观和运动特征的目标点估计

由于视频中目标的运动复杂性以及背景复杂性,仅仅利用目标的外观特征或者运动特征均不能

达到较好的分割结果。如在目标静止或者运动较小的情况下,使用运动特征就会失败,而在背景复杂情况下,外观特征则变得不可靠。因此,融合两种特征进行目标点估计,得到较为鲁棒的初始化结果。

首先对每一帧视频图像进行目标检测,再根据目标的外观和运动特征得到外观边缘与运动边缘,进而进行外观目标点估计和运动目标点估计,然后融合外观与运动的目标点得到较为鲁棒的目标初始化结果。

1.1.1 目标检测

由于背景和噪声的干扰,如果在整幅图像中提取外观边缘,很难得到较干净的目标边缘,不利于目标点估计。如图 2 所示,本文进行目标检测能够有效去除背景的干扰,得到干净的目标边缘,进而得到准确的目标点估计。

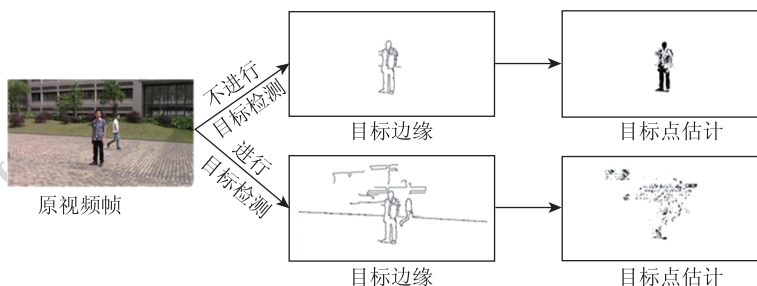


图 2 目标检测在提取外观边缘过程中的作用示意图

Fig. 2 Illustration of the effect of object detection in edge extraction

目前在静态图像中有很多的目标检测方法,并取得了较好的检测结果。Felzenszwalb 等人^[9]提出了混合多尺度可变形部件模型,并训练了大量的不同种类的目标模型,能够较准确地进行目标检测,采用这种方法进行目标检测。但此方法也存在漏检错检的情况,比如在本文实验中多人足球与车的视频中,对人、车进行目标检测出现漏检与错检,但对本文最终分割精度影响不大,具体见实验分析部分。

同时,本文使用 Canny^[10] 边缘检测算法进行边缘检测,并去除一些噪声边缘的干扰。

1.1.2 基于外观特征和运动特征的目标点估计

仅考虑利用目标的外观特征或者运动特征估计目标点,往往得不到较好的分割结果。如 Papazoglou 等人^[6]仅仅利用目标的运动特征估计出目标点,这对一些静止或运动较小的目标常常无法进行有效分割。采用融合外观和运动特征的目标点估计的方法,在运动方面,采用文献[6]的方法提取准确

的运动边缘 E^m 。在外观方面,利用第 1.1.1 中的方法得到干净的外观边缘 E^a 。

采取融合外观和运动的目标估计点,把这些目标估计点合并在一起,达到外观特征与运动特征的信息互补,得到准确的目标点估计 S ,即

$$S = f(E^a) \cup f(E^m) \quad (1)$$

式中, E^a 、 E^m 分别表示外观边缘与运动边缘, f 为目标点估计过程。

估计的目标点用于建立目标和背景的混合高斯模型。对已获取的目标边缘, Papazoglou 等人^[6]提出一种高效的目标点估计方法,此方法基于以下性质,即从一个点发出的射线,经过闭合的曲线时,如果这个点在闭合曲线内,射线经过的边缘总数为奇数。如果这个点在闭合曲线外,那么它发出的射线经过的边缘总数是偶数,如图 3 所示,其中的数字标号是射线所经过曲线边界的总数。这种目标点估计方法速度快,且对不完整复杂的边缘具有较好的鲁

棒性。本文也采用这种方法,并将此过程记为 f 。

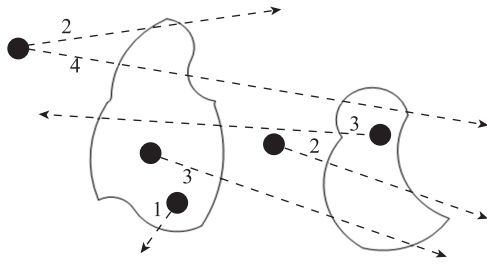


图3 多个闭合曲线中的内、外点判别示意图

Fig. 3 Illustration of distinguish the inside-outside points in many closed curve

1.2 基于 MRF 模型的目标分割

对视频的每一帧进行过分割^[11],生成互不重叠的超像素,用于以下两个方面的处理:1)优化目标估计。由于噪声边缘以及目标边缘不够完整导致估计的目标点过于稀疏,根据估计的目标点计算出每个超像素为目标概率,即目标点在超像素中所占的比例,把概率较大的超像素重新作为目标的估计,提高分割精度;2)以超像素为结点构建 MRF 模型,把目标分割问题转化为能量最小化问题,并用 graph-cut 算法^[12]进行优化求解。其中有些是采用增加空间与时序平滑项进行约束,如文献[6],而在本文的分割框架中,这些约束对分割精度影响较小,同时为了降低复杂度,提高算法效率,故没有考虑这些约束。目标函数定义为

$$\iota^* = \arg \min_{\iota} E(\iota) =$$

$$\arg \min_{\iota} \left(\sum_{i,i} A_i^t(A_i^{t-1}, S_i^t) + a \sum_{i,i} L_i^t(l_i^t) \right) \quad (2)$$

式中, $\iota \in \{L_i^t\}_{i,i}$ 表示所有超像素的标签集, A_i^t 表示第 t 帧 i 个超像素的 GMM 外观模型, S_i^t 表示 t 帧 i 个超像素的目标初始化结果, a 表示均衡因子, $L_i^t \in \{0,1\}$ 表示 t 帧第 i 个超像素标签(0,1 分别表示背景和目標), L_i^t 表示 t 帧第 i 个超像素的位置先验,它是由 $t-1$ 帧对应的超像素经过光流得到的。

1.3 在线视频分割

目前,大部分的目标分割方法^[4-7]都利用了视频的全局信息,能够使视频的每一帧都具有较好的分割结果,但具有较高的时间和空间复杂度,无法进行大量的视频处理。同时,在大多数应用场景中都需要能够在线的处理视频,本文提出的是一种能够在线分割目标的方法,较好的满足这一需求。

本文方法的主要实现步骤如下:

输入:上一帧的 GMM 模型。

1)目标初始化。

2)GMM 估计。用当前帧的目标初始结果优化上一帧的 GMM 作为当前帧外观 GMM。

3)目标分割。根据当前帧的外观 GMM 和超像素的位置先验求解式(2)。

输出:当前帧的 GMM 模型以及分割结果。

2 实验结果与分析

在公共数据集 YouTube-Objects 与自己收集的視頻集上进行实验与分析。其中 YouTube-Objects 公共数据集上包含了大量的视频帧,本文选取了车、飞机、牛、马、摩托车 5 个视频中的一个镜头进行实验。为了进一步验证本文算法的有效性以及实用性,本文也收集了 5 个视频进行实验,其中包括单个行人、两个行人、二人交谈、跑步以及多人足球等 5 个视频,并进行了人工标注。为了实验比较地客观性,在选取这些视频的时候,考虑到文献[6]只对运动目标才能够进行有效分割以及文献[4-5]只适应于单目标分割的情况,本文选取视频中有较大运动的车、飞机、摩托车、单个行人、跑步,而飞机、摩托车、马、牛、单个行人都是视频中单一的目标。总体上这些视频考虑到了目标的个数、目标运动的快慢、镜头远近以及摄像机的运动,能够较全面的评价分割方法的性能。

选取实验比较方法包括现有较流行的方法,其中包括文献[4-6]、文献[13-14]的方法。(注:在文献[14]的分割结果中,选取与手工标注最接近的作为最终分割结果进行比较。)

2.1 对比分析实验

目前大部分的目标分割的精度比较,都是在人工标注的基础上统计错误的像素点个数,即

$$error_i = \sum_{p \in F_i, p' \in GT_i} \text{XOR}(p, p') \quad (3)$$

式中, $error_i$ 表示第 i 帧错误的像素点总数, $p, p' \in \{0,1\}$ (0,1 分别表示背景和目標), F_i 和 GT_i 分别为第 i 帧的分割结果和人工标注的结果, $\text{XOR}(\cdot)$ 为异或操作。

2.1.1 定量分析

表1给出了在 YouTube-Objects 公共数据集上的实验精度比较结果。从表1中的平均错误像素点数可以看出,本文的分割结果明显优于其他 5 种分割结

果,且超过了分割精确度较高文献[5]的方法。图4表示了在本文收集的视频上每帧分割精度比较结果,从图4中可以很明显地看出,本文的分割结果在大部

分视频帧上都要优于其他方法的分割结果。除了多人足球视频和两个行人外,由于文献[4-6]方法都利用了视频全局信息,分割精度的折线整体上平缓。

表1 各种分割方法在 YouTube-Objects 数据集上的精度比较结果

Table 1 The accuracy comparison results of the proposed method with six state-of-the-art methods on YouTube-Objects dataset

方法	车 (0001-shots003)	飞机 (0002-shots014)	牛 (0016-shots002)	马 (0028-shots005)	摩托车 (0014-shots009)	平均错误 像素点数	平均提高精 度百分比/%
文献[4]	6.125 7E+03	3.181 0E+04	5.808 5E+04	3.122 3E+04	6.922 3E+03	2.683 3E+04	74.3
文献[6]	5.356 7E+03	1.258 2E+04	6.896 2E+03	3.215 1E+04	7.915 3E+03	1.298 0E+04	46.9
文献[5]	6.192 8E+03	8.485 6E+03	5.990 5E+03	3.426 7E+04	7.574 0E+03	1.250 2E+04	44.8
文献[13]	2.393 6E+04	4.946 2E+04	7.482 0E+04	3.445 7E+04	3.462 7E+04	4.346 0E+04	84.1
文献[14]	9.538 9E+03	4.164 6E+04	6.114 9E+04	3.116 7E+04	1.910 1E+04	3.252 0E+04	78.8
本文	3.389 9E+03	1.113 2E+04	5.633 3E+03	7.788 4E+03	6.551 1E+03	6.898 9E+03	—

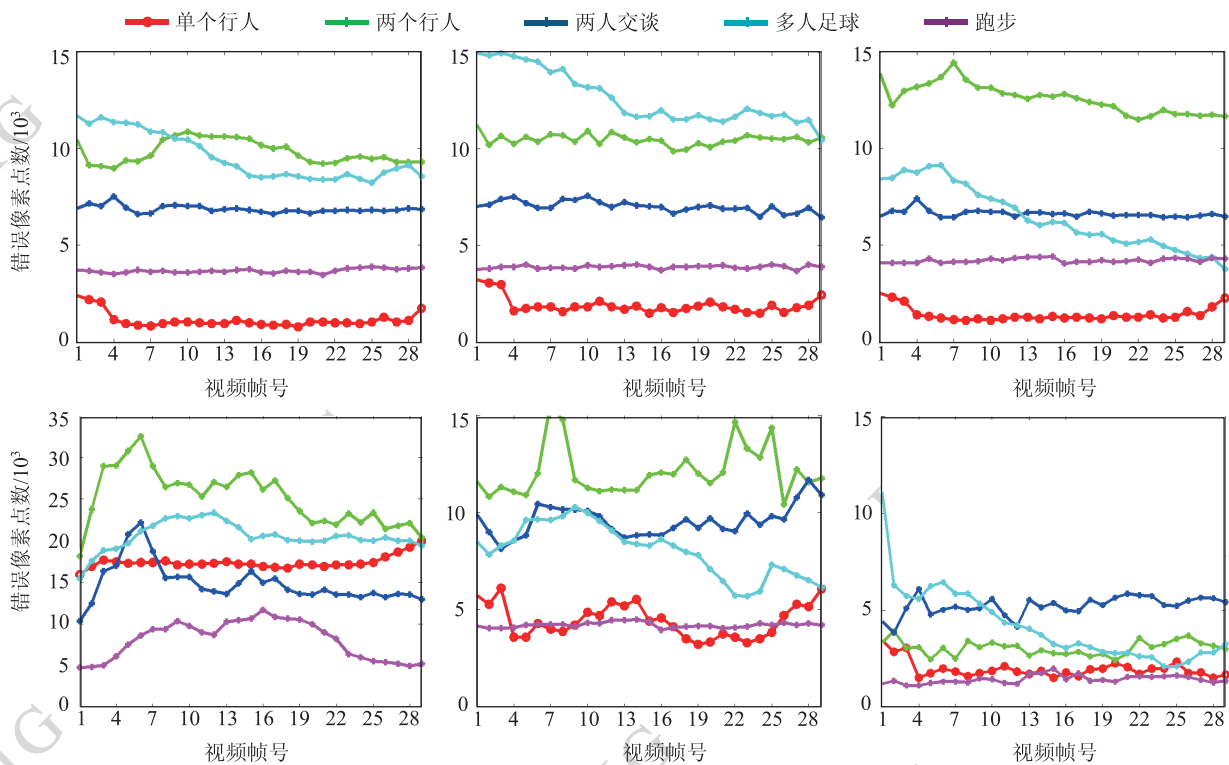


图4 每帧的错误像素点数比较图

Fig. 4 The comparison curves of mismatch points per frame

实验所用的 PC 处理器是 Pentium(R) Dual-Core CPU E6700,主频是 3.20 GHz。视频帧大小为 640 × 360 像素时,检测一帧的目标需要 7 s,得到目标外观边缘需要 0.12 s,运动边缘需要 0.5 s,平均每处理一帧所需时间约为 8.5 s(除去计算光流和过分割时间)。文献[6]处理一帧的时间大概是 1.2 s(除去计算光流和过分割时间),文献[13]处理一帧的时间大

约 10 s,略低于本文方法运行时间,但它们在很多情况下不能得到较好的分割结果。文献[4-5]不仅内存需求大,而且耗时巨大,仅每一帧产生可能的目标区域所需时间就超过 120 s,远大于本文的分割时间。

2.1.2 定性分析

图5展示了包含车、人、动物3种不同种类的视频分割结果。在这4组视频中,前3组表示多目标

在摄像机运动下的分割情况,后3组表示目标的非刚性运动的分割情况。每组视频给出了其中3帧的分割结果,其中每一组视频的第1、2、3、4、5、6行分

别表示文献[4]、文献[13]、文献[14]、文献[6]、文献[5]和本文的分割结果。从图中可以看出,本文的分割结果优于前5种方法的分割结果。



图5 4组视频中的部分帧分割结果比较图

Fig.5 The sample results of object segmentation methods on four video sequences

从4组视频第1行可以看出,文献[4]对存在多目标的多人足球视频中,出现漏分割的情况,同时对于单个大目标的牛视频中,只分割出了牛的一部分,有较大的局限性。第2行看出,文献[13]在摄像机运动情况下的分割结果很差,不能对目标进行有效分割。第3行看出,文献[14]存在错误分割目标的情况,但在多人足球的视频中分割较好。第四行可以看出,文献[6]在目标运动较快时有较好的分割结果,如在第1组视频中有较好的分割出了目标,但它无法分割出运动较小或者静止的目标,尤其是在第2组的视频中,它漏分割大部分的目标。对于时间复杂度较高、分割结果较为准确的文献[5],在第1组视频的单个车辆分割效果较好,但它无法分割视频中的多个目标。即使是在目标检测出现漏检错检,如在多人足球与车的视频中,但本文采用的

是融合外观与运动特征的在线目标分割方法,仍能够根据目标运动的特征进行初始化以及前一帧的分割结果进行目标分割。本文方法对上面存在的问题都能够有效地解决,且分割结果较好。但本文方法也存在一些问题,如在第2组视频中,目标在前几帧的分割结果较差。这是由于本文是一种在线的视频分割方法,前几帧的分割过于依赖于目标初始化结果。

2.2 组成成分分析实验

采用外观模型与位置先验这两个不同的成分建立 MRF 模型进行目标分割。为了验证这两个成分对本文分割精度的影响,这里用多人足球的视频进行组成成分分析。实验结果如图6(a)所示,其中 ourFull、our1、our2 分别代表本文方法,没有外观模型和没有位置先验的分割结果。可以看出,外观模型与位置先验均对本文分割算法精度的提高有较大作用。

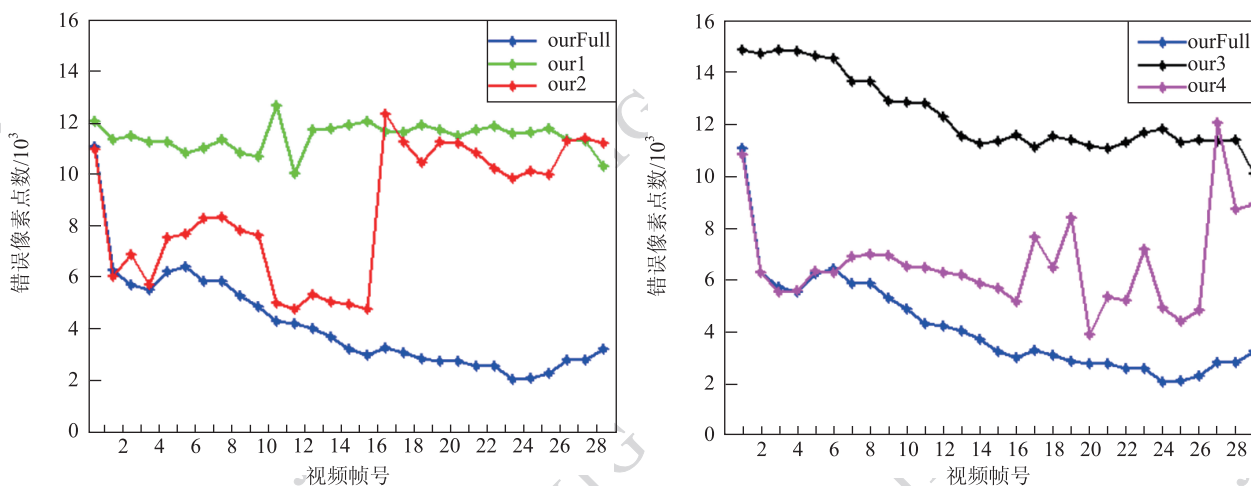


图6 不同组成成分对分割精度影响比较图

Fig. 6 The comparison curves of our components on segmentation accuracy

融合外观与运动的目标估计点对目标进行初始化,为了进一步验证这两部分对分割精度的作用,实验中分别对这两个部分进行实验,结果如6(b)所示,ourFull、our3、our4 分别代表本文方法,运动特征进行初始化,和外观特征进行目标初始化的分割结果。可以看出,仅仅考虑运动或者外观特征对目标进行初始化是不稳定的,分割结果较差,而将两者结合得到较好的目标分割结果。

3 结论

本文提出了一种融合外观和运动特征的在线目标分割方法。该方法能够在摄像机运动情况下在线

地分割出不同类型、不同运动模式下的多个目标,同时具有较高的精度。通过对实验结果的对比分析与组成成分分析,验证了该方法的有效性,其精度在大部分情况下优于其他目标分割算法。同时,本文也分析了该方法的不足,即前几帧受到初始化结果的影响较大,这也是在线视频处理方法通常会遇到的问题。由于该方法受遮挡的影响会比较大,因此,下一步的研究重点是在遮挡情况下使得分割算法具有较高的时序连续性,如以超体素为结点来考虑构图。

参考文献 (References)

- [1] Li B H, Yang G. Adaptive foreground detection approach of

- Gaussian mixture model[J]. *Journal of Image and Graphics*, 2013, 18(12): 1620-1627. [李百惠,杨庚. 混合高斯模型的自适应前景提取[J]. *中国图象图形学报*, 2013, 18(12): 1620-1627.] [DOI:10.11834/jig.20131210]
- [2] Li X C, Zhu S A. A survey of the markov random field method for image segmentation [J]. *Journal of Image and Graphics*, 2007, 12(5): 789-798. [李旭超,朱善安. 图像分割中的马尔可夫随机场方法综述[J]. *中国图象图形学报*, 2007, 12(5): 789-798] [DOI:10.11834/jig.200705129]
- [3] Zhong F, Qin X Y, Peng Q S. Tranductive segmentation of live video with non-stationary background[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA: IEEE, 2010: 2189-2196. [DOI: 10.1109/CVPR.2010.5539899]
- [4] Lee Y J, Kim J, Grauman K. Key-segments for video object segmentation[C]// *Proceedings of IEEE International Conference on Computer Vision*. Barcelona, Spain: IEEE, 2011: 1995-2002. [DOI:10.1109/ICCV.2011.6126471]
- [5] Zhang D, Javed O, Shah M. Video object segmentation through spatially accurate and temporally dense extraction of primary object regions[C]// *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*. Portland, OR, USA: IEEE, 2013: 628-635. [DOI:10.1109/CVPR.2013.87]
- [6] Papazoglou A, Ferrari V. Fast object segmentation in unconstrained video[C]// *Proceedings of IEEE International Conference on Computer Vision*. Sydney, Australia: IEEE, 2013: 1777-1784. [DOI:10.1109/ICCV.2013.223]
- [7] Jain S D, Grauman K. Supervoxel-consistent foreground propagation in video[C]// *Proceedings of European Conference on Computer Vision*. Zurich: Springer, 2014, 8692: 656-671. [DOI: 10.1007/978-3-319-10593-2_43]
- [8] Endres I, Hoiem D. Category independent object proposals [C]// *Proceedings of European Conference on Computer Vision*. Berlin: Springer, 2010, 635: 575-588. [DOI: 10.1007/978-3-642-15555-0_42]
- [9] Felzenszwalb P F, Girshick R B, McAllester D, et al. Object detection with discriminatively trained part based models [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, 32(9): 1627-1645 [DOI:10.1109/TPAMI.2009.167]
- [10] Canny J. A computational approach to edge detection[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(6): 679-698. [DOI: 10.1109/TPAMI.1986.4767851]
- [11] Levinstein A, Stere A, Kutulakos K N, et al. TurboPixels: fast superpixels using geometric flows[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, 31(12): 2290-2297. [DOI:10.1109/TPAMI.2009.96]
- [12] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2001, 23(11): 1222-1239. [DOI: 10.1109/34.969114]
- [13] Barnich O, Van D M. Vibe: a universal background subtraction algorithm for video sequences[J]. *IEEE Transactions on Image Processing*, 2011, 20(6): 1709-1724. [DOI: 10.1109/TIP.2010.2101613]
- [14] Brox T, Malik J. Object segmentation by long term analysis of point trajectories[C]// *Proceedings of European Conference on Computer Vision*. Berlin: Springer, 2010, 6315: 282-295. [DOI: 10.1007/978-3-642-15555-0_21]