



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

2013
12
VOL.18

ISSN1006-8961
CN11-3758/TB



图像分割 P1610



第18卷第12期（总第212期）

2013年12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

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

2013 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@irsa.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@irsa.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

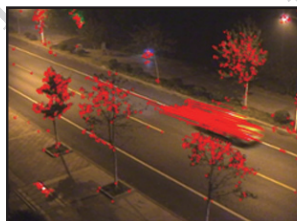
ISSN 1006-8961

CODEN ZTTXFZ

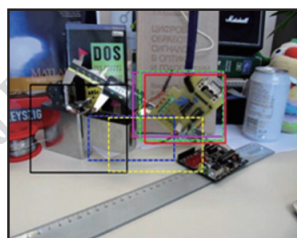
国外发行代号 M1406

国内邮发代号 82-831

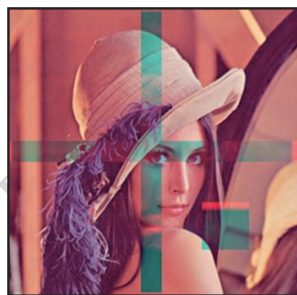
国内定价 50.00元



稀疏光流快速计算的动态目标
检测与跟踪(第1593页)



字典学习联合粒子滤波鲁棒跟踪
(第1628页)



结合SURF特征的多功能彩色
图像水印算法(第1694页)

图像处理和编码

小波子带最优方向性选择和压缩感知的图像编解码

王相海, 程露露, 周夏, 宋传鸣

1551

基于位平面和HVS的信息隐藏算法

张海涛, 姚雪, 陈虹宇, 张晔

1559

混沌密钥调制DFRFT旋转因子的视频加密

金建国, 王乐, 魏明军, 苏晶晶

1567

图像分析和识别

PCA-NLM的纺织品缺陷检测

杨学志, 左海琴, 陈远, 吴克伟, 谢昭

1574

基于主动表观模型姿态矫正和局部加权匹配人脸识别

赵恒, 俞鹏

1582

指纹拓扑模式构建及其在识别中的应用

仲伟波, 吕园, 李敏敏

1587

稀疏光流快速计算的动态目标检测与跟踪

陈添丁, 胡鉴, 吴涛

1593

边缘与灰度检测相结合的场景图像文本定位

何立强, 刘浩, 陈永

1601

Mean Shift图像分割算法的并行化

李宏益, 吴素萍

1610

混合高斯模型的自适应前景提取

李百惠, 杨庚

1620

字典学习联合粒子滤波鲁棒跟踪

查绎, 曹铁勇, 黄辉, 潘竟峰, 尤峻

1628

图像理解和计算机视觉

仿生复眼式全景探测及跟踪策略

罗超, 赵美蓉, 王东, 曹克雄

1637

医学图像处理

利用感兴趣区域从脑部MRI中提取脑组织

江少锋, 万红平, 陈震, 杨素华

1644

“第十届全国智能CAD与数字娱乐会议”专栏

基于Hermite插值的网格拼接和融合

缪永伟, 林海斌, 寿华好

1651

计算机生成中国传统祥云动画

方建文, 彭初, 于金辉

1660

森林动态演替现象的可视化模拟

单梁, 杨刚, 黄心渊

1666

多特征证据理论融合的视觉词典构建

沈项军, 高海迪, 曾兰玲, 毛启容, 詹永照

1676

轨迹和多标签超图配对融合的视频复杂事件检测

柯佳, 詹永照, 陈潇君, 汪满容

1684

结合SURF特征的多功能彩色图像水印算法

朱光, 师文, 朱学芳, 郝世博

1694

采用邻域直方图匹配的矢量纹理图案合成

徐婵婵, 杨刚

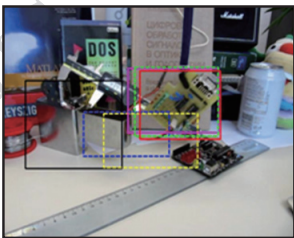
1703

CONTENTS

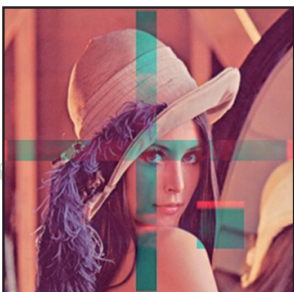
JOURNAL OF IMAGE AND GRAPHICS



Dynamic target detection and tracking based on fast computation using sparse optical flow (P1593)



Object tracking based on learning a dictionary joint practical filter(P1628)



Multi-purpose watermark algorithm of color images based on SURF(P1694)

Image Processing and Coding

Image codec research on the optimum directional selection of wavelet sub-band and compressed sensing

Wang Xianghai, Cheng Lulu, Zhou Xia, Song Chuanming 1551

Research of information hiding algorithm based on bit-plane and HVS

Zhang Haitao, Yao Xue, Chen Hongyu, Zhang Ye 1559

Video encryption based on chaotic key modulating DFRFT rotation factor

Jin Jianguo, Wang Le, Wei Mingjun, Su Jingjing 1567

Image Analysis and Recognition

Fabric defect detection based on PCA-NLM

Yang Xuezhi, Zuo Haiqin, Chen Yuan, Wu Kewei, Xie Zhao 1574

AAM-based alignment and local weighted matching method for face recognition

Zhao Heng, Yu Peng 1582

Fingerprint topological pattern construction and its application in recognition

Zhong Weibo, Lyu Yuan, Li Minmin 1587

Dynamic target detection and tracking based on fast computation using sparse optical flow

Chen Tianding, Hu Jian, Wu Di 1593

Text extraction from scene image based on edge and gray-scale detection collaboration

He Liqiang, Liu Hao, Chen Yong 1601

Parallelization of Mean Shift image segmentation algorithm

Li Hongyi, Wu Suping 1610

Adaptive foreground detection approach of Gaussian mixture model

Li Baihui, Yang Geng 1620

Object tracking based on learning a dictionary joint practical filter

Zha Yi, Cao Tiejong, Huang Hui, Pan Jingfeng, You Jun 1628

Image Understanding and Computer Vision

Bionic compound eye panoramic detection and tracking strategy

Luo Chao, Zhao Meirong, Wang Dong, Cao Kexiong 1637

Medical Image Processing

Automatic brain extraction from cerebral MRI volume using region of interest

Jiang Shaofeng, Wan Hongping, Chen Zhen, Yang Suhua 1644

Issum of CIDE' 2013

Mesh stitching and fusion based on Hermite interpolation scheme

Miao Yongwei, Lin Haibin, Shou Huahao 1651

Computer generation of cloud animation with a Chinese traditional style

Fang Jianwen, Peng Ren, Yu Jinhui 1660

Simulation of the dynamic evolution of spatial distribution of trees in the forest

Shan Liang, Yang Gang, Huang Xinyuan 1666

Visual dictionary construction method based on Dempster-Shafer evidence theory

Shen Xiangjun, Gao Haidi, Zeng Lanling, Mao Qirong, Zhan Yongzhao 1676

The video complex event detection method of matching integration between trajectory and multi-label hypergraphs

Ke Jia, Zhan Yongzhao, Chen Xiaojun, Wang Manrong 1684

Multi-purpose watermark algorithm of color images based on SURF

Zhu Guang, Shi Wen, Zhu Xuefang, Hao Shibao 1694

Vector texture pattern synthesis based on neighborhood histogram matching

Xu Chanchan, Yang Gang 1703

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2013)12-1684-10

论文引用格式: 柯佳, 詹永照, 陈潇君, 汪满容. 轨迹和多标签超图配对融合的视频复杂事件检测[J]. 中国图象图形学报, 2013, 18(12): 1684-1693. [DOI:10.11834/jig.20131218]

轨迹和多标签超图配对融合的视频复杂事件检测

柯佳^{1,2}, 詹永照², 陈潇君³, 汪满容⁴

1. 江苏大学管理学院, 镇江 212013; 2. 江苏大学计算机学院, 镇江 212013;
3. 江苏大学附属医院信息科, 镇江 212013; 4. 江苏大学图书馆, 镇江 212013

摘要: 视频数据中包含丰富的运动事件信息, 从中检测复杂事件, 分析其中的高层语义信息, 已成为视频研究领域的热点之一。视频复杂事件检测, 主要是对事件中多语义概念进行检测分析, 对多运动目标的特征进行描述, 发现底层特征与高层语义概念间的关系, 旨在从各类视频特征及相关的原始视频数据中自动提取视频复杂事件中语义概念模式, 实现“跨越语义鸿沟”的目标。在超图理论的基础上, 提出了针对运动目标特征分别构建轨迹超图和多标签超图, 并对其进行配对融合, 用于检测视频复杂事件。实验结果表明, 同其他方法如基于普通图的事件检测方法和基于超图的多标签半监督学习方法相比, 该方法在检测复杂事件结果中具有更高的平均查准率和平均查全率。

关键词: 超图; 轨迹; 多标签; 配对融合; 谱分割

The video complex event detection method of matching integration between trajectory and multi-label hypergraphs

Ke Jia^{1,2}, Zhan Yongzhao², Chen Xiaojun³, Wang Manrong⁴

1. School of management, Jiangsu University, Zhenjiang 212013, China;
2. School of Computer Science and Telecommunication Engineering, Jiangsu University, Zhenjiang 212013, China;
3. Affiliated Hospital of Jiangsu University, Zhenjiang 212013, China;
4. Jiangsu University Library Zhenjiang 212013, China

Abstract: With the rapid development of the applications of network, enormous multimedia files are emerging every day. Video data is an integration of text, sound, image and other files. Not only being hierarchical, structural and complex, video data is also rich in semantic information. Therefore, extensive attentions have been drawn to how to process video data quickly, extract the video characteristics accurately, and analyze and understand the semantic content deeply. Semantic event detection and analysis could find video information quickly and accurately for users in the vast ocean of video data. It can be applied to the field of video on demand, intelligent monitoring and video mining as well. However, there are still many limitations, such as low recognition rate for multiple moving objects with different characteristics, low accuracy in semantic event detection, difficulties in detecting semantic event correlations, the lack of consistent standards of event semantic description, and so on. The detection and analysis methods of complex event based on matching integration between trajectory and multi-label hypergraphs are proposed. Trajectory and multi-label hypergraphs are constructed for classifying and recognizing the complex events. By matching the trajectory hypergraph and multi-label hypergraph, mapping relationship

收稿日期: 2013-04-15; 修回日期: 2013-07-14

基金项目: 国家自然科学基金项目(61170126); 江苏省 2008 年度普通高校研究生科研创新计划基金项目(CX08B_097Z); 镇江市科技支撑计划(工业部分)基金项目(GY2010023); 教育部人文社会科学研究青年基金项目(12YJCZH020)

第一作者简介: 柯佳(1981—), 女, 讲师, 2013 年于江苏大学获计算机应用技术专业博士学位, 主要研究领域为多媒体技术。E-mail: kejia@ujs.edu.cn

between trajectory and multiple semantic labels is built to extract the complex semantic events. The recognition of low-level features to high-level semantic is made possible for video events. Compared with other methods, such as event detection method based on graph and multi-label semi-supervised learning methods based on hypergraph, the proposed method has a higher mean average recall rate and mean average accuracy rate in the detection result of the complex event. we propose a new event detection method that is trajectory and multi-label hypergraphs model in this paper. This model is a widely-used detection and analysis of complex semantic-based events method. In their clustering process, their number of vertices and clusters has been more than other graph methods. But, the event detection result is better than others.

Key words: hypergraph; trajectory; label multi-label; pairing fusion; spectrum hypergraph cut

0 引言

近年来在视频领域,对运动事件检测分析已成为视频处理过程中的一个重要研究趋势。在视频层次结构中,事件分析一般是在场景级(scene)进行分析,可分为基于关键帧(frame)的事件分析和基于镜头(shot)的事件分析两种,如体育视频分析中,Lu 等人^[1]采用 LDA 方法对篮球比赛中场景进行分类,获得比赛事件。Chen 等人^[2]使用 HMM (hidden Markov model) 来检测篮球比赛事件。还有支持向量机(SVM)^[3]、贝叶斯网络(BN)^[4]、动态贝叶斯(dynamic Bayesian)^[5]、有限自动机(finite state machine)等方法^[6]也用来分析篮球运动轨迹,以对体育视频中的进球事件进行检测。在监控视频分析中,Kalman 等人^[7]使用 Kalman 滤波对多运动目标的轨迹进行跟踪,从中检测出多运动目标事件。Maggiot 等人^[8]提出使用 Meanshift 方法对运动目标进行跟踪发现运动事件。运动信息作为能表示视频内容独有的发展变化特征,对描述理解视频内容具有相对重要的作用。由事件中对象的运动行为特征,可将事件分为简单事件和复杂事件^[9],本文处理的对象是由多运动目标构成的复杂事件。

检测视频事件,并对其高层语义进行分析,是最近几年视频处理领域的研究热点。在基于内容的视频检索中,有人引入了图分割理论中的超图(hypergraph)来对视频进行处理。超图的概念最早由 Berge 于 1973 年提出,通过对普通图在拓扑结构上的泛化,可描述最具一般性的离散结构关系,已经应用在数据库设计、网络分析、运筹学等领域。并有人将该方法引入到了多媒体信息处理中,如 Liu 等人^[10]采用软超图方法进行图像分割,Ding 等人^[11]提出了概率超图方法对图像中的低层特征进行分类以实现图像分割。

在视频事件检测研究领域,如何检测复杂运

动事件并标注出事件,一直是研究人员的重点工作之一。Huang 等人^[12]使用超图分割视频中的运动目标,同时,Hakeem 等人^[13]提出基于普通图分割的方法检测复杂事件,根据事件之间的时序关系对简单事件关联性采用图分割方法进行聚类。在视频中的语义标签检测研究中,Han 等人^[14]提出多超图方法对视频中多语义概念进行检测。姜远等人^[15]提出一种基于随机游走的多标记文档分类方法,可用未标记文档来提高文档分类性能。

旨在对视频复杂事件进行检测,本文根据复杂事件的时空特性,选取了运动目标的轨迹作为分析和理解运动事件的特征,建立了轨迹超图,原因是轨迹体现了物体运动规律,反映了目标的行为目的;同时,为视频中含有的语义概念,建立了多标签语义超图,并采用谱分割聚类方法分别对两个超图进行分割。得到轨迹类和标签类。最后,对两超图进行了配对融合,获得两个超图顶点间关联关系,以检测视频复杂事件,并进行实验分析。

1 超图相关理论

超图的理论基础是图论和集合论。具有共同属性特征的对象属于一个集合,不同的抽象层次可归属于集合的集合;如此构成以集合的包含关系为基础的结构,这种结构可用超图来表示。

1.1 超图的定义

超图 G 中, $V = \{v_1, v_2, \dots, v_n\}$ 是 n 个顶点的有限集合, $E = \{e_1, e_2, \dots, e_m\}$ 是 m 个超边 e 的有限集合。 G 可用 $|V| \times |E|$ 的关联矩阵 H 表示。 E 具有权值集 $W = \{w(e_1), w(e_2), \dots, w(e_m)\}$, 其中每条超边 e_j 都有一个正的权值 $w(e_j)$ 相关联, $w(e_j)$ 的大小反应了超边内顶点之间的聚合程度。

顶点 $v_i \in V$ 的度 $d(v_i) = \sum_{\{e_j \in E | v_i \in e_j\}} w(e_j) = \sum_{e_j \in E} w(e_j) h(v_i, e_j)$, 其中超边 $e_j \in E$ 的度 $\delta(e_j) =$

$|e_j| = \sum_{v_i \in V} h(v_i, e_j)$ 。设 D_V 表示顶点度的对角阵、 D_e 表示超边度的对角阵、 W 表示超边权值的对角阵, 则超图 G 的邻接矩阵 A 可表示为 $A = HWH^T - D_V$, 其中 H 为超图 G 的矩阵表示。

本文建立了轨迹超图 G_T 和多标签超图 G_{ML} 。其中, 在轨迹超图 G_T 中, 顶点为运动目标的每一条轨迹, 根据每条轨迹的空间和速度的特征向量得到轨迹的相似性度量, 再计算出的顶点集 V 的亲矩阵 F , 以得到超边集合。在多标签超图 G_{ML} 中, 每个顶点对应于一类语义标签, 将语义标签在视频集中出现的共现性 (co-occurrence) 作为计算亲和矩阵 F 的依据, 进行超边集合的计算。

1.2 超图的分割

设 $S \subset V$ 表示超图顶点的子集、 S^c 为 S 的补集, 则 $S \cap S^c = \emptyset$, 对超图的一次分割就是将顶点集 V 分成 S 和 S^c 两部分。对于给定的顶点子集 $S \subset V$, 定义 S 的超割边集 ∂S 为将超图分割为 S 和 S^c 两部分后由所有被切割的超边所构成的一个集合, 则 $\partial S = \{e \in E | e \cap S \neq \emptyset, e \cap S^c \neq \emptyset\}$ 。对于给定的顶点子集 $S \subset V$, 定义 S 的卷量 $vol(S) = \sum_{v \in S} d(v)$, S^c 的卷量 $vol(S^c) = \sum_{v \in S^c} d(v)$ 。对于给定的顶点子集 $S \subset V$, 定义超割边集 ∂S 的卷量为

$$vol(\partial S) = \sum_{e \in \partial S} w(e) \frac{|e \cap S| |e \cap S^c|}{\delta(e)} \quad (1)$$

式中, $vol(\partial S) = vol(\partial S^c)$, 式(1)表示超图中每条超边独立存在, 超边内的顶点构成一个全连通子图, 该子图的边为子边, 然后分配相同的权值 $\frac{w(e)}{\delta(e)}$ 给每条子边, 当一个超边被切割时就意味着有 $|e \cap S| |e \cap S^c|$ 个子边被切割, 式(1)右边为所有被切割的子边权值之和。

分割后在同一簇的顶点间关联是最密集的, 而在两个簇间顶点的关联最小, 表示为

$$\arg \min_{\emptyset \neq S \subset V} c(S) = vol(\partial S) \left(\frac{1}{vol(S)} + \frac{1}{vol(S^c)} \right) \quad (2)$$

本文使用了谱分割的方法对轨迹超图和多标签超图进行分割, 分别得到视频事件按照轨迹时空属性和语义标签的语义共现性的分割结果; 根据事件之间的时序依赖关系的配对融合结果, 得到复杂事件的分类结果, 从而检测出复杂事件。

2 轨迹超图和多标签超图的定义

2.1 轨迹定义

随着时间参数变化, 目标运动表现为在连续视频帧上的一系列空间变化, 因此将轨迹定义为同时表现时空变化的形式。我们在文献[16]中提出了改进混合高斯模型的方法识别运动目标并采用不变轮廓矩的方法用于描述连续帧中的多运动目标^[17], 选取运动目标的质心坐标作为对应于第 i 个运动目标 (o_i) 的空间位置 (x_i, y_i), 从目标运动的连续空间位置变化规律中, 提取轨迹。

定义 1 在一段视频中, 令第 i 个运动目标 o_i 的运动轨迹为 T_{o_i} 表示为三元组 (x_i, y_i, t_i), 可解释为时间点 t_i 时运动目标 o_i 的空间位置关系为 (x_i, y_i), x_i 为水平坐标值, y_i 为垂直坐标值。令目标运动空间, 空域, 表示为 R^2 , 时间域表示为 R , 则时空域可以表示为 $R \times R^2$ 。

定义 2 轨迹 T_s 在 t 时刻 ($t_i \leq t \leq t_{i+1}$) 的速度矢量为 $(p_x, p_y) = \left(\frac{x_{i+1} - x_i}{t_{i+1} - t_i}, \frac{y_{i+1} - y_i}{t_{i+1} - t_i} \right)$, 在点 (x_i, y_i) 和 (x_{i+1}, y_{i+1}) 之间的距离上目标的运动速度可以表示为 $\frac{\sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2}}{t_{i+1} - t_i}$ 。

2.2 轨迹的相似性度量

考虑到对不同长度的多运动目标轨迹进行比较, 选择 Hausdroff 距离对轨迹的空间距离进行相似度计算。从视频特征分析中发现, 轨迹的运动特征也是运动目标的主要特征之一, 因此, 采用空间距离和运动特征结合的方法对轨迹的相似性进行度量, 运动特征选取轨迹上运动点的速度方向。

在轨迹集合中选取任意两条轨迹分别为 T_i 和 T_j , 它们的空间距离为 $d_{sp}(T_i, T_j)$, 可表示为

$$d_{sp}(T_i, T_j) = \min \{ d(T_i, T_j), d(T_j, T_i) \} \quad (3)$$

式中, $d(T_i, T_j) = \max_{a \in T_i} \min_{b \in T_j} \|a - b\|$ 。 T_i 和 T_j 越相似, 则它们之间的空间距离就越小, 反之, 空间就越大。

同时, 轨迹除了可用空间距离度量外, 轨迹的速度方向性也成为了轨迹相似性之间的度量指标^[18]。令 a_p 为 T_i 上任一点, a_p 的速度矢量为 $(p_x^{a_p}, p_y^{a_p})$; b_q 为 T_j 上任一点, b_q 的速度矢量为 $(p_x^{b_q}, p_y^{b_q})$, 则速度

方向定义为

$$\cos(\theta) = \frac{(p_x^{a_p}, p_y^{a_p}) \cdot (p_x^{b_q}, p_y^{b_q})}{|(p_x^{a_p}, p_y^{a_p})| \cdot |(p_x^{b_q}, p_y^{b_q})|} \quad (4)$$

轨迹的速度方向距离为

$$d_{ve}(T_i, T_j) = 1 - \cos(\theta) \quad (5)$$

则结合轨迹空间距离和速度方向距离的轨迹相似性可表示为

$$D(T_i, T_j) = \frac{1}{n} \times$$

$$\sum_{p \in T_i} [kgd_{sp}(T_i, T_j) + (1 - k)d_{ve}(T_i, T_j)] \quad (6)$$

n 表示 T_i 中所有点的个数, k 为调整空间距离和速度方向距离对轨迹影响的贡献度系数, 本文取 0.5, 将空间距离和速度方向距离对轨迹相似度影响作等值对半处理, 之后采用超图方法进行聚类分析。

2.3 建立轨迹超图

基于上述超图理论, 对多运动目标轨迹进行轨迹超图的构建, 为某一视频集中的每条轨迹建立一个顶点, 假设某轨迹集合共包含 n 条轨迹为 $\{T_1, \dots, T_i, \dots, T_n\}$, T_i 和 T_j 为其中任意两条轨迹。在得到轨迹超图的顶点集 $V = \{v_1, v_2, \dots, v_n\}$ 且 $|V| = n$ 后, 由每条轨迹的空间和速度的特征向量得到轨迹的相似性度量, 再计算出顶点集 V 的 $|V| \times |V|$ 亲和矩阵 F ; 最后依次将每个顶点作为质心顶点并由该顶点和它的 $m-1$ 个最亲和顶点(由亲和矩阵 F 可求得)一起形成一条包含 m 个顶点的超边 e , 得到轨迹超图的超边集 $E = \{e_1, e_2, \dots, e_m\}$, 完成一个多运动目标的轨迹超图构建。

构建轨迹超图 G_T 的步骤如下:

输入 某段视频中提取到的轨迹集合 $\{T_1, \dots, T_i, \dots, T_n\}$ 。

输出 轨迹超图 G_T 。

1) 确定轨迹超图 G_T 的顶点集

将每条轨迹视为一个顶点, 得到 G_T 的顶点集 $V_i = \{v_1, v_2, \dots, v_n\}$ 。

2) 确定 G_T 的超边集 $E_i = \{e_1, e_2, \dots, e_m\}$

(1) 令顶点 v_i 和顶点 v_j 表示 T_i 和 T_j , 两者之间的亲和度为 F_{ij} , 表示为 $F_{ij} = \exp\left(-\frac{D(T_i, T_j)}{\sigma}\right)$ 。 σ 为 $D(T_i, T_j)$ 计算得到的标准差;

(2) 根据第 1) 步计算所得的顶点集 V_i 的 $|V| \times |V|$ 亲和矩阵 F , 构造超边, 依次视每个顶点为质心顶点并由质心顶点和它的 $m-1$ 个最亲和顶

点一起形成一条包含 m 个顶点的超边 e 。最终得到轨迹超图 G_T 超边集 $E_T = \{e_1, e_2, \dots, e_m\}$ 。

由以上轨迹超图 G_T 的构建过程可知, 多运动目标轨迹在空间距离和速度方向距离上的相似关系将影响轨迹超图 G_T 中顶点间的亲和度。

在此基础上, 分别计算 G_T 的关联矩阵 H_i , 超边权值对角阵 W , 表示为 $w_{ij}(e_i) = \sum_{v_j \in e_i} F_{ij}$ 、顶点度的对角阵 D_v 、超边度的对角阵 D_e 。

使用 4 种超图谱分割方法对轨迹超图 G_T 递归地进行二分, 直到 $c(S)$ 的值以及分割所得各部分内超边的平均权值均不小于相应的阈值。通过这种方式, 可以获取轨迹超图中的轨迹簇, 记为 $TC = \{TC_1, \dots, TC_i, \dots, TC_m\}$ 。由谱分割特性可知, 同一个轨迹簇簇中的各条轨迹间的关联程度很大, 不同轨迹簇中的轨迹之间的关联程度很小。

2.4 构建多标签超图

运动目标的单条轨迹中包含多个子事件, 可采用文献[19]中的方法为事件标注语义概念标签。在视频语义理解领域研究发现, 真实视频中同一场景、同一镜头或同一视频帧中存在多种语义概念, 如图 1 所示分别为 CAVIAR 数据库中一段场景的两帧, 其内容可标注为 2 人行走、交谈、商店、走廊和购物等多个语义标签。



(a) 行人1



(b) 行人2

图1 CAVIAR 行人购物视频帧

Fig. 1 Pedestrian shopping video frame in CAVIAR

可见在一段视频中的多运动目标的多条轨迹,对应着多语义概念标签。多媒体文件中的多标签学习源于中文信息处理中的语义概念分类与学习^[15]。本文建立多标签超图对轨迹中对应的多语义概念标签进行分类与学习。

令多标签超图为 $G_{ML} = (V_{ml}, E_{ml})$, 每个顶点对应于一类语义标签, 用中文信息处理中计算关键词的相似度方法, 对语义概念先进行分类, 得到类别: 令语义标签 l_p 和类别 l_q 之间的相似度为 $D(l_p \cdot$

$l_q)$, $D(l_p \cdot l_q) = \cos(l_p \cdot l_q) = \frac{l_p \cdot l_q}{|l_p| \cdot |l_q|}$ 。每条超边 E_{ml} 包含了 d 个顶点, 可表示为 $E \subseteq V_{ml}^d$, 超边记为 $e_i \in E$, 其中 $e_i = (v_{i1}, \dots, v_{id})$, $v_{ip} \in V$ 。计算顶点集 V_{ml} 的 $|V_{ml}| \times |V_{ml}|$ 亲和矩阵 F 。

为描述每个顶点与其他顶点在视频集中两两出现的概率, 使用共现性^[20]的累加和表示。将 l_p 和 l_q 的共现性记为 $O_{l_p l_q}$, 则有

$$O_{l_p l_q} = o(l_p, l_q) + o(l_q, l_p) \quad (7)$$

$$\text{式中, } o(l_p, l_q) = \frac{\sum_{i=1}^N V_i^p V_i^q}{\sum_{i=1}^N V_i^p}, p, q \in \{1, 2, \dots, P\}, N$$

为轨迹集合大小。对于任意 l_p , V^p 表示每条轨迹中对于是否含有标签 l_p 的向量, 若有, V_i^p 为 1, 反之则为 0。若 $O_{l_p l_q}$ 大于阈值 η , 则认为语义标签 l_p 和 l_q 具有共现性。超边的权值 $w_{ij}(e_i) = \sum_{v_j \in e_i} F_{ij}$ 表示每条超边内的顶点包含的语义标签都两两计算其共现性。

顶点集 V_{ml} 的 $|V_{ml}| \times |V_{ml}|$ 亲和矩阵 F , 记为

$$F = \begin{cases} \exp\left(-\frac{D(l_p \cdot l_q)}{\sigma'}\right) & O_{l_p l_q} \geq \eta \\ 0 & O_{l_p l_q} < \eta \end{cases} \quad (8)$$

式中, $D(l_p \cdot l_q) = O_{l_p l_q}$, σ' 为 $D(l_p \cdot l_q)$ 计算得到的标准差。

通过和 2.3 节中建立轨迹超图 G_T 相同的算法步骤来构建多标签超图 G_{ML} , 其中算法输入为某段视频中提取到的语义标签集合 $\{l_1, \dots, l_i, \dots, l_n\}$, 输出为多标签超图 G_{ML} 。分别计算 G_T 的关联矩阵 H_l , 超边权值对角阵 W , 表示为 $w_{ij}(e_i) = \sum_{v_j \in e_i} F_{ij}$, 顶点度的对角阵 D_v , 超边度的对角阵 D_e 。

使用第 3 节超图谱分割方法对多标签超图 G_{ML} 递归地进行二分, 将 G_{ML} 中的标签类别按照各标签

在视频中出现的共现性进行分类, 从中获得不同概念意义标签之间的共现关联性。

2.5 轨迹和多标签超图的配对融合

若某视频集有多个运动目标形成的 m 条轨迹, 单条轨迹中 T_i 包含 n 个语义标签, 定义轨迹多标签矩阵 ML_{mn} , 用以存储每条轨迹中对应的多个语义标签, 即

$$\begin{matrix} & l_0 & l_1 & l_2 & l_3 \\ \begin{matrix} T_0 \\ T_1 \\ T_2 \\ T_3 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}$$

矩阵中显示 T_1 中有语义概念标签 (l_0, l_1, l_2) , 可表示为在轨迹超图 G_T 中存在顶点 v_i 与多标签超图 G_{ML} 中顶点 $\{v'_j, v'_k, v'_l\}$ 具有联系。如图 2 所示, 可知 G_T 中顶点 v_1, v_2, v_3 分别对应于 G_{ML} 中的顶点集、 $\{v'_1, v'_3, v'_4\}$, $\{v'_2, v'_3\}$, $\{v'_1, v'_2, v'_3, v'_4\}$ 。

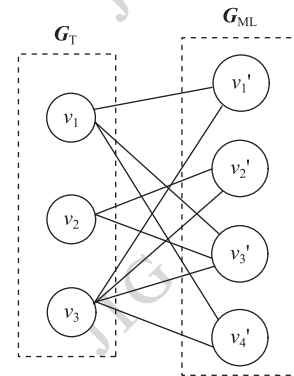


图 2 轨迹超图与多标签超图的顶点配对融合表示

Fig. 2 Representation of matching integration vertices between trajectory and multi-label hypergraphs

研究发现, 语义标签与轨迹之间的配对联系表现为, 在轨迹中包含的多个语义标签所体现的是运动事件语义, 根据事件发生的时序关系特点, 语义标签不仅具有共现性特点, 还体现出时序条件依赖关系。共现性表示在一段视频场景中的多语义标签间联系, 时序条件依赖关系表示在一条轨迹中多语义标签之间的时序关系。因此, 将这种时序条件依赖关系表示为 G_{ML} 中多语义标签在轨迹所对应的事件中依次出现的顺序关联程度, 即在各语义标签表示的事件, 依次发生的条件下某一语义发生的条件概率, 在超图中表示为 G_T 中顶点 v_i 对应于 G_{ML} 中的 v'_i

之后, v'_3 发生的概率。

$$P(v'_i(t) | v'_j(t-1)) = \frac{P(v'_i(t), v'_j(t-1))}{P(v'_j(t-1))} \quad (9)$$

$$\begin{aligned} P(v'_i(t) | v'_j(t-1), v'_k(t-2)) &= \\ \frac{P(v'_i(t), v'_j(t-1), v'_k(t-2))}{P(v'_i(t-1) | v'_k(t-2)) P(v'_k(t-2))} & \quad (10) \\ P(v'_i(t) | v'_j(t-1), v'_k(t-2), v'_l(t-3)) &= \\ \frac{P(v'_i(t), v'_j(t-1), v'_k(t-2), v'_l(t-3))}{P(v'_j(t-1) | v'_k(t-2), v'_l(t-3)) P(v'_k(t-2) | v'_l(t-3)) P(v'_l(t-3))} & \quad (11) \end{aligned}$$

式中, $v'_i(t)$ 为顶点 v'_i 在轨迹中 t 时刻出现的对应于 G_{ML} 中语义标签, $v'_l(t-3)$ 、 $v'_k(t-2)$ 、 $v'_j(t-1)$ 依次为在轨迹中 $t-3$ 、 $t-2$ 、 $t-1$ 时刻出现的 G_{ML} 中语义标签。由此方法可以得到轨迹超图与多语义超图之间的映射关系, 从而将视频集中的轨迹信息转化为多语义标签的表示, 检测出事件语义, 实现了从底层的时空特征到高层语义特征的过渡, 即表现为了运动目标的运动事件描述。

3 轨迹多标签超图的谱分割

正如文献[19]所述, 由式(2)可转化为实值最优化问题, 即

$$\operatorname{argmin}_{f \in \mathbf{R}^{|V|}} \frac{1}{2} \sum_{e \in E} \sum_{u, v \in e} \frac{w(e)}{\delta(e)} \left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}} \right)^2 \quad (12)$$

满足于 $\sum_{v \in V} f^2(v) = 1$, $\sum_{v \in V} f(v) \sqrt{d(v)} = 0$ 。

定义矩阵 $\boldsymbol{\theta} = \mathbf{D}_v^{-1/2} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-1/2}$, $\boldsymbol{\Delta} = \mathbf{I} - \boldsymbol{\theta} = \mathbf{I} - \mathbf{D}_v^{-1/2} \mathbf{H} \mathbf{W} \mathbf{D}_e^{-1} \mathbf{H}^T \mathbf{D}_v^{-1/2}$, 其中, $\mathbf{I}$ 为单位矩阵, $\boldsymbol{\Delta}$ 被称为超图的拉普拉斯矩阵。经计算可证明

$$\sum_{e \in E} \sum_{u, v \in e} \frac{w(e)}{\delta(e)} \left(\frac{f(u)}{\sqrt{d(u)}} - \frac{f(v)}{\sqrt{d(v)}} \right)^2 = 2 \mathbf{f}^T \boldsymbol{\Delta} \mathbf{f} \quad (13)$$

上述最优化问题的解就是超图的拉普拉斯矩阵 $\boldsymbol{\Delta}$ 的最小非零特征值所对应的特征向量 $\boldsymbol{\Phi}$ 。因此超图的一次分割将顶点集 $\mathbf{V}$ 聚类为两个部分: $\mathbf{S} = \{v \in \mathbf{V} | \Phi(v) \geq 0\}$ 和 $\mathbf{S}^c = \{v \in \mathbf{V} | \Phi(v) < 0\}$ 。至此完成了帧超图的一次最优谱分割。

超图的谱分割可对超图递归地进行二分, 直到 $c(\mathbf{S})$ 的值以及分割所得各部分内超边的平均权值均不小于相应的阈值。通过谱分割可以分别获得轨迹类和多标签类别, 结合两者的配对融合可以获得

轨迹与语义标签之间的对应关系, 语义标签组成了视频事件的描述, 完成了从底层的时空特性到语义概念的映射, 并获得了视频复杂事件。用此方法可对视频复杂事件进行检测。

4 实验分析

选择的实验集分别为镇江海关监控系统中采集的监控视频集 (CMV), 镇江市地方海事局长江航道视频监控系统数据集 (YRV) 和视频数据集 CAVIAR^[21], 进行事件检测实验。

表 1 所示为实验视频集详细信息。其中, 本系统对 CMV 中的码头油轮装 (卸) 载事件, YRV 在某航道中的多船只航行事件, CAVIAR 中选择的 26 个行人购物事件和 4 个多人次争执事件等进行分析处理。在 CMV 中, 主要选取了 5 个典型事件; 在 YRV 中选取了 5 个典型事件; 在 CAVIAR 中, 则选取了购物和争执事件进行了实验, 其中, 选择 80% 的数据作为训练集, 20% 的数据作为实验集。为了通过训练建立各个典型事件的模型, 通过视频剪辑分别为每个事件制作了相应的训练视频, 详细信息如表 2 所示。

表 1 实验视频集的详细信息

视频集名称	所含视频数	总帧数	事件数	子事件数
CMV	33	90 548	85	426
YRV	46	45 822	63	152
CAVIAR	30	37 267	87	179

图 3 所示分别为 YRV 数据库中, 采砂船在航道航行的视频帧 Frame0005, Frame0030。图 3 中线条部分为检测到的运动目标轨迹。

为评估本文所提出的超图的正确性, 采用图变换 (graph transformation) 工具 AGG 进行实验。AGG 是一个基于代数方法的图变换系统仿真环境, 它不仅可以图形方式模拟图变换系统的变换过程, 而且还可检查图变换系统的一致性。表 2 中 YRV 数据库里 5 个事件中的多运动对象采砂船 (ship), 小船 (boat) 之间的事件轨迹集合 $\{T_1, T_2, T_3, T_4, T_5, T_6\}$ 关系如图 4(a) 所示, 图 4(b) 为上述视频集中多标签超图中标签集合 $\{l_1, l_2, l_3, l_4, l_5, l_6\}$, 得到上述两种超图后采用配对融合方法对轨迹超图和多标签超图进行关联计算, 再进行谱分割, 从而得到复杂事件

表 2 各视频集中事件描述信息

Table 2 Events description of each video sets

视频集名称	事件编号	事件描述
CMV	C_Event1	油轮靠港,卸油,停止作业
	C_Event2	船只靠近油轮,船只停靠油轮,开始偷油,停止作业
	C_Event3	桥起重机开始移动,移动至油轮上方,停止移动
	C_Event4	船只靠近油轮,检查,船只停留油轮侧方
YRV	Y_Event1	采砂船航行,停航,继续行驶
	Y_Event2	采砂船航行,停航,开始作业
	Y_Event3	采砂船航行,停航,小船靠近,接驳,采砂船继续航行
	Y_Event4	多只船只正常航行
	Y_Event5	多船只航行,某船停航,航道拥堵
CAVIAR	shopping1	有人走入楼层长廊,进店,购物,出店,继续行走
	shopping2	行走,聊天,进店,出店,有人加入,聊天,再次进店
	shopping3	行走,聊天,分别,某人进店,其他人继续行走
	shopping4	行走,店门停留,继续行走
	shopping5	多人在长廊行走,部分人进店
	Fighting1	两人见面,争执,各自走
	Fighting2	两人见面,争执,一人坐地,一人站立,离开
	Fighting3	两人见面,争执,谈话



(a) YPV视频帧Frame0005



(b) YPV视频帧Frame0030

图 3 实验集 YRV 视频帧中运动轨迹检测

Fig. 3 Trajectories detection in YRV

类别,即检测出复杂事件。

在实验过程中,将本文提出的方法轨迹和多标签超图配对融合方法(TG-MLG)和其他方法进行了

对比分析:基于普通图的事件检测方法(CGC)^[16],将采用 Hausdroff 距离度量轨迹相似度的轨迹超图加入到由文献[18]多标签半监督学习方法形成的超图模型中(HG-LGC)。图 5 为上述方法分别在数据集中生成超图的顶点数目,其中 TG-MLG 和 HG-LGC 的顶点数目为所生成轨迹和多标签超图顶点之和。图 6 为生成的聚类数目对比,其中 TG-MLG 和 HG-LGC 的聚类数目为所生成轨迹和多标签超图聚类之和。因此,TG-MLG 的顶点数目和聚类数目都高于其他两种方法。

为了客观地评价本文方法的合理性和有效性,采用最常用的平均查准率(MAP)和平均查全率(MAR)作为衡量指标^[22]。设 $N_{\text{correct}}(c_i)$ 表示用 c_i 正确检测测试集中事件数, $N_{\text{pLabel}}(c_i)$ 表示用 c_i 检测测试集中事件数, $N_{\text{label}}(c_i)$ 表示测试集中与 c_i 相关的实际事件数,则 MAP、MAR 的具体计算为

$$\text{MAP} = \frac{1}{k} \sum_{i=1}^k \text{precision}(c_i) \quad (14)$$

$$\text{MAR} = \frac{1}{k} \sum_{i=1}^k \text{recall}(c_i) \quad (15)$$

$$\text{precision}(c_i) = \frac{N_{\text{correct}}(c_i)}{N_{\text{pLabel}}(c_i)} \quad (16)$$

$$\text{recall}(c_i) = \frac{N_{\text{correct}}(c_i)}{N_{\text{label}}(c_i)} \quad (17)$$

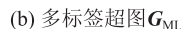
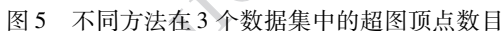


Fig. 4 Simulation analysis of multi-event relationship in YRV using AGG



Method	HG-LGC	TG-MLG	CGC
CMV	67	72	65
YRV	56	57	54
CAVIAR	43	46	40

Fig. 6 Number comparison of clusters in hypergraph with three different methods

确。例如在平均查准率比较中,检测效果提高最多的是在对语义“p_shop”进行的检测,TG-MLG的平均查准率比HG-MLG高6%。和CGC相比,在语义“bridgecrane_work”检测时,TG-MLG

的平均查准率高出 5%。在平均查全率比较中,和 HG-MLG 相比,在语义“oiltank_movein”,“oiltank_loading”,“vessel_collision”和“p_shop”检测时,TG-MLG 的平均查全率要高出 8%。和 CGC 相比,在语义“p_argue”检测时,TG-MLG 的平均查全率要高出 6%。TG-MLG 的检测效

果优于其他两种方法的原因是,本文方法将轨迹与标签之间关系进行映射,并对事件运动属性与标签之间的时序条件依赖关系进行配对融合,可更加准确地将事件标签与运动事件进行配对,同时又能更全面地检测到时序上发生相关的事件。

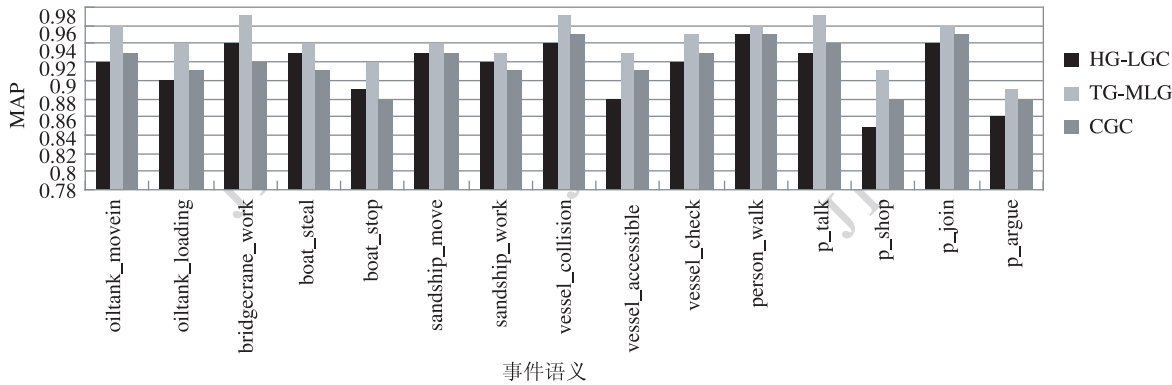


图 7 15 个事件语义的平均查准率对比
Fig. 7 Comparison of MAP of 15 event semantics

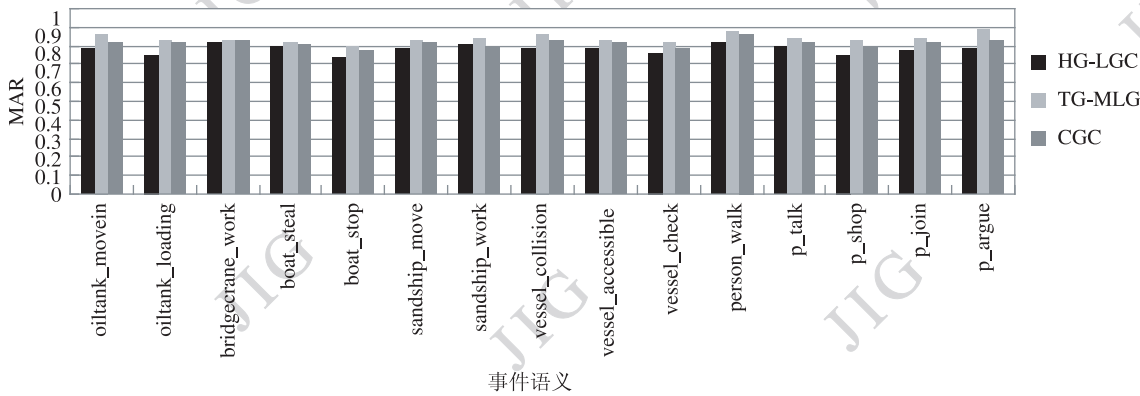


图 8 15 个事件语义的平均查全率对比
Fig. 8 Comparison of MAR of 15 event semantics

5 结 论

如何跨越底层特征与高层语义间的鸿沟,已成为视频研究领域中的热点问题,其中视频事件分析重点在运动事件的概念语义模式发现上进行研究。本文在超图理论基础上,提出了根据运动目标的时空特性提取轨迹为其建立超图模型,同时建立视频概念多标签超图,并对两者进行配对融合,以获得轨迹与多标签间的映射关系,从而检测复杂事件。新方法与其他方法进行了对比实验,结果显示具有更高的准确性和查全率。运用该方法对监控视频中事件检测分析取得了较好结果,今后将在含有其他事

件类型的视频集中展开实验。选择不同特征,构建事件特征超图,并对多超图的融合研究将是下一步展开的主要工作。

参考文献 (References)

[1] Lu H, Tan Y P. Sports video analysis and structuring [C]// Proceedings of IEEE the 4th Workshop on Multimedia Signal Processing. Cannes, France: Institute of Electrical and Electronics Engineers Inc. , 2001 :45-50.

[2] Chen C H, Lien C C. The multi-resolution and flooding based RSST [J]. Image Segmentation Method, 2003, (1) :9-16.

[3] Gao Y, Wang W B, Yong J H, et al. Dynamic video summariza-

- tion using two-level redundancy detection[J]. *Multimedia Tools and Applications*, 2009, (42): 233-250. [DOI: 10.1007/s11042-008-0236-x]
- [4] Huang C L, Shih H C, Chao C Y. Semantic analysis of soccer video using dynamic Bayesian network[J]. *IEEE Transactions on Multimedia*, 2006, 8(4): 749-760. [DOI: 10.1109/TMM.2006.876289]
- [5] Suk H I, Sin B K, Lee S W. Hand gesture recognition based on dynamic Bayesian network framework[J]. *Pattern Recognition*, 2010, 43(9): 3059-3072. [DOI: 10.1016/j.patcog.2010.03.016]
- [6] Assfalg J, Bertini M, Colombo C, et al. Semantic annotation of soccer videos: automatic highlights identification[J]. *Computer Vision and Image Understanding*, 2003, 92(2): 285-305. [DOI: 10.1016/j.cviu.2003.06.004]
- [7] Kalman R E, Bucy R S. New results in linear filtering and prediction theory[J]. *Transactions of the ASME-Journal of Basic Engineering*, 1961, (83): 95-107.
- [8] Maggio E, Cavallaro A. Hybrid particle filter and mean shift tracker with adaptive transition model[C]//*Proceedings of IEEE Signal Processing Society International Conference on Acoustics, Speech, and Signal Processing*. Philadelphia, PA, United States: Institute of Electrical and Electronics Engineers Inc., 2005: 221-224. [DOI: 10.1109/ICASSP.2005.1415381]
- [9] Hongeng S, Navatia R, Bremond F. Video-based event recognition: activity representation and probabilistic recognition methods[J]. *Computer Vision and Image Understanding*, 2004, 96: 129-162. [DOI: 10.1016/j.cviu.2004.02.005]
- [10] Liu Q S, Huang Y C, Dimitris M. Hypergraph with sampling for image retrieval[J]. *Pattern Recognition*, 2011, 44(10): 2255-2262. [DOI: 10.1016/j.patcog.2010.07.014]
- [11] Ding L, Yilmaz A. Interactive image segmentation using probabilistic hypergraphs[J]. *Pattern Recognition*, 2010, (43): 1863-1873. [DOI: 10.1016/j.patcog.2009.11.025]
- [12] Huang Y C, Liu Q S, Dimitris M. Video object segmentation by hypergraph cut[C]//*Proceedings of Intl Conf. on Computer Vision and Pattern Recognition*. Miami, FL, United States: IEEE Computer Society, 2009: 1738-1735.
- [13] Hakeem A, Shah M. Learning, detection and representation of multi-agent events in videos[J]. *Artificial Intelligence*, 2007, 171(8-9): 586-605. [DOI: 10.1016/j.artint.2007.04.002]
- [14] Han Y H, Shao J, Wu F, et al. Multiple hypergraph ranking for video concept detection[J]. *J. Zhejiang Univ. : Sci. C (Comput & Electron)*, 2010, 11(7): 525-537. [DOI: 10.1631/jzus.C0910453]
- [15] Jiang Y, She Q Q, Li M, et al. A transductive multilabel text categorization approach[J]. *Journal of Computer Research and Development*, 2008, 45(11): 1817-1822. [姜远, 余俏俏, 黎铭, 等. 一种直推式多标记文档分类方法[J]. *计算机研究与发展*, 2008, 45(11): 1817-1822.]
- [16] Chen Z J, Chen X J, He H. Moving object detection based on improved mixture gaussian models[J]. *Journal of Image and Graphics*, 2007, 12(9): 1585-1589. [陈祖爵, 陈潇君, 何鸿. 基于改进的混合高斯模型的运动目标检测[J]. *中国图象图形学报*, 2007, 12(9): 1585-1589.]
- [17] Ke J, Zhan Y Z, Chen X J. The research of detect moving target algorithm pseudo invariant line moment-based[C]//*Proceedings of ICIC2009*. Ulsan, South Korea: Springer, 2009: 615-624.
- [18] Piciarelli C, Foresti G L. On-line trajectory clustering for anomalous events detection[J]. *Pattern Recognition Letters*, 2006, 27(15): 1835-1842. [DOI: 10.1016/j.patrec.2006.02.004]
- [19] Ke J, Zhan Y Z, Chen X J, et al. The retrieval of motion event by associations of temporal frequent pattern growth[J]. *Future Generation Computer Systems*, 2013, 29: 442-450. [DOI: 10.1016/j.future.2011.06.004]
- [20] Chen D W, Liu R, Yuan Z M, et al. Research on multi-concept learning based on inter-concept relation[J]. *Computer Science*, 2011, 38(4): 244-248. [陈丹雯, 刘然, 袁志明, 等. 基于概念对关系的视频多语义概念学习方法研究[J]. *计算机科学*, 2011, 38(4): 244-248.]
- [21] Fisher R, Santos-Victor J, Crowley J. CAVIAR: context aware vision using image-based active recognition[DB/OL]. (2013-07-01) [2012-06-10]. <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>.
- [22] Lu H Q, Liu J. Image annotation based on graph learning[J]. *Chinese Journal of Computers*, 2008, 31(9): 1629-1639. [卢汉清, 刘静. 基于图学习的自动图像标注[J]. *计算机学报*, 2008, 31(9): 1629-1639.]