

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **1**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年1月 第17卷 第1期(总第189期)

目 次

综述

压缩感知基本理论:回顾与展望 邵文泽, 韦志辉(1)

图像处理和编码

完全质量保持的 AVS 视频可逆信息隐藏算法 郭宝安, 陈彪, 张飏, 于志强, 张卫明(13)

多成分正则化约束的断层图像重建算法 王丽艳, 韦志辉, 李星秀(20)

基于单幅图像景深和大气散射模型的去雾方法 郭珈, 王孝通, 胡程鹏, 徐晓刚(27)

结合 NSS 和小波变换的无参考图像质量评价 金波, 李朝锋, 吴小俊(33)

结合区域分割和双边滤波的图像去噪新算法 杨学志, 徐勇, 方静, 卢洁, 左美霞(40)

种子像素滤波法去除随机脉冲噪声 单建华(49)

基于纹理和亮度感知特性的率失真优化策略 俞力克, 代锋, 张勇东, 林守勋(54)

单通道散焦模糊置换图像的盲分离 王伟, 方勇(62)

侧扫声呐图像的 3 维块匹配降斑方法 范习健, 李庆武, 黄河, 王敏(68)

新适应模型的提出及分类扩散 李彦宝, 姜广峰, 王治强, 刘薇(75)

图像分析和识别

从神经切片图像中识别功能束类型的研究 钟映春, 罗鹏(82)

融合 SPA 遮挡分割的多目标跟踪方法 丁欢, 张文生(90)

目标识别中的稳定图像特征组合发掘 姜永兵, 彭启民(99)

基于结构特征的高分辨率 TerraSAR-X 图像船舶识别方法研究 殷雄, 王超, 张红, 吴樊(106)

图像理解和计算机视觉

机器人的混合特征视觉环境感知方法 杨俊友, 马乐, 白殿春, 东俊光(114)

纹理分布分析的快速图像修复算法 张晴, 林家骏(123)

计算机图形学

圆的整数反走样生成算法 贾银亮, 张焕春, 经亚枝(130)

遥感图像处理

SAR 图像的快速核主成分分析识别方法 于红芸, 姜涛, 关键 (137)

基于图像内容层次表征的遥感图像分割方法 张学良, 肖鹏峰, 冯学智 (142)

地理信息技术

保持几何模式的城市道路路网简化方法 杨必胜, 张云菲, 栾学晨 (150)

中国图象图形学报

刊名题字: 宋 健 月刊(1996 年创刊)

第 17 卷 第 1 期

2012 年 1 月 16 日出版

主管单位 中国科学院
主 办 中国科学院遥感应用研究所
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
北京 9718 信箱 邮编 100101
电子信箱: jig@irsa. ac. cn
电话: 010-68407995 010-82614429
网 址: www. cjig. cn

印刷装订 北京北林印刷厂
广告经营许可证 京朝工商广字第 0346 号
总 发 行 北京报刊发行局
订 购 全国各地邮局
国外发行 中国国际图书贸易总公司
(中国国际书店)
(北京 399 信箱 邮编 100044)

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing Application,
CAS China Society of Image and Graphics
Institute of Applied Physics and Computational
Mathematics
Chief editor LI Xiaowen
Editor, Publisher Editorial and Publishing Board
of Journal of Image and Graphics
(P. O. Box 9718, Beijing 100101, China)
E-mail: jig@irsa. ac. cn
Distributed by Beijing Bureau for Distribution of Newspapers
and Journals
Domestic All Local Post Offices in China
Foreign China International Book Trading Corporation
(P. O. Box 399, Beijing 100044, China)
Printed by Beijing Beilin Printing House

Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 1 January 2012

Contents

Review

Advances and perspectives on compressed sensing theory *Shao Wenze, Wei Zhihui*(1)

Image Processing and Coding

Complete video quality-preserving reversible data-hiding based on AVS *Guo Bao'an, Chen Biao, Zhang Biao, Yu Zhiqiang, Zhang Weiming*(13)

Morphological component analysis for tomography reconstruction *Wang Liyan, Wei Zhihui, Li Xingxiu*(20)

Single image dehazing based on scene depth and physical model *Guo Jia, Wang Xiaotong, Hu Chengpeng, Xu Xiaogang*(27)

No-reference image quality assessment based on natural scene statistics and wavelet *Jin Bo, Li Chaofeng, Wu Xiaojun*(33)

New filter based on region segmentation and bilateral filtering *Yang Xuezhi, Xu Yong, Fang Jing, Lu Jie, Zuo Meixia*(40)

Removal of random-valued impulse noise based on seed pixels filter *Shan Jianhua*(49)

Novel rate distortion optimization strategy based on perceptual properties of texture and luminance *Yu Like, Dai Feng, Zhang Yongdong, Lin Shouxun*(54)

Blind separation of single-channel permuted defocus blurred image *Wang Wei, Fang Yong*(62)

Side-scan sonar image despeckling based on block-matching and 3D filtering *Fan Xijian, Li Qingwu, Huang He, Wang Min*(68)

New well-posed model and classified diffusion *Li Yanbao, Jiang Guangfeng, Wang Zhiqiang, Liu Wei*(75)

Image Analysis and Recognition

Type recognition of fascicular groups from nerve slice image *Zhong Yingchun, Luo Peng*(82)

Multi-target tracking approach combined with SPA occlusion segmentation *Ding Huan, Zhang Wensheng*(90)

The mining of stable image feature-compositions in object recognition *Jiang Yongbing, Peng Qimin*(99)

Vessel recognition with high resolution terraSAR-X image based on structure feature *Yin Xiong, Wang Chao, Zhang Hong, Wu Fan*(106)

Image Understanding and Computer Vision

Robot vision environmental perception method based on hybrid features *Yang Junyou, Ma Le, Bai Dianchun, Toshimitsu Higashi*(114)

Fast image inpainting algorithm using color texture distribution analysis *Zhang Qing, Lin Jiajun*(123)

Computer Graphics

Integral algorithm for circle anti-aliasing *Jia Yinliang, Zhang Huanchun, Jing Yazhi*(130)

Remote Sensing Image Processing

SAR images recognition based on kernel principal component analysis *Yu Hongyun, Jiang Tao, Guan Jian*(137)

Remotely sensed imagery segmentation based on the hierarchical representation of image content *Zhang Xueliang, Xiao Pengfeng, Feng Xuezhi*(142)

Geoinformatics

Pattern preserving method for grid simplification in road networks *Yang Bisheng, Zhang Yunfei, Luan Xuechen*(150)

中图分类号: TP391.4 文献标志码: A 文章编号: 1006-8961(2012)01-0090-09

论文引用格式: 丁欢, 张文生. 融合 SPA 遮挡分割的多目标跟踪方法 [J]. 中国图象图形学报, 2012, 17(1): 90-98

融合 SPA 遮挡分割的多目标跟踪方法

丁欢, 张文生

中国科学院自动化研究所控制国家重点实验室, 北京 100190

摘要: 复杂环境下的多目标视频跟踪是计算机视觉领域的一个难点, 有效处理目标间遮挡是解决多目标跟踪问题的关键。将运动分割方法引入目标跟踪领域, 提出一种融合骨架点指派 (SPA) 遮挡分割的多目标跟踪方法。由底层光流信息得到骨架点, 并估计骨架点遮挡状态; 综合使用目标外观、运动、颜色信息等高级语义信息, 将骨架点指派给各个目标; 最后以骨架点为核, 对运动前景密集分类, 得到准确的目标前景像素; 在粒子滤波器跟踪框架下, 使用概率外观模型进行多目标跟踪。在 PETS2009 数据集上的实验结果表明, 文中方法能够改进现有多目标跟踪方法对目标间交互适应性较差的缺点, 更好地处理动态遮挡问题。

关键词: 多目标跟踪; 遮挡分割; 粒子滤波; 骨架点; 运动分割

Multi-target tracking approach combined with SPA occlusion segmentation

Ding Huan, Zhang Wensheng

State Key Laboratory of Intelligent Control and Management of Complex Systems

Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China

Abstract: Multiple-target tracking in complex scenes is one of the most complicated problems in computer vision. Handling occlusions between objects is the key issue in multiple-target tracking. This paper introduces a method of motion segmentation into the object tracking system, and presents a SPA (skeleton points assign) based occlusion segmentation approach to track multiple people through complex situations which are captured by static monocular cameras. In the proposed method, we select the skeleton points and evaluate their occlusion states by bottom information like optical flow; then we assign these points to different objects using advanced semantic information, such as appearance, motion, and color. Finally a dense classification of foreground pixels is used to accomplish occlusion segmentation. Object tracking is handled by a particle filter-based tracking framework, and a probabilistic appearance model is used to find the best particle. Experiments are performed using the public challenging data set PETS 2009. The experimental results show that our approach can improve the performance of the existing tracking approach and handle dynamic occlusions better.

Key words: multiple-target tracking; occlusion segmentation; particle filter; skeleton points; motion segmentation

0 引言

多目标跟踪是计算机视觉领域的一个重要应

用, 被广泛用于智能安防、机器人视觉、智能人机交互、交通监控、军事制导等多个方面, 是视频分析与监控系统的基础。但是, 在单目固定摄像头环境下, 由于目标高速运动、动态交互、复杂背景等因素, 尤其是外观相似目标间的遮挡, 使跟踪的鲁棒性面临

收稿日期: 2011-03-15; 修回日期: 2011-06-08

基金项目: 国家自然科学基金项目 (90924026); 国家高技术研究发展计划 (863) 项目 (2008AA01Z121, 2007AA01Z338); 首都科技条件平台中国科学院研发实验服务基地研发实验基金项目

第一作者简介: 丁欢 (1985—), 男, 现为中国科学院自动化研究所复杂系统与智能科学重点实验室模式识别与智能系统专业博士研究生, 主要研究方向为图像处理、视频跟踪和行为识别。E-mail: huan.ding@ia.ac.cn

巨大挑战。因此,如何在跟踪中进行遮挡处理,克服相似目标之间遮挡、交互对跟踪结果的影响(图1),成为当前一项亟待解决的问题。

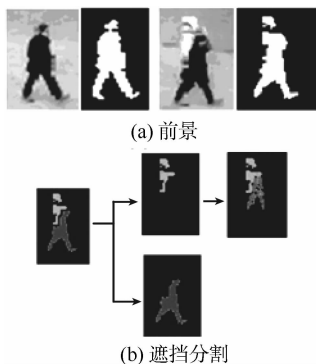


图1 目标动态遮挡及遮挡分割补偿

Fig.1 Dynamic occlusion and occlusion segmentation

目前主流的多目标跟踪方法主要包括多假设跟踪器(MHT)^[1]以及联合概率数据关联滤波器(JPDAF)^[2],而遮挡使MHT及JPDAF方法对目标的观测值与实际目标出现较大偏差,造成跟踪结果偏移甚至错误。一种解决遮挡问题的方法是使用多个摄像头补偿单摄像头视野^[3]。然而,在很多实际应用中,从成本及计算效率两方面考虑,都无法满足在同一地点部署多个摄像头的要求。

现有单目固定摄像头多目标跟踪方法大多将遮挡作为模型的一部分隐含入跟踪过程中,使用颜色^[4]、外观^[4-7]、运动^[8-10]信息对目标建模,但这些模型侧重跟踪过程中的目标描述,对各种遮挡区分能力较差。

对遮挡进行专门处理的系统中,较有代表性的工作包括: Hai等人^[11]将视频帧分割成为多个运动层,采用EM方法估计目标形状、运动; Qian等人^[12]使用混合高斯模型估计目标的运动与外观,以此作为MCMC框架下跟踪假设与观测数据关联的依据; Papadourakis等人^[13]利用颜色信息构建像素距离,并对前景像素聚类,尝试解决目标长时间遮挡问题。上述遮挡处理方法都简单假设目标的颜色满足单高斯或混合高斯分布,没有考虑到目标颜色外观分布对遮挡处理的重要作用,同时对目标在场景中的历史信息利用较少,导致对颜色相近目标之间的相互遮挡处理效果不佳。

在最近的文献中,运动分割领域的许多方法被引入跟踪框架中,以得到更准确的跟踪结果。如 Papadakis等人^[14]采用位置、运动信息构造能量函

数,采用图分割方法切分运动目标; Silva等人^[15]根据光流信息提取运动特征点,并对轨迹进行聚类。这些方法依旧依托于图像运动分割方法,从最底层的像素进行运动分割,缺少与跟踪目标高级语义特征的关联印证,跟踪结果误差较大,同时这些方法对所有像素进行均等分析,时间代价巨大。

针对上述问题,提出一种新颖的融合骨架点指派(SPA)的遮挡分割方法,提取目标的光流特征点作为骨架点,根据每个骨架点在一段时间序列中的存在情况,将其分类成为可见点、遮挡点等不同类型分别处理,之后充分利用运动目标外观^[7]、颜色及运动等高阶特征,将骨架点指派给各运动目标。在此基础上,将所有运动前景密集分割,最后在粒子滤波器框架下进行多目标跟踪。

1 方法概述

本文中采用的单目摄像头多目标跟踪方法主要由3个步骤组成:

1)背景建模与运动前景提取 采用混合高斯模型^[16]对监控场景建模,并提取出运动前景。本步骤将单目固定摄像头采集的图像作为输入,输出为二值化的运动前景图像。

2)融合SPA的遮挡分割 本步骤是本文的主要工作,根据当前跟踪目标情况,将上步得到的前景图像中的运动前景点进行分割指派,分别作为不同目标的运动前景。分割过程主要可以分为如下3个部分:

(1)骨架点提取与匹配(见第2节) 根据场景光流速度信息,提取出目标的骨架点,这里我们考察过去一帧图像、当前图像以及预读一帧图像,在这连续3帧的图像序列中,计算每两帧图像中骨架点的匹配情况,得到骨架点匹配矩阵,用于估计骨架点类别。

(2)骨架点指派(见第3节) 根据连续3帧图像中骨架点的匹配矩阵,构造骨架点状态转移矩阵,用于得到当前帧骨架点的状态,之后与当前存在跟踪各目标的颜色、空间、运动等信息比较,将骨架点指派给各目标。

(3)运动前景密集分类(见第4节) 将其他运动前景以已指派的骨架点为核分类,从而实现所有运动前景的分割,得到每个目标的运动前景。

3)粒子滤波框架下的多目标跟踪 采用当前比较成熟的粒子滤波器跟踪框架^[17-18],使用概率外观模型^[7]估计系统状态似然度,进而得到最终多目

标跟踪结果。

2 骨架点的提取与匹配

本文提出的遮挡分割方法是围绕“骨架点”(skeleton points)进行的。这里的骨架点指较均匀分布于目标周围,能够代表运动目标信息的特征点。通过考察骨架点在连续 3 帧视频图像序列中的出现情况,判断骨架点状态(被遮挡点、形态学出现等)。

2.1 骨架点的提取

本文所需的骨架点主要用于为遮挡分割提供依据,因此应当具有如下 3 个特征:

1) 骨架点数目应当适当,且均匀分布在目标及其周围,这样才能保证骨架点能够包含目标的大部分信息且计算成本较小;

2) 骨架点位置应当尽量稳定,这样通过骨架点匹配才能正确判断出被遮挡的特征点;

3) 由于被跟踪目标,尤其是行人目标,大部分属于非刚体,因此骨架点应当在处理非刚性目标匹配上有较好性能。

目前最常用的特征点提取方法,如 Harris、KLT 及 SIFT 等,均是在目标亮度变化较大处分布较多,而在目标亮度变化较平滑的地方分布较少,同时这些特征点主要用于处理刚性目标间的匹配,而在非刚性目标匹配上容易出现较大误差。因此,这些特征不适合选为本文方法中的骨架点。

本文中我们选用 Sand 等人^[19]采用的光流特征点作为骨架点(图 2)。这种特征点基于帧间光流速度,并充分考虑了尺度变化与遮挡对光流计算的影响,特征点密度自适应性强,场景中不存在运动时,密度稀疏,而在存在运动的部分密度非常稠密,能够充分覆盖运动目标,便于后续分析。



图 2 光流特征点示例

Fig. 2 The optical flow feature points

设 t 时刻通过文献[19]得到的特征点构成集合 P_t , 本文主要考虑对运动前景的分割问题,因此忽略场景中背景物体中的特征点,只考虑运动前景中的特征点,因此可以构成 t 时刻骨架点集合 $SP_t = \{p | p \in P_t \cap F_t\}$, 其中 F_t 为 t 时刻运动前景像素集合。

2.2 骨架点匹配

为确定骨架点在不同时刻图像中的出现情况,需要对连续时间序列图像中的骨架点进行匹配。为能够真正确定当前帧骨架点的状态,我们预取一帧图像 I_{t+1} , 加上当前帧图像 I_t 及上一帧图像 I_{t-1} , 计算这 3 帧图像两两之间的骨架点匹配矩阵。

定义 1 设第 t 帧图像中,骨架点个数为 M_t , 则定义 t_1, t_2 两帧图像间的匹配矩阵为一个 $M_{t_1} \times M_{t_2}$ 的二值矩阵 $MatchMap^{t_1, t_2}$, 其中每个矩阵元素取值为

$$MatchMap_{i,j}^{t_1, t_2} = \begin{cases} 1 & \text{若 } SP_{t_1} \text{ 中第 } i \text{ 个点与 } SP_{t_2} \\ & \text{中第 } j \text{ 点匹配} \\ 0 & \text{其他} \end{cases} \quad (1)$$

采用一种贪心的匹配策略(算法 1)寻找骨架点的匹配点,即首先在骨架点集合 SP_{t_1} 中取一个点,在 SP_{t_2} 中寻找匹配相似度最大的骨架点,找到后将这两个点从待匹配集合中删除,直到 SP_{t_1} 中所有的骨架点都完成匹配运算。此处,我们使用骨架点的颜色信息与空间位置信息作为判断匹配似然度的标准,定义匹配似然度函数为

$$SpLikelihood(sp_{t_1}, sp_{t_2}) = e^{-spaceDis(sp_{t_1}, sp_{t_2})^2 / s_{space}^2} e^{-colorDis(sp_{t_1}, sp_{t_2})^2 / s_{col}^2} \quad (2)$$

式中, $spaceDis(sp_{t_1}, sp_{t_2})$ 、 $colorDis(sp_{t_1}, sp_{t_2})$ 分别表示两骨架点在欧氏空间与 RGB 颜色空间中的距离,定义如下

$$spaceDis(SP_{t_1}, SP_{t_2}) = (x_{SP_{t_1}} - x_{SP_{t_2}})^2 + (y_{SP_{t_1}} - y_{SP_{t_2}})^2 \quad (3)$$

$$colorDis(SP_{t_1}, SP_{t_2}) = (r_{SP_{t_1}} - r_{SP_{t_2}})^2 + (g_{SP_{t_1}} - g_{SP_{t_2}})^2 + (b_{SP_{t_1}} - b_{SP_{t_2}})^2 \quad (4)$$

具体骨架点匹配(算法 1)可描述如下:

初始化 匹配矩阵 $MatchMap^{t_1, t_2}$ 所有元素初始化为 0
Do

骨架点集合 SP_{t_1} 中取元素 i ;

For 骨架点集合 SP_{t_2} 中所有元素

以式(2)作为似然度函数,计算两骨架点之间

距离;

找到似然度最大的元素 j , 似然度为 $maxlikelihood$;

If $maxlikelihood$ 大于阈值

$MatchMap_{i,j}^{t_1,t_2} = 1$;

将 i, j 分别从 SP_{t_1}, SP_{t_2} 中去除;

Else

将 j 从 SP_{t_1} 中去除;

Until

SP_{t_1} 或 SP_{t_2} 为空。

如图3所示, t_1 帧图像中, 共有5个骨架点, t_2 图像中有6个骨架点, 经算法1得到的骨架点匹配关系如图中黑线所示, 则这两帧图像骨架点匹配矩阵如表1所示。矩阵中的1表示某骨架点同时在 t_1, t_2 两帧图像中出现, 矩阵中全为0的行(列)表示 $t_1(t_2)$ 中未能与 $t_2(t_1)$ 找到匹配的骨架点, 如 t_2 中的骨架点 sp_4 。通过上述算法, 我们可得到连续3帧图像中两两间的骨架点匹配矩阵 $MatchMap^{t,t-1}$ 、 $MatchMap^{t,t+1}$ 及 $MatchMap^{t-1,t+1}$, 能够表述出骨架点在各图像中的出现情况, 进而判断每个骨架点的实际状态。

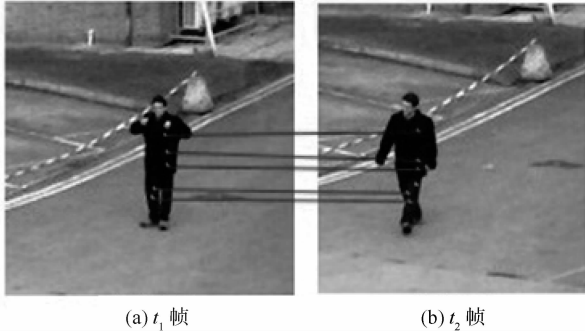


图3 骨架点匹配示意图

Fig. 3 Skeleton points match between two objects

表1 图3两帧图像匹配矩阵 $MatchMap^{t_1,t_2}$

Table 1 $Matchmap^{t_1,t_2}$ of skeleton points in Fig. 3

| | sp_1 | sp_2 | sp_3 | sp_4 | sp_5 | sp_6 |
|--------|--------|--------|--------|--------|--------|--------|
| sp_1 | 1 | 0 | 0 | 0 | 0 | 0 |
| sp_2 | 0 | 1 | 0 | 0 | 0 | 0 |
| sp_3 | 0 | 0 | 1 | 0 | 0 | 0 |
| sp_4 | 0 | 0 | 0 | 0 | 1 | 0 |
| sp_5 | 0 | 0 | 0 | 0 | 0 | 1 |

3 骨架点指派

得到骨架点匹配矩阵 $MatchMap^{t,t-1}$ 、

$MatchMap^{t,t+1}$ 及 $MatchMap^{t-1,t+1}$ 后, 我们可根据骨架点在连续3帧图像中的存在情况, 即状态转移情况, 判断出骨架点属于新出现的骨架点、被遮挡骨架点、原有骨架点等状态, 这一过程通过构造骨架点状态转移矩阵并对骨架点的状态转移模式进行深入分析得到; 之后, 依据当前目标的颜色、空间、速度等信息, 对所有可见骨架点分类, 将其指派到不同跟踪目标。

3.1 骨架点状态分析

在一段连续视频图像序列中, 骨架点可能的状态有如下几种: 始终可见的骨架点、被遮挡的骨架点、由于目标形态变化而新出现的骨架点。本文构造骨架点状态转移矩阵 $STMap^t$ 记录连续3帧图像中骨架点出现情况及其对应的骨架点状态。

定义2 定义某骨架点状态转移情况可表示为向量 $ST^l = (s_{t-1}, s_t, s_{t+1})$, 其中当该骨架点存在于第 l 帧图像时 s_l 取1, 否则取0, $l = t-1, t, t+1$ 。则可定义第 t 帧图像中骨架点状态转移矩阵为二值矩阵 $STMap^t = [ST_1^t \quad ST_2^t \quad \dots \quad ST_{N_t}^t]^T$, 其中 N_t 为第 $t-1, t, t+1$ 帧中骨架点个数总和。

通过对 $STMap^t$ 所有取值情况的详细分析(表2)可知, 通过考察 $t-1, t, t+1$ 3帧图像中骨架点状态转移矩阵, 可以完全确定出第 t 帧图像中骨架点的状态。通过骨架点匹配矩阵 $MatchMap^{t,t-1}$ 、 $MatchMap^{t,t+1}$ 及 $MatchMap^{t-1,t+1}$ 构造 $STMap^t$, 具体骨架点状态转移矩阵构造方法(算法2)如下:

初始化 $STMap^t$ 中所有元素初始化为0

1) For $STMap^t$ 第 i 行, $1 < i < M_t$, M_t 为第 t 帧中骨架点个数

(1) 每行代表第 t 帧中的一个骨架点, 因此 $STMap_{i, s_t}^t = 1$;

(2) 考察 $MatchMap^{t,t-1}$ 第 i 行, 若 $\exists MatchMap^{t,t-1}[i][j] = 1$, 则 $STMap_{i, s_{t-1}}^t = 1$, 并在 $t-1$ 帧的骨架点中标记 j 点, 否则 $STMap_{i, s_{t-1}}^t = 0$;

(3) 考察 $MatchMap^{t,t+1}$ 第 i 行, 若 $\exists MatchMap^{t,t+1}[i][k] = 1$, 则 $STMap_{i, s_{t+1}}^t = 1$, 并在 $t+1$ 帧的骨架点中标记 k 点, 否则 $STMap_{i, s_{t+1}}^t = 0$ 。

2) For $STMap^t$ 第 i 行, $M_t + 1 < i < M_t + N_t$, N_t 为第 $t-1$ 帧中未标记骨架点个数

(1) 每行代表第 $t-1$ 帧中的一个骨架点, 且 t 帧中没有该点, 因此 $STMap_{i, s_{t-1}}^t = 1, STMap_{i, s_t}^t = 0$;

(2) 考察 $MatchMap^{t-1,t+1}$ 中对对应行, 若存在非0点, 则在 $t+1$ 帧骨架点中标记对应点, 否则 $STMap_{i, s_{t+1}}^t = 0$;

3) For $STMap^t$ 第 i 行, 设 K_t 为第 $t+1$ 帧中未标记骨架

表 2 骨架点状态分析
Table 2 The analysis of skeleton points states

| 骨架点状态转移矩阵 | | | 出现情况描述 | 骨架点状态 |
|-----------|-----|-------|--------------|----------------------------|
| $t-1$ | t | $t+1$ | | |
| 0 | 0 | 0 | 3 帧图像中均无该骨架点 | 无需考虑该情况 |
| 0 | 0 | 1 | 仅在第 3 帧中出现 | $t+1$ 帧中新出现的骨架点或错误点 |
| 0 | 1 | 0 | 仅在第 2 帧中出现 | t 帧中的错误点 |
| 0 | 1 | 1 | 连续两帧中出现 | t 中新出现的骨架点 |
| 1 | 0 | 0 | 连续两帧中消失 | $t-1$ 中的特征点被遮挡 |
| 1 | 0 | 1 | 仅在第 2 帧中消失 | t 帧中出现的错误点 |
| 1 | 1 | 0 | 仅在第 3 帧中消失 | t 帧中的特征点被遮挡或 $t+1$ 出现错误点 |
| 1 | 1 | 1 | 在 3 帧图像中均出现 | 始终可见的骨架点 |

点个数,若 $M_t + N_t + 1 < i < M_t + N_t + K_t$, 代表仅在第 $t+1$ 帧出现的骨架点,因此 $STMap_i^t \cdot s_{t-1} = 0, STMap_i^t \cdot s_t = 0, STMap_i^t \cdot s_{t+1} = 1$

根据得到的 $STMap^t$, 将第 t 帧图像中所有骨架点分成 3 个集合: 始终可见骨架点集合 $V_t = \{sp_i | ST_i^t = (1, 1, 1)\}$, 新出现的骨架点集合 $N_t = \{sp_i | ST_i^t = (0, 1, 1)\}$ 以及被遮挡特征点集合 $O_t = \{sp_i | ST_i^t = (1, 0, 0)\}$, 其中 O_t 中元素坐标为该骨架点在第 $t-1$ 帧中的位置。

3.2 骨架点指派

完成骨架点状态划分后, 使用颜色、空间、运动信息将 V_t 、 N_t 中的点指派给当前跟踪目标。

使用 Nummiaro 等人^[20]中的方法构造目标颜色概率直方图: 通过映射函数 $h(x)$ 将 RGB 颜色空间按照 3 个分量平均划分成 $m = 8 \times 8 \times 8$ 个区间, 将每个目标前景像素统计入颜色概率直方图, 即第 n 个目标在 t 时刻颜色概率分布 $p_{col,n}(t) = \{p_{col,n}(t)^{(u)}\}_{u=1,\dots,m}$ 可通过下式计算:

$$p_{col,n}(t)^{(u)} = f \sum_{\Omega_n} F(r,t) d\{h[I(r,t)] - u\} \quad (5)$$

式中, $f = 1 / \sum_{\Omega_n} F(r,t)$, δ 为 Kronecker delta 函数, Ω_n 是第 n 个目标坐标范围, $r \in \Omega_n$ 为像素图像中的坐标, I, F 分别代表原始图像及前景像素图像。在此基础上可构造骨架点基于颜色统计信息的似然度函数, 即骨架点 sp 出现在第 n 个目标中的颜色置信度

$$likelihood_{col} = p_{col,n}(t)^{(h[I(sp,t)])} \quad (6)$$

同时, 构造目标的空间及运动置信度:

$$likelihood_{space} = e^{-[(x_{sp}-x_n)^2+(y_{sp}-y_n)^2]/\sigma_{space}^2} \quad (7)$$

$$likelihood_{motion} = e^{-[(u_{sp}-u_n)^2+(v_{sp}-v_n)^2]/\sigma_{motion}^2} \quad (8)$$

式中, x_n, y_n 为第 n 个目标的质心, 由第 $t-1$ 帧中 Ω_n 内所有前景像素位置加权得到, u_n, v_n 为目标的平均速度, 由第 $t-1$ 帧中 Ω_n 内所有前景像素光流速度加权得到。

将这 3 种信息融合, 得到某骨架点属于目标 n 可见点集合 V^n 的置信度

$$p(sp \in V^n) = likelihood_{col} \times likelihood_{space} \times likelihood_{motion} \quad (9)$$

依据式(9)计算每个骨架点属于当前场景目标的概率, 并选择概率最大的目标作为该骨架点目标指派的最终结果。骨架点指派结果如图 4 所示, 可见当两目标发生交互时, 采用本文的骨架点指派方法, 能够很好地区分属于不同目标的骨架点。



图 4 骨架点指派结果

Fig. 4 Skeleton points assignment results

4 运动前景密集分类

遮挡分割要求得到目标分割的密集表示, 因此

本文提出的融合 SPA 的遮挡分割方法的最后一步是将每一个运动前景像素指派给一个目标。我们已经知道每帧中的骨架点位置及其对应的目标,所以只需要将每个运动前景像素指派给最近邻的骨架点。为完成这个任务,我们构造基于颜色、空间及运动信息的相似度函数,将运动前景像素与骨架点几何进行比较并分类。

设 $P = \{sp_1, sp_2, \dots, sp_n\}$ 为第 t 帧图像中所有已经被指派到各目标的骨架点集合,为骨架点 sp 指定的目标编号为 $\lambda_{sp} \in \{1, \dots, K\}$ 。我们为 P 中的每个点建立一个包含颜色、运动、空间信息的混合高斯模型,如式(10)所示,以得到每个运动前景像素与该骨架点属于同一目标的概率。

$$G_{sp}^t(x, y, u, v, c, t) = \frac{1}{(2\pi)^2 |\Sigma|^{1/2}} \times \exp \left(-\frac{1}{2} \begin{bmatrix} x - x_{sp}^t \\ y - y_{sp}^t \\ u - u_{sp}^t \\ v - v_{sp}^t \\ c - c_{sp}^t \end{bmatrix}^T \Sigma^{-1} \begin{bmatrix} x - x_{sp}^t \\ y - y_{sp}^t \\ u - u_{sp}^t \\ v - v_{sp}^t \\ c - c_{sp}^t \end{bmatrix} \right) \quad (10)$$

式中, $x_{sp}^t, y_{sp}^t, u_{sp}^t, v_{sp}^t, c_{sp}^t$ 分别表示该骨架点的水平竖直位置坐标、 uv 方向速度以及颜色,协方差矩阵 Σ 定义如下

$$\Sigma = \begin{bmatrix} \sigma_s & 0 & 0 & 0 & 0 \\ 0 & \sigma_s & 0 & 0 & 0 \\ 0 & 0 & \sigma_M & 0 & 0 \\ 0 & 0 & 0 & \sigma_M & 0 \\ 0 & 0 & 0 & 0 & \sigma_c \end{bmatrix}$$

$\sigma_s, \sigma_M, \sigma_c$ 分别表示空间、运动、颜色标准差。

由式(10)可知,当运动前景点距离骨架点空间位置较近、运动速度矢量接近、颜色接近时会产生较大的 G_{sp}^t , 否则 G_{sp}^t 会很小。我们将每个运动前景像素与 P 中的所有骨架点进行比较,按照最大似然策略选择该像素点的标号 λ_p , 最大化整个场景似然度

$$\max_N \sum_{\{sp | \lambda_{sp} = N\}} G_{sp}^t.$$

本文提出的融合 SPA 的遮挡分割方法最终结果如图5所示,运动前景像素由混合高斯背景建模方法得到,通过本文的遮挡分割,可以很好地将前景像素指派到不同的目标。这在粒子滤波多目标跟踪框架中具有非常重要的意义:1)粒子后验概率计算时,若前景中存在大量不属于目标的像素,会极大影

响概率外观似然度计算;2)在概率外观模型模板更新阶段,会把一部分遮挡物体的信息保存入概率外观模型,导致概率外观模型的误差,从而产生跟踪错误。我们的遮挡分割方法很大程度上排除了遮挡的影响,运动前景能够较准确地反映出被遮挡目标的信息,有利于改善跟踪结果。



图5 运动分割结果

Fig. 5 Motion segmentation results

5 实验结果与分析

5.1 遮挡分割结果验证

单目视觉相机中出现的遮挡通常可分为静态遮挡及动态遮挡,我们选取了多种不同的遮挡类型对本文提出的融合 SPA 的遮挡分割方法进行了实验验证。如图6(a)所示,在静态遮挡情况下,本文提出的遮挡分割方法能够把运动目标与场景中的遮挡物区分开;对于相向而行的两个目标,由于运动速度存在较大差别,本文方法也具有较好的区分能力,如图6(b)所示;而当场景中存在多个目标,且目标行进方向相差不多时如图6(c)所示,本文方法也能够较好地处理。在图6(c)中右边两目标出现部分的分割错误是由于两目标颜色非常接近,在计算光流的时候容易出现误差,导致运动信息提取误差。

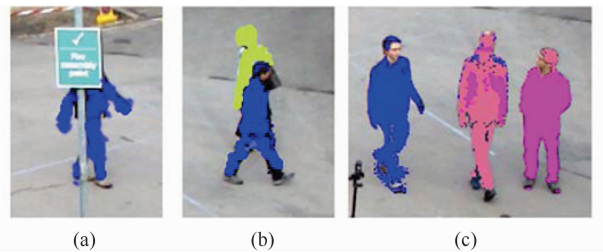


图6 遮挡分割结果

Fig. 6 Occlusion segmentation results

5.2 多目标跟踪结果验证

为验证本文方法的有效性,在 PETS 2009 (performance evaluation of tracking and surveillance

2009) S2. L1 数据集上测试, 获得了满意的结果。PETS 2009 包含各个时段的测试视频, 视频中包含多个运动目标, 目标之间存在大量交互, 经常发生动态遮挡。我们对数据集中的多组视频进行测试, 实验环境为奔腾双核 E5200 处理器, 主频为 2.5 GHz, 内存为 2 GB; 实验平台采用 Visual studio 2005, 并调用了 OpenCV 函数库中的数据结构及部分函数。

实验数据集是分辨率为 768×576 的 RGB 彩色图片, 以 View1 为例, 长度为 795 帧。不计目标在视频中重复出现, 共有 19 个目标进入视野, 各目标行为较复杂, 主要内容见表 3。

表 3 测试视频行为事件统计

Table 3 Challenging events

| 事件类型 | 事件数目 | 事件类型 | 事件数目 |
|---------|------|-------|------|
| 目标交叉 | 52 | 静态遮挡 | 44 |
| 成组行动 | 2 | 新出现目标 | 8 |
| 消失后重新出现 | 9 | 3 人交互 | 2 |

图 7 展示了部分有代表性的跟踪结果, 完整的测试视频序列数据集可在 <http://www.cvg.rdg.ac.uk/PETS2009/a.html> 中得到。图 8 为在 PETS 2009 数据集另一组测试视频上的跟踪结果。

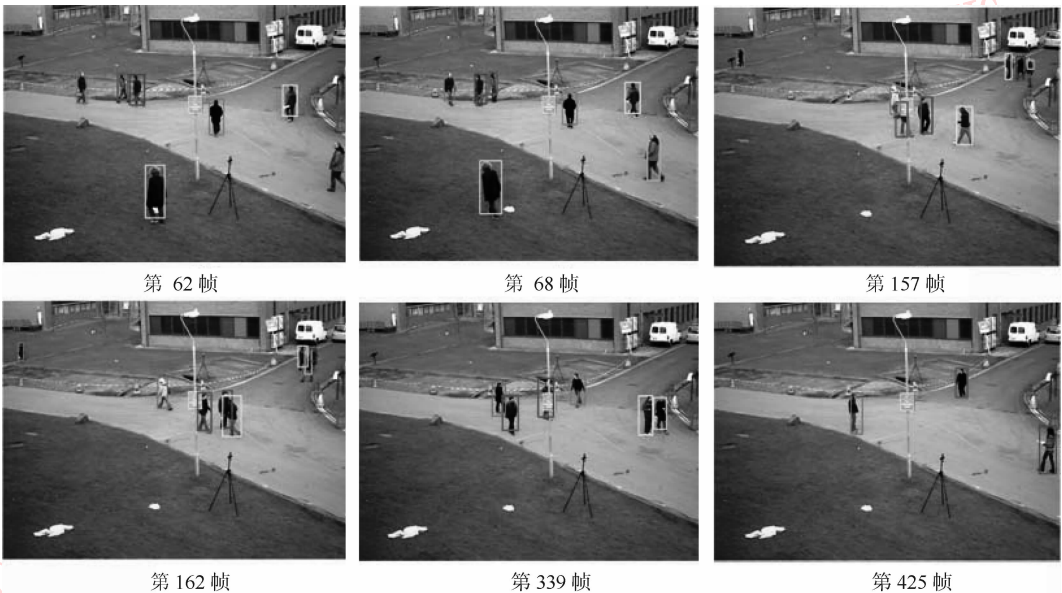


图 7 跟踪结果(PETS 2009 数据集, 任务 S2, Video L1, View1)

Fig. 7 Tracking results (PETS 2009, Task S2, Video L1, View1)

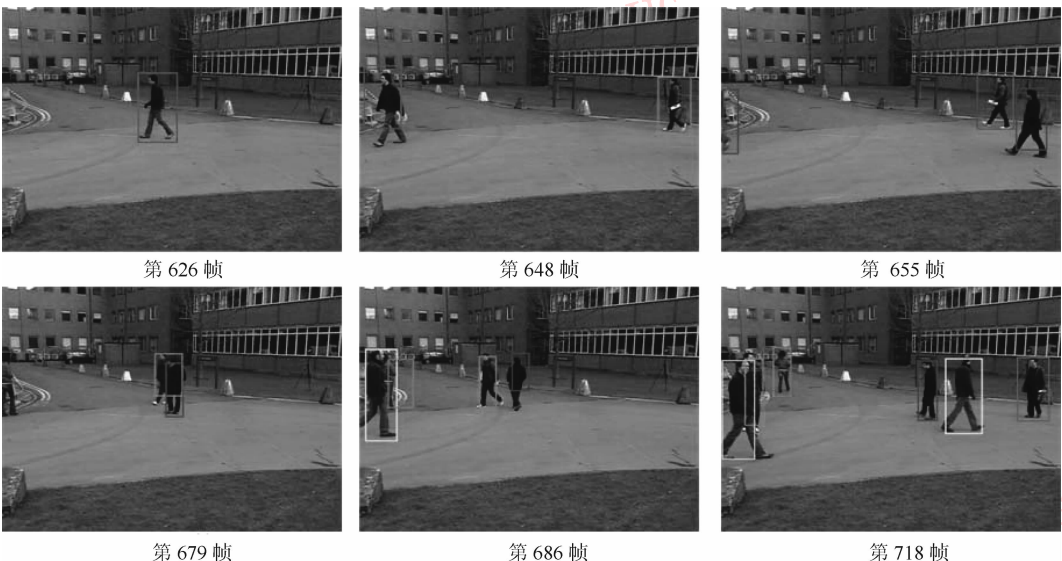


图 8 跟踪结果(PETS 2009 数据集, 任务 S2, Video L1, View5)

Fig. 8 Tracking results (PETS 2009, Task S2, Video L1, View5)

从图7可以看出,当目标不存在遮挡及交互时,能够跟踪目标(第425帧);第62—68帧中,左上角有两目标交叉,本文算法能够准确跟踪;第157、339帧中,一个目标被静态标志牌遮挡,通过粒子滤波器概率外观模型中保存的历史信息,可以判断出目标的位置;第162帧中,被遮挡的目标重新出现,能够继续跟踪,同时两个外观相似的目标有交互行为,跟踪假设也能较准确地判断出两目标的位置。

由图8可看出,摄像机镜头角度偏低,目标之间遮挡更为严重。如第679—686帧,第700—720帧都出现了大范围的遮挡。从实验结果上来看,本文提出的方法对短时间的遮挡有比较好的适应性。

采取文献[21]提出的评价体系对实验结果进行定量分析,并与最常用的颜色直方图模型、改进的概率外观模型^[7]以及基于K-均值遮挡分割的多目标跟踪方法在相同测试集中得到的跟踪结果进行比较。结果如图9所示。

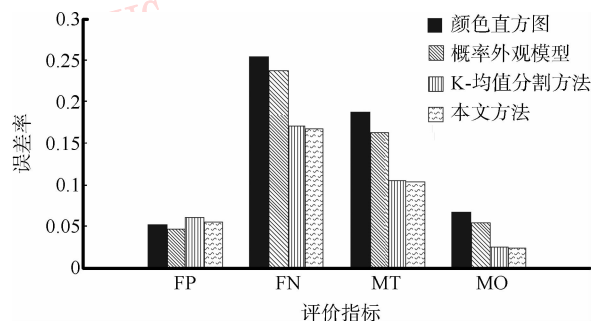


图9 跟踪误差分析

(PETS 2009 数据集,任务 S2, Video L1, View1)

Fig.9 Tracking error analysis

(PETS 2009, Task S2, Video L1, View1)

跟踪效果评价的主要参数是FP(false positives)、FN(false negatives)、MO(multiple objects)、MT(multiple trackers),其中FP指误跟,即一个跟踪框没有跟踪到目标;FN指漏跟,即一个目标没有被任何一个跟踪框跟踪到;MO指额外目标,即一个跟踪假设同时跟踪到两个或两个以上的目标;MT指多个跟踪假设,即两个或两个以上跟踪假设同时跟踪一个目标。由图9及表4可以看出,概率外观模型由于在颜色信息基础上加入了空间分布信息,错误率比直接使用颜色分布低,而本文提出的方法从FN、MT、MO三个方面获得了提高。这是因为本文方法基于运动目标的骨架点进行遮挡分割,获得了目标更为准确的运动前景图像,从而减少了一个跟踪假设同时跟踪多个

目标(MO)及跟踪假设目标漂移(MT)的情况;同时,由于得到了被遮挡目标更为准确的信息,减少了漏跟率(FN);但是,从实验结果看,本文方法的误跟率比原有方法稍高,产生这种结果的原因是由于对运动前景进行划分,在某些情况下,当某个目标运动不大或两个目标光流颜色、运动、速度信息均相近时,会产生遮挡分割不准确的情况,进而引发跟踪错误,提高了误跟率(FP)。

表4 跟踪误差分析 (PETS 2009, 任务 S2, Video L1)

Table 4 Tracking error analysis (PETS 2009, Task S2, Video L1)

| 场景 | 跟踪方法 | FP | FN | MT | MO |
|------|----------|--------------|--------------|--------------|--------------|
| 场景 1 | 颜色直方图 | 0.052 | 0.253 | 0.187 | 0.065 |
| | 概率外观模型 | 0.046 | 0.235 | 0.162 | 0.054 |
| | K-均值分割方法 | 0.061 | 0.169 | 0.104 | 0.026 |
| | 本文方法 | 0.055 | 0.167 | 0.104 | 0.024 |
| 场景 5 | 颜色直方图 | 0.065 | 0.283 | 0.257 | 0.175 |
| | 概率外观模型 | 0.058 | 0.279 | 0.245 | 0.164 |
| | K-均值分割方法 | 0.066 | 0.269 | 0.201 | 0.086 |
| | 本文方法 | 0.062 | 0.266 | 0.199 | 0.091 |
| 场景 7 | 颜色直方图 | 0.055 | 0.258 | 0.252 | 0.085 |
| | 概率外观模型 | 0.042 | 0.251 | 0.244 | 0.082 |
| | K-均值分割方法 | 0.051 | 0.250 | 0.215 | 0.073 |
| | 本文方法 | 0.048 | 0.248 | 0.211 | 0.068 |
| 场景 8 | 颜色直方图 | 0.051 | 0.287 | 0.198 | 0.094 |
| | 概率外观模型 | 0.052 | 0.280 | 0.187 | 0.087 |
| | K-均值分割方法 | 0.055 | 0.272 | 0.164 | 0.045 |
| | 本文方法 | 0.053 | 0.276 | 0.162 | 0.042 |

6 结论

将运动分割的方法引入目标跟踪领域,提出一种融合骨架点指派遮挡分割的多运动目标跟踪方法。利用图像底层的光流速度信息,提取出运动目标的骨架点,再充分利用不同目标颜色、外观以及运动信息等高级语义特征,采用将骨架点指派给各目标,最后对所有运动前景像素密集分类,较好地完成了多目标跟踪过程中需要的遮挡分割任务,并在粒子滤波跟踪器框架下,实现多目标跟踪。本文提出的方法在国际通用的测试数据集中进行实验,取得了很好的跟踪效果,能够满足复杂场景下多目标跟

踪任务,尤其对目标间动态遮挡问题有较好的效果。与现有工作相比,本文使用底层运动信息得到骨架点并与目标高级信息结合,完成遮挡分割,得到目标更准确的观测外观,提高跟踪鲁棒性,同时本文方法重点处理骨架点指派问题,比现有的运动分割方法在执行效率上有较大提高。具体贡献主要体现在以下3个方面:1)定义了骨架点匹配方式及状态转移矩阵,判断骨架点类型;2)定义了基于目标外观、颜色、运动等高级语义信息的骨架点指派方法;3)以骨架点为核,将运动前景进行密集分类。经多组实验证明,本文提出的方法能够在两个以至多个目标之间相互遮挡的情况下取得很好的遮挡分割效果,能够改进单目固定摄像头下多目标跟踪的效果。

实验过程中发现,本文方法仍存在的不足及改进思路如下:1)遮挡分割时,由于光流计算方法的限制及目标相互的影响,物体光流速度计算不完全准确,影响遮挡分割效果,需要探索更为鲁棒的运动信息获取途径;2)通过本文方法可以估计被遮挡的骨架点坐标,包含目标被遮挡的信息,在后续工作中可以从这方面入手,对运动目标被遮挡部分进行补偿,进一步提升跟踪效果。

志谢 感谢 Sand P 博士提供用于文献[19]中的光流代码。

参考文献 (References)

[1] Reid D. An algorithm for tracking multiple targets [J]. IEEE Transactions on Automatic Control, 1979, 24(6) : 843-854.

[2] Bar-Shalom Y, Fortmann T, Scheffe M. Joint probabilistic data association for multiple targets in clutter [C] // Proceedings of 1980 International Conference of Information Sciences and Systems, 1980.

[3] Khan S M, Shah M. Tracking multiple occluding people by localizing on multiple scene planes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(3) : 505-519.

[4] Zhao T, Nevatia R. Tracking multiple humans in complex situations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(9) : 1208-1221.

[5] Xu L, Puig P. A hybrid blob-and appearance-based framework for multi-object tracking through complex occlusions [C] // Proceedings 2nd Joint IEEE International Workshop on VS-PETS. Beijing, China; IEEE Press, 2005 : 73-80.

[6] Senior A, Hampapur A, Tian Y, et al. Appearance models for occlusion handling [J]. Image and Vision Computing, 2006, 24(11) : 1233-1243.

[7] Yang J, Vela P A, Shi Z. Probabilistic multiple people tracking through complex situations [C] // Proceedings of IEEE International Workshop on PETS. Miami, USA; IEEE Press, 2009 : 79-86

[8] Du W, Piater J. Tracking by cluster analysis of feature points using a mixture particle filter [C] // Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance. Como, Italy; IEEE Press, 2005 : 165-170.

[9] Bugeau A, Perez P. Detection and segmentation of moving objects in highly dynamic scenes. [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, USA; IEEE Press, 2007 : 2102-2109.

[10] Kristan M, Pers J, Kovacic S, et al. A local-motion-based probabilistic model for visual tracking [J]. Pattern Recognition, 2009, 42(9) : 2160-2168.

[11] Hai T, Sawhney H S, Kumar R. Object tracking with Bayesian estimation of dynamic layer representations [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(1) : 75-89.

[12] Qian Y, Medioni G. Multiple-target tracking by spatiotemporal Monte Carlo Markov chain data association [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009, 31(12) : 2196-2210.

[13] Papadourakis V, Argyros A. Multiple objects tracking in the presence of long-term occlusions [J]. Computer Vision and Image Understanding, 2010, 114(7) : 835-846.

[14] Papadakis N, Bugeau A. Tracking with occlusions via graph cuts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(1) : 144-157.

[15] Silva, L S, Scharcanski J. Video segmentation based on motion coherence of particles in a video sequence [J]. IEEE Transactions on Image Processing, 2010, 19(4) : 1036-1049.

[16] Stauffer C, Grimson W E L. Learning patterns of activity using real-time tracking [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8) : 747-757.

[17] Wang Z, Yang X, Xu Y, et al. Camshift guided particle filter for visual tracking [J]. Pattern Recognition Letters, 2009, 30(4) : 407-413.

[18] Isard M, Blake A. Condensation-Conditional density propagation for visual tracking [J]. International Journal of Computer Vision, 1998, 29(1) : 5-28.

[19] Sand P, Teller S. Particle video: long-range motion estimation using point trajectories [J]. International Journal of Computer Vision, 2008, 80(1) : 72-91.

[20] Nummiaro K, Koller-meier E, Gool L V. An adaptive color-based particle filter [J]. Image and Vision Computing, 2003, 21(1) : 99-110.

[21] Smith K, Gatica-Perez D, Odobez J, et al. Evaluating multi-object tracking [C] // Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. San Diego, California USA; IEEE Press, 2005 : 36-43.