



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
09
VOL.18

ISSN1006-8961
CN11-3758/TB



城市交通规划 P1197



第18卷第9期（总第209期）

2013年9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 010-82614429

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 010-82614429

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

综述

高分辨率遥感影像空间结构特征建模方法综述

秦昆, 陈一祥, 甘顺子, 冯霞, 任文力

1055

简缩极化SAR数据信息提取与应用

张红, 谢镭, 王超, 张波, 吴樊, 汤益先

1065

图像处理和编码

融合CCCTI码和谱系聚类的仿射配准

张桂梅, 江少波, 刘丕玉, 张松

1074

面向虚拟视点绘制的深度图滤波及上采样方法

高凯, 陈贺新, 赵岩, 耿英楠, 王钢

1085

图像分析和识别

包空间多示例图像自动分类

王科平, 杨艺, 王新良

1093

表情和姿态的双模态情感识别

闫静杰, 郑文明, 辛明海, 邱伟

1101

等级关联的粒子群迭代多目标检测跟踪

程旭, 李拟珺, 张索非, 吴镇扬

1107

改进后的TLD视频目标跟踪方法

周鑫, 钱秋朦, 叶永强, 王从庆

1115

多特征稳健主成分分析的视频运动目标分割

甘超, 王莹, 王向阳

1124

融合全局最好和声搜索算法的模糊C均值聚类图像分割

崔兆华, 高立群, 欧阳海滨, 李文娜

1133

从运动中恢复目标结构的改进因子分解法

邱少华, 文贡坚, 回丙伟, 张鹏

1142

计算机图形学

并行扫描转换结构中的状态管理

殷诚信, 韩俊刚, 黄虎才

1149

医学图像处理

3维颅骨表面模型的复杂孔洞修补

包佳蕊, 梁荣华, 吴福理, 黄鲜萍, 蒋莉, 陈矛

1156

归一化Radon变换和统计测度的超声图像肌束方向自动跟踪

李乔亮, 李振宇, 张会生, 张小飞, 陈昕, 陈思平, 汪天富

1164

遥感图像处理

改进的多光谱双边滤波图像融合

吴川, 杨冬

1170

长度归一化扫描的合成孔径雷达图像船舶尾迹恒虚警率检测

南杰, 张红, 王超, 吴樊, 张波

1176

地理信息技术

3D可视化输电线路中杆塔空间挂点计算

何国辉, 周聪

1184

地图符号服务与地图服务的耦合

焦东来, 张海涛, 闫国年, 姜杰

1190

不同表达粒度对城市路网结构健壮性评价的影响

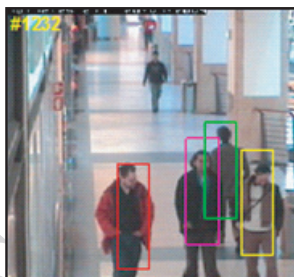
段滢滢, 陆锋

1197

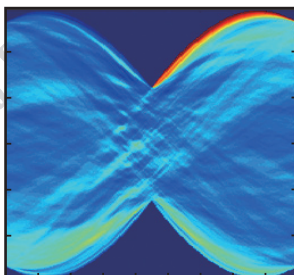
DEM保真精度量化描述模型

张锦明, 游雄

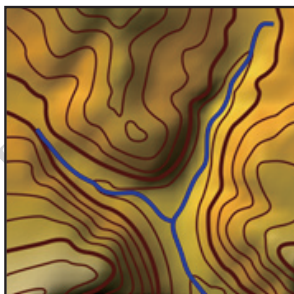
1206



等级关联的粒子群迭代多目标检测跟踪(第1107页)



归一化Radon变换和统计测度的超声图像肌束方向自动跟踪(第1164页)



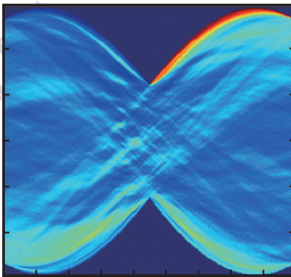
DEM保真精度量化描述模型(第1206页)

CONTENTS

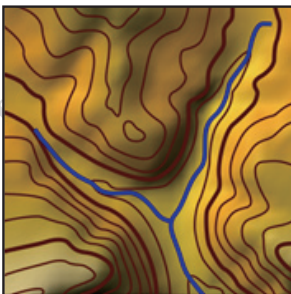
JOURNAL OF IMAGE AND GRAPHICS



Hierarchical associative multi-target detection and tracking with particle swarm optimization(P1107)



Automatic tracking of muscle fascicle orientation in ultrasound images: a approach based on a normalized Radon transform and statistical measure (P1164)



Quantitative descriptive model of DEM fidelity accuracy(P1206)

Review

Review on methods of spatial structural feature modeling of high resolution remote sensing images

Qin Kun, Chen Yixiang, Gan Shunzi, Feng Xia, Ren Wenli 1055

Information extraction and application of compact polarimetric SAR data

Zhang Hong, Xie Lei, Wang Chao, Zhang Bo, Wu Fan, Tang Yixian 1065

Image Processing and Coding

Affine registration based on CCCTI and hierarchical clustering

Zhang Guimei, Jiang Shaobo, Liu Piyu, Zhang Song 1074

Depth map filtering and upsampling method for virtual view rendering

Gao Kai, Chen Hexin, Zhao Yan, Geng Yingnan, Wang Gang 1085

Image Analysis and Recognition

Automatic classification of multiple-instance image based on the bag space

Wang Keping, Yang Yi, Wang Xinliang 1093

Bimodal emotion recognition based on body gesture and facial expression

Yan Jingjie, Zheng Wenming, Xin Minghai, Qiu Wei 1101

Hierarchical associative multi-target detection and tracking with particle swarm optimization

Cheng Xu, Li Nijun, Zhang Suofei, Wu Zhenyang 1107

Improved TLD visual target tracking algorithm

Zhou Xin, Qian Qiumeng, Ye Yongqiang, Wang Congqing 1115

Multi-feature robust principal component analysis for video moving object segmentation

Gan Chao, Wang Ying, Wang Xiangyang 1124

Improved fuzzy C-means clustering combined with the global best harmony search algorithm for image segmentation

Cui Zhaohua, Gao Liqun, Ouyang Haibin, Li Wenna 1133

Improved factorization method for structure from motion

Qiu Shaohua, Wen Gongjian, Hui Bingwei, Zhang Peng 1142

Computer Graphics

Efficient state management for parallel rasterization

Yin Chengxin, Han Jungang, Huang Hucai 1149

Medical Image Processing

Various hole repairing algorithm for 3D mesh surface models of human skull

Bao Jiarui, Liang Ronghua, Wu Fuli, Huang Xianping, Jiang Li, Chen Mao 1156

Automatic tracking of muscle fascicle orientation in ultrasound images: a approach based on a normalized Radon transform and statistical measure

Li Qiaoliang, Li Zhenyu, Zhang Huisheng, Zhang Xiaofei, Chen Xin, Chen Siping, Wang Tianfu 1164

Remote Sensing Image Processing

Improved multispectral bilateral filter video fusion algorithm

Wu Chuan, Yang Dong 1170

Ship wake constant false alarm rate detection algorithm in synthetic aperture radar images based on length normalized scanning

Nan Jie, Zhang Hong, Wang Chao, Wu Fan, Zhang Bo 1176

Geoinformatics

Tower hanging point calculation for transmission line on three-dimensional visualization

He Guohui, Zhou Cong 1184

Coupling of map symbol web service and web map service

Jiao Donglai, Zhang Haitao, Lv Guonian, Jiang Jie 1190

The impact of different granularity representations on robustness evaluation of city road network

Duan Yingying, Lu Feng 1197

Quantitative descriptive model of DEM fidelity accuracy

Zhang Jinming, You Xiong 1206

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2013)09-1124-09

论文引用格式: 甘超, 王莹, 王向阳. 多特征稳健主成分分析的视频运动目标分割[J]. 中国图象图形学报, 2013, 18(9): 1124-1132. [DOI: 10.11834/jig.20130909]

多特征稳健主成分分析的视频运动目标分割

甘超, 王莹, 王向阳

上海大学通信与信息工程学院, 上海 200072

摘要: 提出一种多特征稳健主成分分析(MFRPCA)算法, 该算法融合多种视觉特征进行视频运动目标分割, 分割的目的即将运动目标从静止信息中提取出来, 分割的主要过程是将多特征视频矩阵分解为低秩矩阵和稀疏矩阵。矩阵分解过程是求解一个带受限条件的核范数与 $L_{2,1}$ 范数组合的最小化问题, 此最小化问题可以通过增广拉格朗日乘子法(ALM)有效求解。与其他算法相比, 本文算法融合了图像的颜色、边缘和纹理特征等多个特征, 通过对变化检测基准数据集进行检测, 本文算法获得的查全率为 0.486 0 和 F 度量为 0.559 7, 实验结果表明, 本文算法的稳健性和可靠性均优于其他算法。

关键词: 视频分割; 稀疏和低秩矩阵; 多特征稳健主成分分析; 增广拉格朗日乘子法; 核范数最小化

Multi-feature robust principal component analysis for video moving object segmentation

Gan Chao, Wang Ying, Wang Xiangyang

College of Communication & Information Engineering, Shanghai University, Shanghai 200072, China

Abstract: In this paper, we propose an algorithm named Multi-feature Robust Principal Component Analysis (MFRPCA), to integrate multiple visual features for Video Moving Object Segmentation. The aim of Video Moving Object Segmentation is to separate the moving objects from the static information. Its main process is to decompose the multiple feature video matrices into low rank and sparse matrices. The decomposition is to solve a minimization problem formulated as a constrained combination of nuclear norm and $L_{2,1}$ norm, which can be solved efficiently by Augmented Lagrange Multiplier (ALM) method. Compared to other methods developed recently, the proposed method integrates color, edge and texture features. The quantitative results for Recall and F-measure obtained from experiments on the change detection benchmark dataset are 0.486 0 and 0.559 7 respectively. The results which outperform other methods well show that the proposed MFRPCA can achieve more robust and reliable performance.

Key words: video segmentation; sparse and low rank; multi-feature robust principle component analysis; augmented Lagrange multipliers; nuclear norm minimization

收稿日期: 2012-10-29; 修回日期: 2013-02-27

基金项目: 国家自然科学基金项目(60975024)

第一作者简介: 甘超(1986—), 男, 2013 年于上海大学获电路与系统专业工学硕士学位, 主要研究方向为计算机视觉, 模式识别。

E-mail: kobeganchao@163.com

通讯作者: 王向阳, E-mail: wangxiangyang@shu.edu.cn

0 引言

目前,对视频序列的分析和理解是一个相当活跃的研究领域。在诸如视频监控、基于视觉的人体行为分析等许多应用中,首先需要的是能够探测到由静态相机拍摄的一个场景中的物体的移动。最基本的要求就是从静态信息(背景)中分离出运动的物体(前景)。视频运动目标分割就是这样的一类方法,是自动视频分析的关键技术之一,尤其是在视频监控领域里,它是静态视频处理的第一步,它的输出结果的好坏对后续处理,如目标识别、追踪和行为识别起着重大的影响。

近年来,共5种基于稳健主成分分析(RPCA)理论的方法被提出用来进行视频运动目标分割:RPCA via Robust Subspace Learning (RSL)^[34], RPCA via

Principal Component Pursuit(PCP)^[5], RPCA via Inexact Augmented Lagrange Multiplier (IALM)^[6], RPCA via Templates for First-Order Conic Solvers(TFOCS)^[7]和RPCA via Bayesian Robust Principal Component Analysis(BRPCA)^[8]。一般情况下,上述方法非常有效,但它们都只考虑了一个特征描述子。

作为一种数据聚类问题,视频运动目标分割的性能取决于特征空间的选择。然而,很难找到一个单一的特征对不同场景的各种视频图像都能进行较好地描述,因为每个特征描述子一般都有各自的优点和局限性:颜色描述子在描述颜色时能携带较多的信息,但不适合描述其他的视觉信息;梯度描述子能较好的保留目标的边缘和轮廓,但可能会造成信息丢失。因此,整合不同类型的特征对获得精确、可靠的前景是非常关键的解决方法。图1说明了融合多个特征的优越性。



图1 融合多特征的必要性和优越性

Fig.1 Illustration of the necessity and superiority of fusing multiple types of features

为了有效地利用多特征,将视频运动目标分割看成稀疏追寻问题,提出了一种新的算法多特征稳健主成分分析(MFRPCA),通过恢复多特征描述矩阵的低秩和稀疏部分来进行多特征融合。给定一个

由多特征描述的图像(视频)序列,背景通常被认为是低秩的,而前景占整个图像的少部分,则被认为是稀疏的。前景就是通过将多特征联合矩阵分解为一系列的低秩和稀疏矩阵后的稀疏部分,该推导过程

就是一个受限的核范数与 $L_{2,1}$ 范数的最小问题, 为一个凸优化问题, 并且可以由拉格朗日乘子法 (ALM) 有效地进行求解。与近期提出的方法相比, 本文方法在单一的推理步骤中融合图像的颜色、边缘和纹理特征等多个特征, 从而产生更准确、更可靠的结果。

1 相关工作

1.1 主成分分析

主成分分析 (PCA) 是一种高效的高维数据处理、分析、压缩和可视化工具, 它是一种通过降维技术把多个变量化为少数几个主成分 (即综合变量) 的统计分析方法。这些主成分能够反映原始变量的绝大部分信息, 它们通常表示为原始变量的某种线性组合, 假设给定的高维数据实质上可由低维线性子空间进行建模, PCA 的目标就是高效、准确地估计低维子空间。

假设所给的数据是组成一个大矩阵 $D \in \mathbf{R}^{m \times n}$ 的列。寻找低维子空间的数学模型是找到一个低秩矩阵 A , 使得 A 和 D 之间的差异最小化, 从而变成约束优化问题

$$\min_{A,E} \|E\|_F \text{ 使得 } D = A + E, \text{rank}(A) \leq r \quad (1)$$

式中, E 为误差矩阵 $r \ll \min(m, n)$ 是子空间的目标维数, $\|\cdot\|_F$ 就是矩阵平方和的平方根。如果定义两个矩阵之间的欧几里得内积为: $\langle X, Y \rangle = \text{tr}(X \times Y)$, 那么 $\|X\|_F^2 = \langle X, X \rangle$ 。该最小化问题可以通过奇异值分解 (SVD) 方便地得到解决^[6]。

PCA 已经成为计算机数据分析的基本操作。它与低维子空间的数据相符合。然而, 由于它对异常或损坏的观察数据矩阵缺乏稳健性, 因此, 在真实场景中的性能和适用性有限。

传统的 PCA 在同一分布的高斯噪声的损坏下给出了最优估计, 只要噪声的幅度比较小, 工作效果就比较理想。但是, 在大的噪声损坏下, 即使只影响了小部分的观察数据, 其效果也比较差。然而, 如异常数据和大幅度噪声等损坏原始数据的情况普遍存在于各种图像和视频, 有些数据甚至可能由于遮挡被任意破坏, 或者与设法确定的低维结构根本不相干^[5]。因此, 有必要对当附加误差 E 可能是任意大的条件下, 是否可以从受损数据 $D = A + E$ 中恢复出低秩矩阵 A 进行了解和研究。

1.2 稳健主成分分析

稳健主成分分析通过解决凸优化问题对损坏的低秩矩阵的进行恢复^[9]。Candes 等人^[5]已证明了在一些必要的条件下精确地恢复低秩矩阵可能性。

给定一个矩阵 D , RPCA 的目标是通过最小化

$$\min_{A,E} \text{rank}(A) + \lambda \|E\|_0 \text{ 使得 } D = A + E \quad (2)$$

这个约束条件的问题把 D 分解成一个低秩矩阵 A 和稀疏矩阵 E 。式中, $\|\cdot\|_0$ 是 L_0 范数, 它是指矩阵的非零项的数目, 即表示矩阵的稀疏性。但是, 这是一个难以解决的高度非凸优化问题。幸运的是, 可以通过用核范数和 L_1 范数分别取代秩的最小化和 L_0 范数, 即

$$\min_{A,E} \|A\|_* + \lambda \|E\|_1 \text{ 使得 } D = A + E \quad (3)$$

式中, $\|A\|_*$ 表示矩阵的核范数 (一个矩阵的奇异值的总和)^[10], $\|\cdot\|_1$ 是 L_1 范数, 参数 $\lambda > 0$ 是一个正的加权参数, 用于权衡这两部分的比重。

Wright 等人^[59]已经证明, 只要误差矩阵 E 足够稀疏 (相对于 A 的秩), 就有可能通过解凸优化问题式 (3) 准确地恢复出低秩和稀疏部分, 也就是最小化核范数和 L_1 范数的加权组合。

1.3 稳健主成分分析的应用

在图像和视频处理方面, 稳健主成分分析 (RPCA) 应用广泛, 如背景建模、阴影消除、人脸图像识别^[9]、批量图像校准^[11]、纹理不变性表示^[12]、稳健光度立体^[13]、摄像机标定^[14]、3 维重建^[15]。

对于视频运动目标分割的应用, 由于帧与帧之间的相关性, 可以很自然的将背景近似建模为低秩部分。前景一般只占据图像像素的小部分, 并可以被视为稀疏误差。在文献^[5, 9]已证明, 即使在光照变化的情况下, RPCA 仍然可以通过低秩和稀疏的分离得到更好的视觉效果。结果表明凸优化方法作为一种工具, 在视频分析中具有潜在能力。

2 MFRPCA 算法

图 2 为提出的 MFRPCA 算法流程框图: 给定一组视频序列, 对每帧图像进行多尺度特征提取, 包括颜色、梯度、Gabor 滤波器特征, 形成一个多维特征矩阵。然后用 MFRPCA 算法对多维特征矩阵分解为低秩和稀疏两部分, 其中稀疏部分即为运动目标等前景。

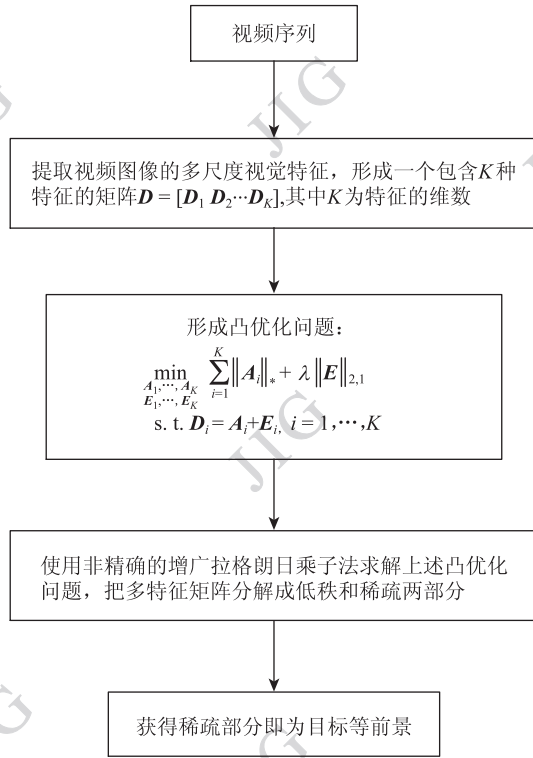


图2 MFRPCA 的视频运动目标分割流程图

Fig. 2 The flow chart of video moving object segmentation by MFRPCA

2.1 问题阐述

对于多特征, 视频运动目标分割问题可归结如下:

定义 1 假设 $D_1, D_2, \dots, D_K$ 为表示 K 种特征的特征矩阵。 D_i 的大小为 $d_i \times N$, d_i 是特征维数, N 是像素的数目。背景分割任务就是通过融合多特征 $D_1, D_2, \dots, D_K$ 将特征矩阵分解为低秩部分 (对应于背景) 和稀疏部分 (对应于前景的运动)。

2.2 多特征稳健主成分分析

在定义 1 中, MFRPCA 通过解决凸优化问题来寻找联合稀疏矩阵 E , 即

$$\min_{A_1, \dots, A_K, E_1, \dots, E_K} \sum_{i=1}^K \|A_i\|_* + \lambda \|E\|_{2,1} \quad \text{s.t.} \quad D_i = A_i + E_i, i = 1, \dots, K \quad (4)$$

式中, $E = [E_1, \dots, E_K]$ 是通过矩阵 $E_1, \dots, E_K$ 的列垂直连接形成的, $\|\cdot\|_{2,1}$ 表示 $L_{2,1}$ 范数, 即矩阵的列的 L_2 范数的和: $\|E\|_{2,1} = \sum_j \sqrt{\sum_i ([E]_{ij})^2}$, $[E]_{ij}$ 是矩阵 E 的第 i 行、第 j 列的项。由于最小化 $L_{2,1}$ 范数能够使矩阵 E 的列为 0, 多特征的整合就可以通过最小化式 (4) 进行无缝实现。即方法将使 $[E_1, E_2, \dots, E_K]$ 的列有联合一致的幅度, 如同时最

大或者最小。图 3 阐明了多特征的思想。

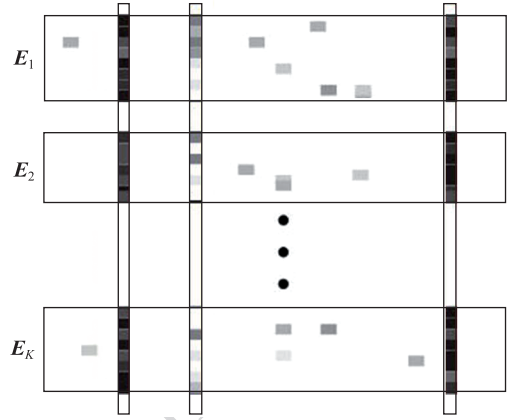
图3 矩阵 E 的 $L_{2,1}$ 范数的最小化Fig. 3 Illustrating the minimization of the $L_{2,1}$ norm on E

图 3 所示方法使得矩阵 $E_1, \dots, E_K$ 的项有联合稀疏一致的列。由于不同矩阵的同一下标的列对应一个图像像素点, 该方法也能使得不同的特征产生一致的结果。因此, 交叉特征信息自然而然的建模出来, 多个特征也将共同起作用。注意到, 式 (4) 与将 RPCA 单独应用于每个特征矩阵 D 的方法不同, 即

$$\min_{A_1, \dots, A_K, E_1, \dots, E_K} \sum_{i=1}^K (\|A_i\|_* + \lambda \|E_i\|_1) \quad \text{s.t.} \quad D_i = A_i + E_i, i = 1, \dots, K \quad (5)$$

2.3 解决 MFRPCA 凸优化问题

通过拉格朗日乘法 (ALM) [6] 来解决函数的最小化问题, 即

$$L = \lambda \|E\|_{2,1} + \sum_{i=1}^K (\|A_i\|_* + \langle Y_i, D_i - A_i - E_i \rangle + \frac{\mu}{2} \|D_i - A_i - E_i\|_F^2) \quad (6)$$

式中, $Y_1, \dots, Y_K$ 是拉格朗日乘子, $\mu > 0$ 是一个惩罚参数。通过最小化式 (5) (如: 非精确 ALM), 可以用算法 1 来解决式 (4) 中的问题。

算法 1: 通过非精确的拉格朗日乘法解多特征稳健主成分分析问题

输入: 数据矩阵 $\{D_i\}$, 参数 λ, μ 取值为 $\frac{1}{2 \times k \sqrt{\min(m, n)}}$, 其中 k 为多特征矩阵的长度, m 和 n 为单特征矩阵的行数和列数

初始化: $A_i = 0, E_i = 0, Y_i = 0, \mu = 10^{-6}, \max_{\mu} = 10^{10}, \rho = 1.1, \varepsilon = 10^{-8}$ 。

While not converged do {

1) 固定其他值, 通过 $A_i = \arg \min_A \frac{1}{\mu} \|A\|_* + \frac{1}{2} \|D_i - A_i - E_i + \mu^{-1} Y_i\|_F^2$, 更新 $A_1, \dots, A_K$ 。

2) 固定其他值, 通过 $E = \arg \min_E \frac{\lambda}{\mu} \|E\|_{2,1} + \frac{1}{2} \|E - M\|_F^2$, 其中 M 通过垂直连接矩阵 $D_i - A_i + \frac{Y_i}{\mu}, i=1, \dots, K$ 的列形成, 更新 $E = [E_1, E_2, \dots, E_K]$ 。

3) 更新乘子 $Y_i = Y_i + \mu(D_i - A_i - E_i)$ 。

4) 更新参数 $\mu = \min(\rho\mu, \max_\mu)$ 。

5) 检查收敛条件 $\|D_i - A_i - E_i\|_\infty < \varepsilon, i=1, \dots, K$ 。

End While

输出: A, E

如算法 1 中所示, 步骤 1) 通过奇异值阈值操作子 (SVT) 来解^[16], 即将软阈值操作子用于奇异值。

软阈值算子定义为

$$\delta_\varepsilon = \begin{cases} x - \varepsilon & x > \varepsilon \\ x + \varepsilon & x < -\varepsilon \\ 0 & \text{其他} \end{cases} \quad (7)$$

式中, $x \in \mathbf{R}, \varepsilon > 0$ 。根据这种记法, 有

$$D_\varepsilon(W) := U \delta_\varepsilon[S] V^T = \arg \min_X \varepsilon \|X\|_* + \frac{1}{2} \|X - W\|_F^2,$$

$$\delta_\varepsilon[W] = \arg \min_X \varepsilon \|X\|_1 + \frac{1}{2} \|X - W\|_F^2 \quad (8)$$

式中, USV^T 是 W 的奇异值分解。

要求解算法 1 中的步骤 2), 先回顾一下文献[17]中的引理 3.2。

引理 3.2 $Q = [q_1, q_2, \dots, q_i, \dots]$ 是一个给定的矩阵, $\|\cdot\|_F$ 是 F-范数。假设

$$\min_W \lambda \|W\|_{2,1} + \frac{1}{2} \|W - Q\|_F^2 \quad (9)$$

的最优解为 W^* , 那么 W^* 的第 i 列是

$$W^*(:, i) = \begin{cases} \frac{\|q_i\| - \lambda}{\|q_i\|} q_i & \lambda < \|q_i\| \\ 0 & \text{其他} \end{cases} \quad (10)$$

式中, $\|q_i\|$ 是 L_2 范数。

步骤 2) 的解可以逐列的解决:

$$E(:, i) = \arg \min_E \frac{\lambda}{\mu} \|E\|_2 + \frac{1}{2} \|E - M(:, i)\|_2^2 \quad (11)$$

式中, $E(:, i)$ 是矩阵 E 的第 i 列, $i=1, \dots, N$ 。

$$E(:, i) = \begin{cases} \frac{\|M(:, i)\|_2 - \lambda}{\|M(:, i)\|_2} M(:, i) & \lambda < \|M(:, i)\|_2 \\ 0 & \text{其他} \end{cases} \quad (12)$$

3 实验

3.1 实验设置

3.1.1 数据集

在实验中, 使用变化检测基准数据集^[18], 目前, 包括热成像在内, 该数据集包含 31 个实时视频, 共有超过 80 000 帧、跨越 6 个类别的数据, 其中包含了不同的运动和变化检测的视频, 每个类别包含 4~6 个视频序列。有代表的数据集包括室内和室外的视觉捕获的数据监控和智能环境的数据。数据集包含了一套全面仔细的人工分割的标准结果 GT (ground truth) 以便能够进行各种不同算法的精确定量比较。文中使用的数据集覆盖 4 种视频类别 (基准、相机的抖动、阴影和热成像) 中的 6 个视频序列来同几个其他算法进行比较。对于每个视频序列, 使用了每个视频序列中的几百帧。对于单个不同特征之间的性能比较实验, 从以上 6 个视频序列所得的结果中随机选取一些结果进行定量比较, 一共选择了 273 帧。

3.1.2 特征

对于给定的图像, 提取 3 个不同类型的视觉特征来描述: 颜色、梯度和 Gabor 滤波器。

1) 颜色特征 从每个像素点中提取 R、G、B 3 个颜色值产生 3 个颜色特征, 每个特征称为特征矩阵。

2) 梯度特征 在向量的微积分中, 一个标量场的梯度是指向标量场的最大速率增加的方向的矢量场, 其大小是增长速度的幅度。简单而言, 空间的任何数量的变化 (如图形方式) 可以由一个斜率表示。梯度表示斜坡的陡度和方向。这里梯度被用于描述每个像素及其相邻像素中的相对倾斜的水平、获得目标物体的边缘轮廓。实验中通过 Prewitt 算子计算梯度。

3) Gabor 滤波器^[19] 在图像处理中, Gabor 滤波器是一种用来边缘检测的线性滤波器。Gabor 滤波器的频率及方向特征和人类数据系统相似。综合不同频率和方向的 Gabor 滤波器可以从图像中提取有用的特征, 已经发现它们特别适合描述纹理特征。

实验中使用的 Gabor 滤波器的参数如下: 波长 Wavelength($\lambda = 8$), 方向($\theta = 0$), 相位偏移($\varphi = [0, \pi/2]$), 纵横比($\gamma = 0.5$), 带宽($b = 1$), 方向数量($N = 8$)。

3.1.3 评估

为了证明融合多种特征的优势, 考虑算法采用不同特征情况下的结果: RPCA(颜色)、RPCA(梯度)、MFRPCA(梯度+颜色)和 MFRPCA(梯度+颜色+Gabor 滤波器)。从像素级上的表现来判断结果的好坏。用于衡量和评价前景和标准结果(GT)好坏的 4 个术语为

1) True Positive(TP) 前景像素点中那些被正确地标记为前景的像素数目。

2) False Positive(FP) 背景像素点中那些被错误地标记为前景的像素数目。

3) True Negative(TN) 背景像素点中那些被正确地标记为背景的像素数目。

4) False Negative(FN) 前景像素点中那些被错误地标记为背景的像素数目。

计算上述 4 个术语后, 接着计算 3 个重要的评价参数: 查全率(R)、精确率(P)和 F 度量(F)。

查全率计算前景像素点中那些被正确地标记为前景的像素数目占标准结果中的前景像素总数的比重, 即

$$R = \frac{TP}{TP + FN}$$

精确率计算前景像素点中那些被正确地标记为前景的像素数目占整个结果被认为是前景的像素总数的比重, 即

$$P = \frac{TP}{TP + FP}$$

F 度量则是查全率和精确率的一个综合, 是反应整个结果的好坏的一个综合指标, 即

$$F = 2 \times \frac{R \times P}{R + P}$$

3.2 实验结果

3.2.1 主要结果

如 3.1 节所述, 实验在统一配置的环境下对不同算法的性能进行评估和比较。实验结果见表 1 和表 2, 表 1 和表 2 中的结果证实了融合多特征的 MFRPCA 算法的优越性。与将 RPCA 用于某些单特征的方法相比, 表 1 中的结果清楚地显示功能融合多特征的优势。图 4 给出了一些实验结果, 也显示

了多特征融合的优势。图 5 为关于 false positive 和 truth positive 的 ROC(receiver operating characteristic curve)曲线, 该结果的实验数据为视频种类 baseline 下的 pedestrians 视频序列, 选取了 310 帧图像进行实验, 可以看出 MFRPCA 得出的结果最佳。图 6 则给出了表 2 的图示结果, 可以看出, 与其他的基于 RPCA 理论的方法相比, 由 MFRPCA 分割产生的结果是更可靠、准确的。

表 1 变化检测基准数据集上单类特征之间结果比较

Table 1 Results comparing between single features on change detection benchmark dataset

算法	平均结果		
	R	P	F
MFRPCA(颜色+梯度)	0.334 4	0.846 9	0.479 4
RPCA(颜色)	0.211 0	0.896 6	0.341 6
RPCA(梯度)	0.171 2	0.603 9	0.266 8

表 2 变化检测基准数据集上不同算法之间结果比较

Table 2 Results comparing between different algorithms on change detection benchmark dataset

算法	平均结果		
	R	P	F
MFRPCA(颜色+梯度+Gabor 滤波器)	0.483 0	0.853 0	0.559 7
MFRPCA(颜色+梯度)	0.399 9	0.938 8	0.529 5
IALM	0.346 4	0.940 5	0.480 3
BRPCA	0.474 1	0.891 4	0.555 0
PCP	0.386 9	0.918 2	0.521 0
RSL	0.381 6	0.893 2	0.521 7
TFOCS	0.163 6	0.967 9	0.250 6

3.2.2 与其他算法的比较

为了评价所提出的算法的优势, 与其他 5 种基于 RPCA 理论的算法进行比较, 分别为 RSL、PCP、IALM、TFOCS 和 BRPCA。在这些算法中, 将所有原始的图像转换为灰度图像。从表 2 和图 4 的平均结果来看, 本文算法优于其他算法, 而且对于不同的视频序列比其他方法具有更强的稳健性。正如表 2 中所示, 本文算法得到的查全率为 0.483 0、F 为 0.559 7, 在所有的算法中排在第 1。实验结果表明, 本文算法在用于视频运动目标分割时是可靠、准确和稳健的。

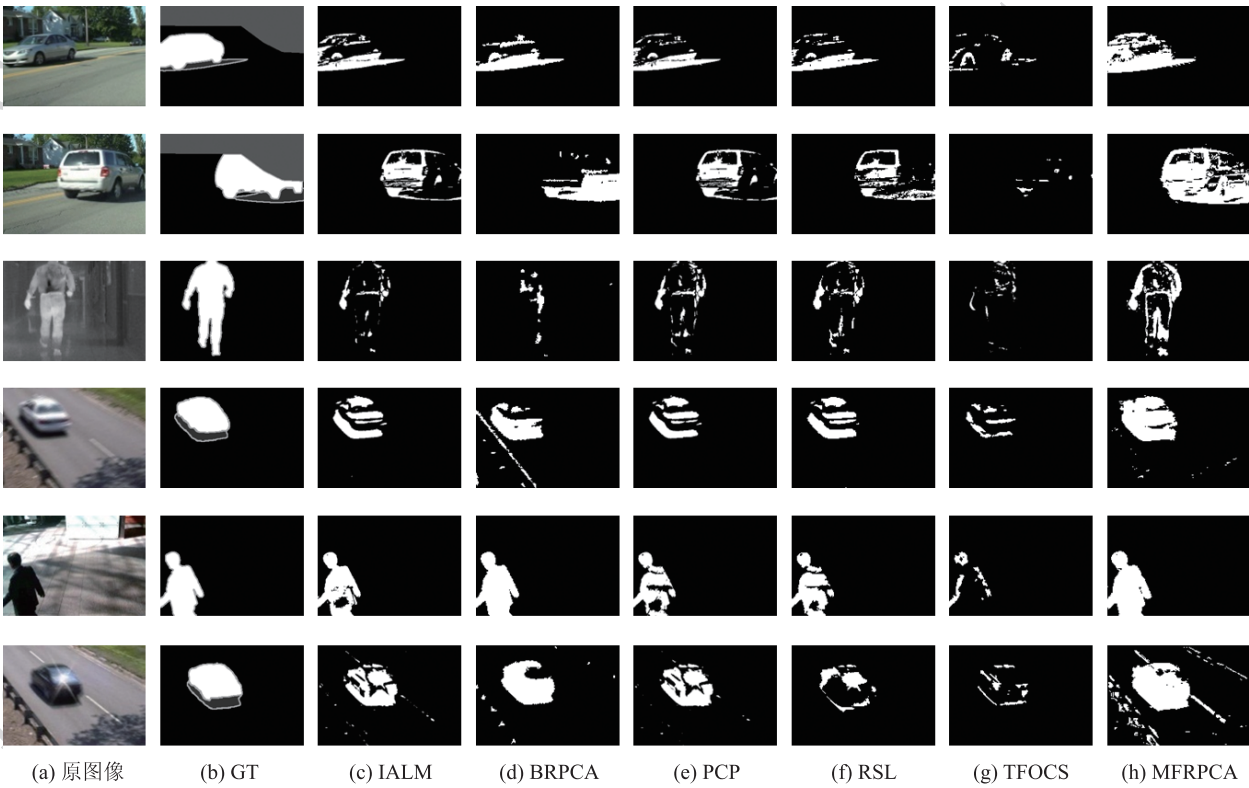


图 4 不同方法产生的结果
Fig. 4 Results by different methods

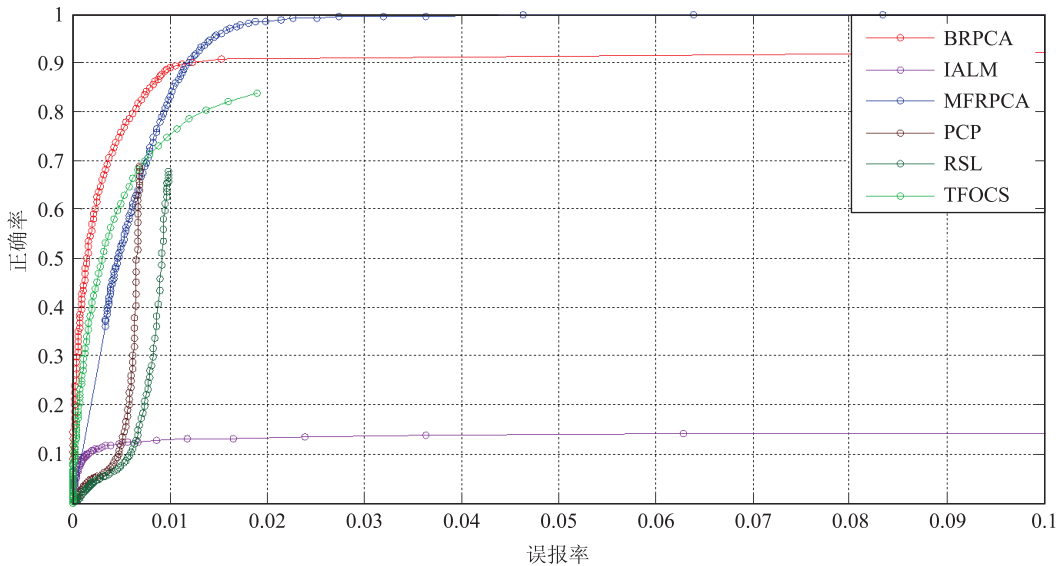


图 5 根据视频序列 pedestrians 的结果获得 ROC 曲线
Fig. 5 Roc curve of results from pedestrians

3.2.3 算法运行效率比较

本节将讨论不同算法的效率,为了统一其他客观因素,在同一台电脑上使用同样的数据对每种算法进行实验,实验数据为视频序列 pedestrians 下的

200 幅视频图像。图像的原尺寸大小为 240×360 , 统一缩小尺寸为 80×120 大小。表 3 列出了不同算法的运行所需时间。

实验中每种算法均为离线处理模式,这种以批

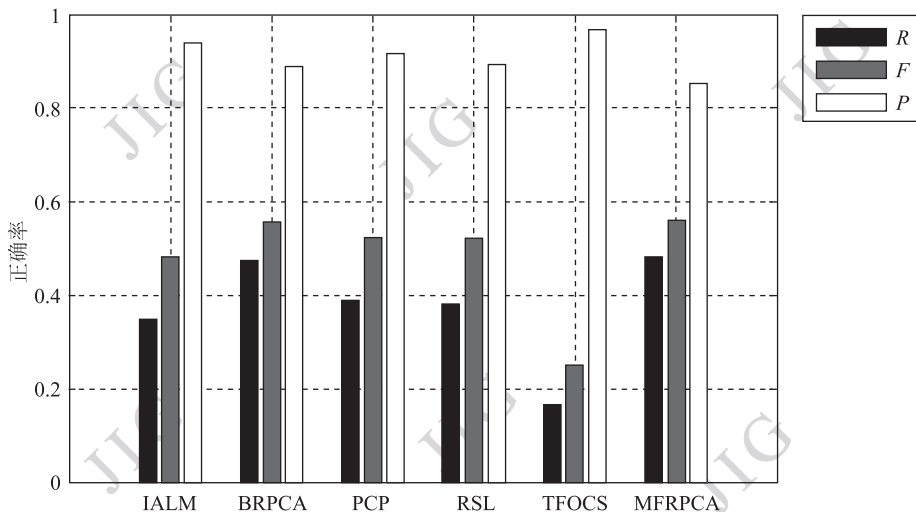
图6 在变化检测基准数据集上的获得的平均查全率(R)、 F 度量(F)和精确度(P)

Fig. 6 Average Recall, F-measure, Precision on change detection benchmark dataset

表3 不同算法运行所需时间比较

Table 3 Elapsed time comparing between different algorithms

算法	时间/s
MFRPCA(颜色 + 梯度 + Gabor 滤波器)	1 817.6
IALM	14.2
BRPCA	1 019.9
PCP	21.1
RSL	232.5
TFOCS	39.8

处理的离线分割模式对计算机内存的消耗有较高的要求,而本文算法是对多特征矩阵进行分割处理,对内存消耗有着更高的要求,当保存多特征矩阵消耗了大部分内存后,只有少部分内存用于进行算法计算等操作,导致计算速度变慢,其他算法对内存消耗较低,因此都能取得较快的计算速度。

目前出现了许多并行处理的方式来减少计算时间,利用硬件加速的方法来提高计算效率,如文献[20-21]。同时也产生了一些在线的处理方法,如文献[22],但是这种在线实时的方法需要事先进行训练获得背景,在这过程中会消耗较多的时间,如果不对背景进行训练,取得的分割效果会比较差,而且当背景发生变化后,就得重新训练背景,也会消耗较多的时间。

4 结 论

提出了一种新的用于视频运动目标分割的算

法,多特征稳健主成分分析。与其他方法相比,MFRPCA 将多种特征的信息融合到一个统一的推理过程中,可以有效地通过一个凸优化问题进行求解。该方法无缝融合了多种特征,在一个单一的推理步骤中联合的产生相似矩阵,从而产生更有效、更可靠的结果。

参考文献 (References)

- [1] Bouwmans T, Baf F E, Vachon B. Statistical background modeling for foreground detection: a survey[J]. Handbook of Pattern Recognition and Computer Vision, World Scientific Publishing, 2010, 4(2):181-189.
- [2] Bouwmans T, Baf F E, Vachon B. Background modeling using mixture of gaussians for foreground detection-a survey[J]. Recent Patents on Computer Science, 2008, 1(3):219-237.
- [3] Torre F D L, Black M. Robust principal component analysis for computer vision[C]// Proceedings of the International Conference on Computer Vision. Vancouver, Canada: University of British Columbia, 2001:362-369. [DOI:10.1109/ICCV.2001.937541]
- [4] Torre F D L, Black M. A framework for robust subspace learning[J]. International Journal on Computer Vision, 2003, 54(1-3):117-142. [DOI:10.1023/A:1023709501986]
- [5] Candes E J, Li X, Ma Y. Robust principal component analysis? [J]. Journal of the ACM, 2011, 58(3):1-37.
- [6] Lin Z, Chen M, Wu L. The augmented lagrange multiplier method for exact recovery of corrupted low-rank matrices: UIIU-ENG-09-2215[R]. University of Illinois at urbanu-Champaign, 2009.
- [7] Becker S, Candes E, Grant M. TFOCS: flexible first-order methods for rank minimization, low-rank matrix optimization symposium[C]//Proceedings of the SIAM Conference on Optimization

2011. Darmstadt, Germany: SIAM Activity Group on Computational Science and Engineering, 2011.
- [8] Ding X, He L, Carin L. Bayesian robust principal component analysis[J]. IEEE Transactions on Image Processing, 2011, 20(12):3419-3430. [DOI:10.1109/TIP.2011.2156801]
- [9] Wright J, Peng Y G, Ma Y. Robust principal component analysis: exact recovery of corrupted Low-rank matrices by convex optimization[C]//Proceedings of Advances in Neural Information Processing Systems. Vancouver British, Columbia Canada, 2009: 2080-2088.
- [10] Fazel M. Matrix rank minimization with applications[D]. Washington: Stanford University, 2002.
- [11] Peng Y G, Ganesh A, Wright J. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, Beijing, China: Tsinghua University, 2010: 763-770. [DOI:10.1109/TPAMI.2011.282]
- [12] Zhang Z D, Liang X, Ganesh A. TILT: transform Invariant low-rank textures[C]//Proceedings of Asian Conference on Computer Vision. Queenstown, New Zealand: Springer US, 2010:314-328.
- [13] Wu L, Ganesh A, Shi B X. Robust photometric stereo via low-rank matrix completion and recovery[C]//Proceedings of Asian Conference on Computer Vision. Queenstown, New Zealand: Springer US, 2010:703-717.
- [14] Zhang Z D, Maysushita Y, Ma Y. camera calibration with lens distortion from low-rank textures[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Beijing, China: Tsinghua University, 2011:2321-2328. [DOI: 10.1109 / CVPR.2011.5995548]
- [15] Mobahi H, Zhou Z H, Yang A Y. Holistic 3D reconstruction of urban structures from low-rank texture[C]//Proceedings of IEEE Conference on Computer Vision Workshops. Urbana, IL, USA: University of Illinois, 2011:593-600.
- [16] Cai J F, Candes E J, Shen Z W. A singular value thresholding algorithm for matrix completion[J]. SIAM Journal on Optimization, 2010, 20(4):1956-1982.
- [17] Liu G C, Lin Z C, Yu Y. Robust subspace segmentation by low-rank representation[C]// Proceedings of the International Conference on Machine Learning. Haifa, Israel: Omnipress, 2010: 663-670.
- [18] Goyette N, Jodoin P M, Porikli F. Change detection. net: a new change detection benchmark dataset [C]//Proceedings of IEEE workshop on Change Detection. Providence, RI:IEEE, 2012:1-8. [DOI: 10.1109/CVPRW.2012.6238919]
- [19] Feichtinger H G, Strohmer T. Gabor Analysis and Algorithms: Theory and Applications[M]. Boston: Birkhauser, 1998.
- [20] Mu, Y, Dong J, Yuan X. Accelerated low-rank visual recovery by random projection[C]//Proceedings of the 11th International Conference on Computer Vision. Beijing, China: IEEE Computer Society, University of Colorado Springs, 2011:2609-2616.
- [21] Anderson M, Ballard G, Demme J. Communication-avoiding QR decomposition for gpus[C]// IEEE International Parallel and Distributed Processing Symposium. Anchorage, Washington, IEEE Computer Society, 2011: 48-58. [DOI: 10.1109/IPDPS.2011.15]
- [22] He J, Balzano L, Szlam A. Incremental gradient on the grassmannian for online foreground and background separation in sub-sampled Video[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington: IEEE Computer Society, 2012:1568-1575.