

中图分类号: TP394.1 文献标识码: A 文章编号: 1006-8961(2015)11-1482-10

论文引用格式: Zhou Z W, Jin Z. Weighted nonconvex nuclear norm and its application in the moving target detection[J]. Journal of Image and Graphics 2015 20(11): 1482-1491. [周宗伟, 金忠. 非凸加权核范数及其在运动目标检测中的应用[J]. 中国图象图形学报, 2015, 20(11): 1482-1491.] [DOI: 10.11834/jlg.20151107]

非凸加权核范数及其在运动目标检测中的应用

周宗伟, 金忠

南京理工大学计算机科学与工程学院, 南京 210094

摘要: 目的 近年来, 低秩矩阵分解被越来越多的应用到运动目标检测中。但该类方法一般将矩阵秩函数松弛为矩阵核函数优化, 导致背景恢复精度不高; 并且没有考虑到前景目标的先验知识, 即区域连续性。为此提出一种结合非凸加权核范数和前景目标区域连续性的目标检测算法。方法 本文提出的运动目标检测模型以鲁棒主成分分析(RPCA)作为基础, 在该基础上采用矩阵非凸核范数取代传统的核范数逼近矩阵低秩约束, 并结合了前景目标区域连续性的先验知识。该方法恢复出的低秩矩阵即为背景图像矩阵, 而稀疏大噪声矩阵则是前景目标位置矩阵。结果 无论是在仿真数据集还是在真实数据集上, 本文方法都能够取得比其他低秩类方法更好的效果。在不同数据集上, 该方法相对于RPCA方法, 前景目标检测性能提升25%左右, 背景恢复误差降低0.5左右; 而相对于DECOLOR方法, 前景目标检测性能提升约2%左右, 背景恢复误差降低0.2左右。结论 矩阵秩函数的非凸松弛能够比凸松弛更准确的表征出低秩特征, 从而在运动目标检测应用中更准确的恢复出背景。前景目标的区域连续性先验知识能够有效地过滤掉非目标大噪声产生的影响, 使得较运动目标检测的精度得到大幅提高。因此, 本文方法在动态纹理背景、光照渐变等较复杂场景中均能够较精确地检测出运动目标区域。但由于区域连续性的要求, 本文方法对于小区域多目标的检测效果不甚理想。

关键词: 运动目标检测; 低秩矩阵分解; 非凸加权核范数; 区域连续性; 矩阵恢复

Weighted nonconvex nuclear norm and its application in the moving target detection

Zhou Zongwei, Jin Zhong

School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

Abstract: **Objective** Several low-rank matrix decomposition-based approaches have been proposed for moving object detection in recent years. However, most of these methods use the nuclear norm to substitute rank functions for optimization. As a result, the precision of background recovery is relatively low. Another problem is the failure of these methods to use prior knowledge of the regional continuity of foreground objects, which is important information for object detection. To solve these issues, we propose a novel object detection method that combines the weighted non-convex nuclear norm and the regional continuity of the foreground object. **Method** The new object detection model is designed on the basis of the robust

收稿日期: 2015-04-04; 修回日期: 2015-06-09

基金项目: 国家自然科学基金项目(61373063, 61233011, 61125305, 61220301, 61375007); 国家重点基础研究发展计划(2014CB349303)

第一作者简介: 周宗伟(1994—), 男, 现于南京理工大学攻读模式识别与智能系统专业硕士学位, 主要研究方向为图像处理与计算机视觉。E-mail: cumtzhouzongwei@gmail.com

Supported by: National Natural Science Foundation of China(61373063, 61233011, 61125305, 61220301, 61375007); National Basic Research Program of China(2014CB349303)

principal component analysis. The proposed model uses the weighted non-convex nuclear norm to replace the traditional nuclear norm for low-rank constraints. Furthermore, the prior knowledge of the regional continuity of the foreground object is added to restrain the clustered objects. By using this model, the recovered low-rank matrix becomes the background image matrix, and the large sparse noise matrix becomes the foreground object matrix. **Result** Experiments demonstrate that the proposed method outperforms other low-rank decomposition-based approaches in both the simulated data and real sequences. Specifically, the proposed methodology shows an increased projected target detection performance that is 25% and 2% greater than that of RPCA and DECOLOR. With respect to the two approaches, the proposed method reduces background recovery errors by about 0.5 and 0.2. **Conclusion** The non-convex relaxation of rank functions possesses better properties than the convex one in approximating matrix ranks, which is useful in restoring background images in motion object detection. The regional continuity of foreground objects allows the efficient exclusion of scattered outliers to enhance the effect of the objects detected. Therefore, this method can detect moving targets accurately in complex scenes, such as those with dynamic backgrounds and illumination-changing scenarios. However, the proposed method is not ideal for multi-object detection in small areas because of regional continuity requirements.

Key words: moving object detection; low-rank matrix decomposition; weighted nonconvex nuclear norm; regional continuous; matrix recovery

0 引言

智能视频分析在计算机视觉领域应用非常广泛,比如视频监控、汽车导航、交通控制等。智能视频分析包含3个基本步骤^[1]:目标检测、目标追踪和目标分析。其中目标检测是后续步骤的基础。

传统的目标检测方法包括光流法^[2-4]、帧差法^[5]、目标检测子法和背景差分法^[1]。光流法计算量较大,且易受噪声影响,不利于实时处理。帧差法简单易实现,但对运动场景稳定性要求较高,易受噪声影响,常出现“空洞”和“鬼影”现象。目标检测子法首先需要先在训练集上训练出特定目标的分类器,然后使用滑动窗口扫描每帧图像,并用分类器对窗口区域进行前景和背景的分类。该类方法能够有效地检测出固定目标,但分类器的预先训练限制了该类方法在实际应用中的推广。背景差分法首先构建出目标场景的背景图像,然后以每帧图像减除背景图像得到前景目标图像。该类方法能够同时检测出多种前景目标,且简单易实现。

背景差分法中背景图像的构建包含两种方法。一种方法是通过场景图像训练出背景图像^[6],这种方法要求训练集中不能包含运动目标,限制了该方法的实际应用。另一种方法是背景建模方法,这类方法包括单高斯模型^[7]、混合高斯模型^[8]、核密度估计模型^[9]、隐马尔可夫模型^[10]、线性自回归模型^[11]、码本模型^[12],以及 Ling 等人^[13]提出的基于

鲁棒弹性网络特征学习的方法等。

近些年提出的基于鲁棒主成分分析^[14](RPCA)的运动目标检测方法属于背景建模方法的一种。该方法利用运动视频序列中连续帧之间背景的强相关性构建出低秩背景矩阵恢复模型,而运动目标可以认为是低秩背景矩阵上的稀疏大噪声。鲁棒主成分分析的解决算法主要包括拉格朗日乘子法(ALM)、非精确拉格朗日乘子法(inexact ALM)和加速近端梯度法(APG)。该方法能够有效地克服背景环境周期性变化产生的目标误检,同时对光照变化等具有较好鲁棒性。

但是RPCA方法没有考虑到运动目标的先验知识,会把一些零散的非目标大噪声误判为前景目标区域。Zhou 等人^[15]认为运动目标具有区域连续性,只有满足区域连续性条件的稀疏大噪声区域才被认为是前景目标,因此 Zhou 等人在RPCA方法的基础上融入前景目标的区域性信息提出了基于低秩表示中连续野点检测的运动目标检测(DECOLOR)方法。该方法相对于RPCA方法大大提高了运动目标检测的正确率。

DECOLOR方法相对于其他常用的背景差分法能够同时分离运动场景的前景和背景,并且前景的检测效果较好,但是该方法时间复杂度和空间复杂度都较大,难以实现实时运动目标检测,因此考虑使用DECOLOR方法作为一种预处理方法。首先通过DECOLOR方法能够对运动背景图像进行建模,然后可以使用其他实时性背景减除类方法进行实时运

动目标检测。而 DECOLOR 方法中对运动背景的建模仍然可以改善。

RPCA 方法和 DECOLOR 方法中均使用矩阵的核范数去逼近矩阵的秩函数,但 Chen 等人^[16]指出非凸加权核范数能够比核范数更好的逼近矩阵的秩函数,即更准确地反映出连续图像帧之间运动背景的强相关性。因此,本文使用非凸加权核范数取代核范数去逼近秩函数,同时对前景目标加入区域连续性约束,提出了一种基于非凸加权核范数的运动目标检测方法。

1 相关知识

希望结合非凸加权核范数和 DECOLOR 方法的优势,找到一种能够同时准确分离出运动目标和背景图像的方法。因此,首先给出 DECOLOR 方法和非凸加权核范数的相关介绍。

1.1 DECOLOR 方法

将由 n 帧图像组成的观测矩阵表示为 $D = [I_1, I_2, \dots, I_n] \in \mathbf{R}^{m \times n}$, $n \leq m$, 其中 $I_i \in \mathbf{R}^m$ 表示第 i 帧图像量化后的列向量,每帧图像有 m 个像素点。用于运动目标检测的鲁棒主成分分析(RPCA)方法认为观测矩阵 D 由 3 部分组成:低秩的背景矩阵 $B \in \mathbf{R}^{m \times n}$ 、前景目标构成的稀疏大噪声矩阵 $F \in \mathbf{R}^{m \times n}$ 和真实环境产生的高斯小噪声矩阵 $E \in \mathbf{R}^{m \times n}$ 。因此运动目标检测问题可以通过以下模型求解,即

$$\begin{aligned} \min_{B, F} \quad & \frac{1}{2} \|E\|_F^2 \\ \text{s. t.} \quad & D = B + F + E \\ & r(B) \leq K, \|F\|_0 \leq s \\ & E_{ij} \sim N(0, \sigma^2) \end{aligned} \quad (1)$$

式中 σ^2 表示高斯噪声的方差,取决于序列图像的质量。 $\|E\|_F = \sqrt{\sum_{ij} E_{ij}^2}$ 表示矩阵 E 的 Frobenius 范数。 $r(B)$ 表示矩阵 B 的秩,即非零奇异值的个数。 $\|F\|_0$ 表示矩阵 F 的 l_0 范数,即矩阵中非零元素的个数。 K 限制了背景矩阵的秩,即序列图像之间背景的相关性,该值取决于场景变化程度,比如光照等。 s 约束前景目标矩阵的稀疏程度,即前景目标像素点的个数,该值取决于运动目标的面积。

由于矩阵的秩函数和 l_0 范数的非凸性,目标式(1)是一个 NP 难问题,为了对式(1)求解,使用

Lagrange 乘子法将式(1)转变为无约束优化问题,并将矩阵的秩函数松弛为核范数, l_0 范数松弛为矩阵的 l_1 范数,即得到鲁棒主成分分析(RPCA)的一般形式

$$\min_{B, F} \frac{1}{2} \|D - B - F\|_F^2 + \alpha \|B\|_* + \beta \|F\|_1 \quad (2)$$

式中, $\|B\|_* = \sum_i \sigma_i$ 表示矩阵 B 的核范数, σ_i 表示矩阵 B 的第 i 个奇异值。 $\|F\|_1 = \sum_{ij} |F_{ij}|$ 表示矩阵 F 的 l_1 范数。 α, β 为平衡参数。

但 RPCA 方法并没有考虑运动目标检测中前景目标具有的先验信息。Zhou 等人注意到前景目标像素点具有集群的特点,即运动目标是区域连续的特点。因此, Zhou 等人在 RPCA 方法的基础上加入该先验知识,提出了 DECOLOR 方法。

DECOLOR 方法使用马尔可夫随机场^[17]来建模前景目标位置矩阵 S , $S \in \mathbf{R}^{m \times n}$, 位置矩阵 S 的元素定义为

$$S_{ij} = \begin{cases} 0 & D_{ij} \text{ 是背景像素点} \\ 1 & D_{ij} \text{ 是前景像素点} \end{cases} \quad (3)$$

使用图 $G = \langle v, \rho \rangle$ 刻画像素点之间关系,其中 v 是所有像素点的集合, ρ 是任意两个相邻像素点间的关系。由尹辛模型得到前景目标 S 的能量函数,即

$$\sum_{ij \in v} u_{ij}(S_{ij}) + \sum_{(ij, kl)} \lambda_{ij, kl} |S_{ij} - S_{kl}| \quad (4)$$

式中 $u_{ij}(S_{ij})$ 用来约束 S_{ij} 属于前景目标的可能性,定义

$$u_{ij}(x) = \begin{cases} 0 & x = 0 \\ \lambda_{ij} & x = 1 \end{cases} \quad (5)$$

式中 λ_{ij} 是惩罚因子。 $\lambda_{ij, kl}$ 是对相邻像素点 $\langle ij, kl \rangle$ 像素变化的惩罚,用以约束目标的区域连续性。而该算法对运动目标的位置没有任何先验约束,所以可以令 $\lambda_{ij}, \lambda_{ij, kl}$ 均为常量 λ 。

因此 DECOLOR 方法要解决的目标优化问题可以写成

$$\begin{aligned} \min_{B, S} \quad & \frac{1}{2} \sum_{ij: S_{ij}=0} (D_{ij} - B_{ij})^2 + \\ & \beta \sum_{ij} S_{ij} + \lambda \sum_{(ij, kl) \in e} |S_{ij} - S_{kl}| \\ \text{s. t.} \quad & r(B) \leq K \end{aligned} \quad (6)$$

式中 β, λ 为平衡参数,分别约束了图像局部的平滑

性和运动目标的区域连续性。

对于低秩约束,DECOLOR方法和RPCA方法一样将秩函数松弛为矩阵的核范数,将式(6)写成矩阵形式得到DECOLOR的最终求解模型,即

$$\min_{\mathbf{B}, \mathbf{S}} \frac{1}{2} \|P_S(\mathbf{D} - \mathbf{B})\|_F^2 + \alpha \|\mathbf{B}\|_* + \beta \|\mathbf{S}\|_1 + \lambda \|\mathbf{A}v(\mathbf{S})\|_1 \quad (7)$$

式中 \mathbf{A} 是像素点之间约束矩阵, $v(\mathbf{S})$ 表示将位置矩阵 \mathbf{S} 矢量化。 $P_S(\cdot)$ 为投影到矩阵 \mathbf{S} 上的投影算子,即

$$P_S(\mathbf{X}) = \begin{cases} X_{ij} & S_{ij} = 0 \\ 0 & S_{ij} = 1 \end{cases} \quad (8)$$

式中 $\mathbf{X} \in \mathbf{R}^{m \times n}$ 。

但理论表明核范数并不是秩函数的最优替代函数,只有当满足一定的非相关假设时,使用核函数代替秩函数才能够得到近似的最优低秩解^[18],而这些非相关假设在实际应用中很难满足,因此希望使用条件更加宽松却更加紧致的非凸加权核范数替代秩函数。

1.2 非凸加权核范数

对于向量的 l_0 范数的松弛函数,Candes等人^[19]指出使用向量的非凸惩罚函数能比传统 l_1 范数松弛取得更精确的结果。非凸惩罚函数 $f(t)$ 定义为

$$f(t) = \eta(|t|) \quad (9)$$

式中 $\eta(t)$ 是实数域 \mathbf{R} 上的单调不减非凸函数,且 $\eta(0) = 0$ 。

图1给出了1维情况下非凸惩罚函数与 l_1 范数对 l_0 范数逼近程度的对比。其中 $f_1(t)$ 表示向量的 l_1 范数, $f_0(t)$ 表示向量的 l_0 范数, $f(t)$ 表示非凸惩罚函数。由图1直观看到 $f(t)$ 比 $f_1(t)$ 更精确地逼近 $f_0(t)$ 函数。即非凸惩罚函数能够比 l_1 范数更好地描述稀疏特征。

而矩阵的秩相当于矩阵奇异值向量的 l_0 范数,矩阵的核范数相当于矩阵奇异值向量的 l_1 范数,因此很自然的将向量的非凸惩罚函数扩展到矩阵的奇异值向量上,得到矩阵的非凸加权核范数^[20]。

矩阵的非凸加权核范数定义为矩阵奇异值经过非凸函数映射后的和,即矩阵 $\mathbf{X} \in \mathbf{R}^{m \times n}$ 的非凸加权核范数定义为

$$\|\mathbf{X}\|_{nw} = \sum_{i=1}^n g(\sigma_i^X) \quad (10)$$

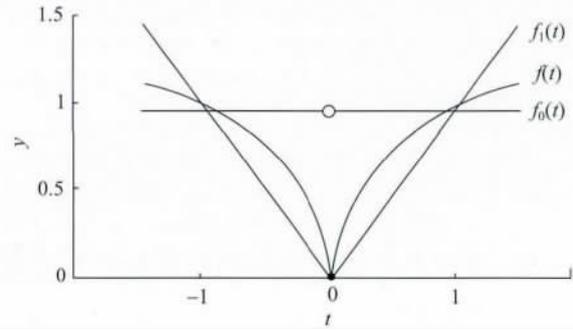


图1 1维情形下 l_1 范数和非凸惩罚函数对 l_0 范数逼近程度的比较

Fig. 1 The comparison of l_0 -norm and nonconvex penalty function as the approximation of l_0 -norm under one-dimensional situation

式中 $n \leq m$, σ_i^X 表示矩阵 \mathbf{X} 的第 i 个奇异值, $\|\mathbf{X}\|_{nw}$ 表示矩阵 \mathbf{X} 的非凸加权核范数,与式(9)中非凸惩罚函数定义类似,函数 $g(t)$ 是作用在矩阵奇异值上的非凸惩罚函数,而矩阵的奇异值均是非负值,因此实函数 $g(t)$ 只需满足两个条件: 1) $g(0) = 0$; 2) 在 $[0, \infty)$ 区间是单调不减非凸函数。

在向量稀疏约束中,非凸惩罚函数能够比 l_1 范数取得更好的稀疏效果。与之类似,在矩阵的低秩约束中,使用矩阵的非凸加权核范数逼近秩函数能够比核范数取得更好的效果。但是矩阵的非凸加权核范数优化问题比向量的非凸稀疏问题更具有挑战性。

2 基于非凸核范数的运动目标检测

DECOLOR方法通过在RPCA基础上加入前景目标的先验知识,大大提升了前景目标检测的正确率,但是该方法运行时间较长不适用于运动目标实时性检测,可以首先离线使用该方法作为背景建模,然后再使用其他实时性算法进行运动目标的实时性检测。

和RPCA方法一样,DECOLOR方法仍是通过矩阵的核范数逼近矩阵的秩函数计算背景矩阵,通过第1节知道,使用加权非凸核范数能够比核范数更好地逼近矩阵的秩函数,因此使用加权非凸核范数取代矩阵的核范数去逼近优化矩阵的秩函数,并且类似于DECOLOR方法在优化目标函数中加入运动目标的区域连续性,从而在保持前景目标检测性

能前提下,提升了背景恢复的准确度,对后续的实时性目标检测具有重要意义。

想解决的目标式为

$$\min_{B, S_{ij} \in \{0,1\}} \frac{1}{2} \|P_S(\mathbf{D} - \mathbf{B})\|_F^2 + \alpha \sum_i g(\sigma_i^B) + \beta \|\mathbf{S}\|_1 + \lambda \|\mathbf{A}v(\mathbf{S})\|_1 \quad (11)$$

式中, $\mathbf{S} \in \mathbf{R}^{m \times n}$ 是希望得到的前景目标的位置矩阵。

目标式(11)中既含有连续变量又含有离散变量,且包含非凸函数,而且基于非凸加权核范数的低秩优化问题本身就很难求解,因此无法直接求解式(11)。需使用交替方向法求解式(11)。

2.1 低秩背景矩阵估计

当 \mathbf{S} 给定时,式(11)中 \mathbf{B} 的求解可以看做是矩阵完备化问题,即

$$\min_B \frac{1}{2} \|P_S(\mathbf{D} - \mathbf{B})\|_F^2 + \alpha \sum_i g(\sigma_i^B) \quad (12)$$

因为函数 $g(t)$ 在 $[0, \infty)$ 区间是非凸函数 $g(t)$ 满足

$$g(\sigma_j^B) + \langle w_i, \sigma_j^B - \sigma_i^B \rangle \geq g(\sigma_i^B) \quad (13)$$

式中 w_i 表示函数 $g(t)$ 在 $t = \sigma_i^B$ 处的次梯度,即 $w_i \in \partial g(\sigma_i^B)$, $\partial g(\sigma_i^B)$ 是函数 $g(t)$ 在 $t = \sigma_i^B$ 的次梯度集。

由于 $g(t)$ 在 $[0, \infty)$ 上的非凸性,问题(12)难以直接求解,但式(13)给出了非凸函数的线性逼近方法。因此问题式(12)可以通过迭代式(14)求解。

$$\begin{aligned} \mathbf{B}^{k+1} &= \arg \min_B \frac{1}{2} \|P_S(\mathbf{D} - \mathbf{B})\|_F^2 + \\ &\alpha \sum_i (g(\sigma_i^{B^k}) + w_i^k (\sigma_i^B - \sigma_i^{B^k})) = \\ &\arg \min_B \frac{1}{2} \|P_S(\mathbf{D} - \mathbf{B})\|_F^2 + \alpha \sum_i w_i^k \sigma_i^B \quad (14) \end{aligned}$$

式中 k 是当前迭代次数, $\sigma_i^{B^k}$ 是第 k 迭代得到的低秩背景矩阵 \mathbf{B}^k 的第 i 个奇异值, w_i^k 是 $[0, \infty)$ 上的非凸单调不减函数 $g(t)$ 在 $t = \sigma_i^{B^k}$ 处的一个次梯度值。

但是在式(14)中,矩阵投影算子 $P_S(\cdot)$ 依然是非凸函数,因此为了便于求解,采用一阶 Taylor 展开式线性化式(14)中第1项对误差的约束。函数 $h(\mathbf{B}) = \|P_S(\mathbf{D} - \mathbf{B})\|_F^2$ 在矩阵 $\mathbf{Y} \in \mathbf{R}^{m \times n}$ 处的一阶 Taylor 展开为

$$\begin{aligned} h(\mathbf{B}) &\approx h(\mathbf{Y}) + \nabla h(\mathbf{Y}) \langle \mathbf{B} - \mathbf{Y} \rangle + \\ &\frac{\mu}{2} \|\mathbf{B} - \mathbf{Y}\|_F^2 = \end{aligned}$$

$$\langle P_S(\mathbf{D} - \mathbf{Y}) \mathbf{B} - \mathbf{Y} \rangle + \frac{\mu}{2} \|\mathbf{B} - \mathbf{Y}\|_F^2 \quad (15)$$

式中 $\mu > L(h)$, $L(h)$ 表示函数 $h(\mathbf{B})$ 的 Lipschitz 常数。

因此迭代式(14)的最终迭代式为

$$\begin{aligned} \mathbf{B}^{k+1} &= \arg \min_B \alpha \sum_i w_i^k \sigma_i^B + \\ &\frac{\mu}{2} \left\| \mathbf{B} - \left(\mathbf{B}^k + \frac{1}{\mu} (P_S(\mathbf{D} - \mathbf{B}^k)) \right) \right\|_F^2 \quad (16) \end{aligned}$$

问题式(16)依然是一个非凸优化问题,但根据下述定理,该问题有解析解。

定理^[21]对任意的 $\lambda > 0$, $\mathbf{Y} \in \mathbf{R}^{m \times n}$, 当 $0 \leq w_1 \leq w_2 \leq \dots \leq w_n$ 时 ($n \leq m$), 最小化问题

$$\min_X \lambda \sum_{i=1}^n w_i \sigma_i^Y + \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2$$

可以通过加权奇异收缩算子得到全局最优解

$$\mathbf{X}^* = \mathbf{U} \mathbf{S}_{\lambda w}(\boldsymbol{\Sigma}) \mathbf{V}^T$$

式中 $\mathbf{Y} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T$ 是 \mathbf{Y} 的奇异值分解过程,且 $\sigma_1^Y \geq \sigma_2^Y \geq \dots \geq \sigma_n^Y \geq 0$, $\mathbf{S}_{\lambda w}(\boldsymbol{\Sigma})$ 是以 $(\sigma_{ii} - \lambda w_i)_+$ 为元素的对角矩阵 $i \in \{1, 2, \dots, n\}$ $(t - a)_+ = \begin{cases} t - a & t > a \\ 0 & t \leq a \end{cases}$ 。

因为 $g(t)$ 在 $[0, \infty)$ 上是单调不减函数,所以 σ_i^Y 越大时 w_i 越小,即式(16)满足定理条件,可以得到解析解。其迭代过程在算法1中给出。

算法1 低秩背景矩阵估计

输入: $\mu > L(f)$, $\mathbf{D} \in \mathbf{R}^{m \times n}$, $\mathbf{S} \in \mathbf{R}^{m \times n}$, $n \leq m$ 。

初始化: $k = 0$, $\mathbf{B}_k = \mathbf{0}$, $w_i^k = 0$, $i = 1, \dots, n$ 。

迭代:

1) 根据定理,计算式(16)的解析解,得到 \mathbf{B}^{k+1} 。

2) $w_i^{k+1} = \partial g(\sigma_i^{B^{k+1}})$, 即计算函数 $g(t)$ 在 $t = \sigma_i^{B^{k+1}}$ 处的次梯度 $\sigma_1^{B^{k+1}} \geq \dots \geq \sigma_n^{B^{k+1}}$ 。

直到收敛。

输出: $\mathbf{B}^* = \mathbf{B}^{k+1}$ 。

2.2 稀疏前景位置矩阵估计

在背景低秩矩阵 \mathbf{B} 已经估计得到的前提下,前景位置矩阵的求解即最优化能量函数

$$\begin{aligned} F(\mathbf{S}) &= \frac{1}{2} \|P_S(\mathbf{D} - \mathbf{B})\|_F^2 + \beta \|\mathbf{S}\|_1 + \\ &\lambda \|\mathbf{A}v(\mathbf{S})\|_1 = \\ &\frac{1}{2} \sum_{ij} (D_{ij} - B_{ij})^2 (1 - S_{ij}) + \\ &\beta \sum_{ij} S_{ij} + \lambda \|\mathbf{A}v(\mathbf{S})\|_1 \quad (17) \end{aligned}$$

该类目标函数可以通过图割法^[22-24]获得精确最

优解。

在图的构造过程中,应该同时考虑到空间和时间上相邻像素点的关系,但是这样会导致图的结构非常庞大而且复杂,难以计算。因此实验只考虑空间上,即每一帧图像上相邻像素点之间的关系。实验结果表明,这种简化依然能够取得理想的效果。

2.3 基于非凸加权核范数的运动目标检测算法

本文提出的基于非凸加权核范数的运动目标检测问题可以建模为式(11)的优化问题,而该问题可以通过低秩背景矩阵的恢复和前景位置矩阵的估测两个过程交迭得到。本文提出的模型解决方法在算法2中给出。

算法2 基于非凸加权核范数的运动目标检测

输入: 观测矩阵 $D \in \mathbf{R}^{m \times n}$ 。

初始化: $\beta, \alpha, \lambda, k=0, B^k=0, S^k=D$

迭代

1) 使用算法1解决式(12)中的优化问题,得到低秩背景矩阵 B^{k+1} 。

2) 使用图割法求解式(17)模型,得到前景位置矩阵 S^{k+1} 。

直至算法收敛。

输出: $B^* = B^{k+1}, S^* = S^{k+1}$ 。

算法2中 α, β, λ 分别是连续帧序列的背景相关性、总体前景目标面积大小和单一前景目标区域连续性的权衡参数。其中 β 取决于前景目标的稀疏性,前景目标所占区域越大 β 取值越小,反之亦然。 λ 取决于每个目标区域面积的大小。当 λ 较大时,算法会把一些零散的稀疏大噪声连续区域裁定为前景目标区域,因此当待检测目标中最小目标的面积越大时 λ 应该越小。 α 取决于连续帧序列图像中场景的变化情况,比如光照变化、树影摇曳、水纹波动等。若场景变化缓慢,那么每帧图像场景的真实背景相关性较强 α 应该取较大的值,否则 α 取较小的值。在实际应用中,可以使 β 由一个较小的值逐渐增大 α 由一个较大的值逐渐减小,即在算法开始时被检测为前景区域的点很多,随着算法的执行,对这些点进行逐步筛选,最终得到较精确的结果。

3 实验结果与分析

设计实验比较鲁棒主成分分析方法(RPCA)、基于加权非凸核范数的主成分分析方法(WNPCA)、

DECOLOR方法和本文方法分别在仿真数据和真实数据上对于运动背景图像恢复和运动目标检测的性能。大量的研究表明,在向量稀疏优化问题中,常用的非凸惩罚函数,即 $f(t)$,取非凸指数函数^[16,19]时具有较好的效果,因此在实验中令矩阵奇异值的映射函数 $g(t)$ 取指数函数,并令该函数非凸单调递增。即

$$g(x) = x^p \quad (18)$$

式中 $0 \leq p \leq 1$ 保证了函数在 $[0, \infty)$ 区间内的非凸性。

3.1 评价准则

为了定量的比较4种低秩算法的在背景恢复和目标检测上的性能,使用均方根误差(RMSE)衡量背景恢复的效果,RMSE为

$$\text{RMSE}_s = \frac{\|B - B_0\|_F^2}{\|B_0\|_F^2} \quad (19)$$

式中 B 是算法恢复出来的背景图像,而 B_0 是真实的背景图像,下标 s 表示仿真数据背景恢复的评价准则。但对于真实数据中,难以获得准确的真实背景图像,因此通过评估运动场景中去除运动目标后的其他区域的恢复程度来近似作为背景的恢复评价指标。

$$\text{RMSE}_T = \frac{\|P_\Omega(B - D)\|_F^2}{\|P_\Omega(D)\|_F^2} \quad (20)$$

式中,下标 T 表示真实数据运动背景恢复的评价准则, Ω 是人工标定出的非目标区域。 $P_\Omega(\cdot)$ 是矩阵投影算子。

而为了衡量前景目标检测的正确率,把目标检测问题看作是前景像素和背景像素的分类问题,即将前景像素点作为正样本,而背景像素点作为负样本,并且由前景目标分类的准确率和召回率组合得到一个新的评价度量(Fm)^[25],即

$$Fm = \frac{2 \cdot \gamma_{\text{pre}} \cdot \gamma_{\text{rec}}}{\gamma_{\text{pre}} + \gamma_{\text{rec}}} \quad (21)$$

$$\gamma_{\text{pre}} = \frac{TP}{TP + FP}, \quad \gamma_{\text{rec}} = \frac{TP}{TP + FN}$$

式中 TP, FP, TN, FN 分别表示真正样本、误检正样本、真实负样本、误检负样本的个数。 $\gamma_{\text{pre}}, \gamma_{\text{rec}}$ 分别表示分类的准确率和召回率。

3.2 仿真数据

为了模拟真实数据,实验中仿真数据 $D \in \mathbf{R}^{m \times n}$ 也由3部分组成:背景低秩矩阵前景 $B_0 \in \mathbf{R}^{m \times n}$,稀

疏大噪声矩阵 $F \in \mathbf{R}^{m \times n}$ 和高斯小噪声矩阵 $E \in \mathbf{R}^{m \times n}$ 。其中低秩矩阵 B_0 的秩为 r , $B_0 = UV$, 其中 $U \in \mathbf{R}^{m \times r}$, $V \in \mathbf{R}^{r \times n}$ 是随机矩阵, m 为每帧图像中像素点的个数, n 为视频序列的图像帧数, $m > n$, 实验中取 $m = 200$, $n = 100$; 目标矩阵 F 由宽度为 W 的连续区域构造, 前景目标像素灰度值服从 $\mu(-c, \epsilon)$ 均匀分布, ϵ 为一个较大的常量; 高斯小噪声矩阵 E 由信噪比 SNR 约束, 即

$$SNR = \sqrt{\frac{\text{var}(B_0)}{\text{var}(E)}} \quad (22)$$

图 2 分别给出了仿真数据的观测矩阵 D 、背景矩阵 B_0 、目标矩阵 F 和小噪声矩阵 E 。图 3 给出了 $r = 10$, $SNR = 2$, $W = 60$ 条件下分别使用 RPCA、WMPCA、DECOLOR 和本文方法得到的背景图像和前景目标位置图像。发现在没有加入前景目标区域连续性条件时, 检测前景结果包含大量的小区域的噪声, 同时恢复出的背景图像上仍然能够看到前景目标的痕迹, 而加入前景目标区域连续性条件后, 前景目标区域更加准确和完整, 背景恢复也相对与未加该条件时更好。而比较图 3 能发现使用非凸加权核范数恢复出的背景矩阵比使用核范数恢复出的背

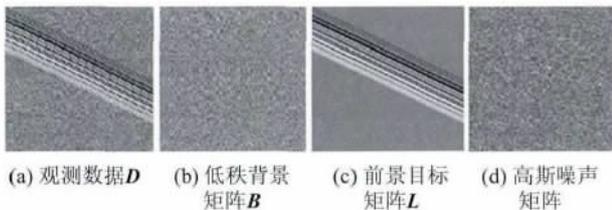


图 2 仿真数据
Fig. 2 Simulation Data

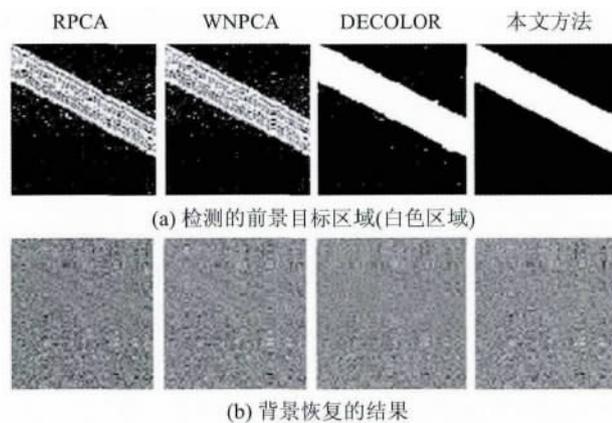
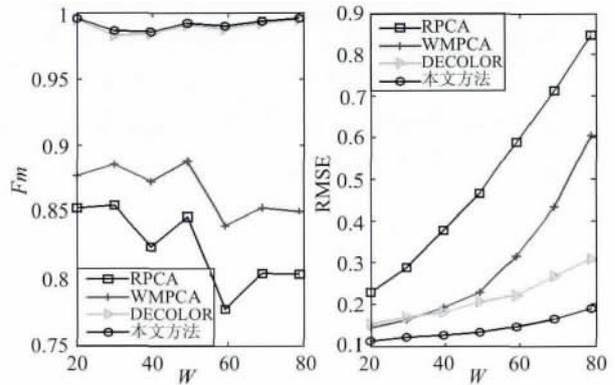


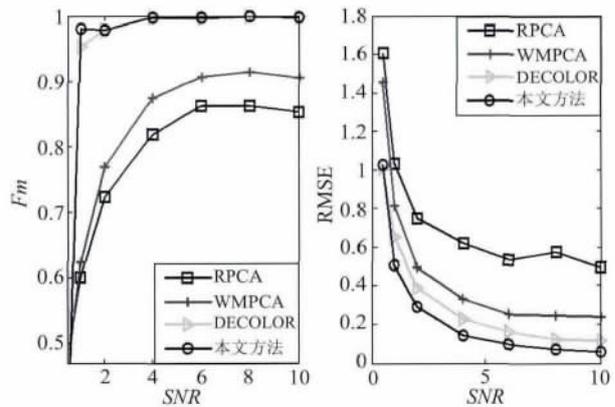
图 3 不同算法下目标检测和背景恢复的结果
Fig. 3 Result of object detection and background recovery under different methods

景矩阵更加精确, 进而前景目标的检测效果也得到提升。直观上可以感觉本文方法能够取得比其他方法更好的效果。

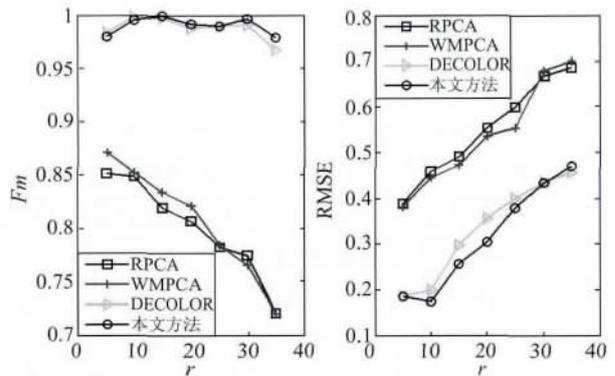
为了定量分析算法对于不同条件下的鲁棒性, 分别取不同值的 r , SNR , W 进行 500 次实验, 图 4 给出了不同条件下的 4 种算法平均性能比较。



(a) 前景区域大小对算法的影响



(b) 高斯噪声对算法的影响



(c) 背景环境变化复杂度对算法的影响

图 4 不同算法对不同参数的鲁棒性

Fig. 4 Robustness of different algorithms for different parameters

分析图 4 (a) 发现, W 的变化, 即前景目标的大小变化, 对本文方法和 DECOLOR 方法的前景检测准确率影响不大, 但对 RPCA 方法和 WMPCA 方法

的影响较大,而且 DECOLOR 方法和本文方法的效果要明显优于 RPCA 和 WMPKA 方法。另外虽然和 DECOLOR 相比,本文方法对于前景目标检测准确度提高的不大,但是对于背景恢复的准确性而言,本文方法要优于 DECOLOR 方法,随着前景目标面积的增大,这种趋势越明显。同样,分析图 4(b) 发现,对于不同程度的噪声环境,本文方法能够取得不低于 DECOLOR 方法的前景目标检测效果,取得优于 DECOLOR 方法的背景恢复效果。本文方法要显著优于 RPCA 和 WMPKA 方法。图 4(c) 验证了不同的背景复杂度下,本文方法相对于其他 3 种对比方法的优越性。

3.3 真实数据

在 5 种公共数据集上测试了 RPCA、WMPKA、DECOLOR 和本文算法的性能,表 1 给出了 5 种公共数据集的基本参数和来源。

表 1 实验使用图像序列数据集信息

Table 1 Information of the sequences used in experiments

数据集	大小 × 帧数	描述	来源文献
offices	[160, 120] × 48	拥挤场景	文献[26]
pedestrian	[238, 158] × 24	拥挤场景	文献[27]
hall	[160, 128] × 24	拥挤场景	文献[28]
waterSurface	[160, 128] × 48	动态背景	文献[28]
airport	[160, 128] × 48	动态背景	文献[28]

实验中,对于不同的算法,均调整参数使其取得最好的效果,其中 RPCA 和 DECOLOR 方法的参数取 Zhou 等人推荐的参数设置。表 2、表 3 分别给出了真实数据集在不同的基于低秩的运动目标检测方法下的前景检测和背景恢复的精度。图 5 给出了拥挤场景下数据集 hall 和动态背景下数据集 airport 最后一帧的图像数据和人工标定的数据,以及对应 RPCA、WMPKA、DECOLOR 和本文方法的检测结果。

由表 2 和表 3 可以看出,WMPKA 算法相对于 RPCA 算法能够取得更好的效果,究其原因在于前者使用非凸加权核范数,而后者使用传统的核范数,逼近矩阵的秩函数,能够更好地体现运动序列连续帧图像间背景的强相关性。DECOLOR 算法相对于 RPCA 算法无论对于前景目标检测还是背景图像恢复都有较大的改善,其根本原因在于 DECOLOR 算

表 2 不同算法下前景目标检测正确率 (F_m)

Table 2 The accuracy of foreground object detection under different algorithms (F_m)

数据集	RPCA	WMPKA	DECOLOR	本文方法
offices	0.598 8	0.612 9	0.879 3	0.905 0
pedestrian	0.605 5	0.613 4	0.773 7	0.796 9
hall	0.610 2	0.610 9	0.814 1	0.833 5
waterSurface	0.459 1	0.481 4	0.853 4	0.885 6
airport	0.489 9	0.505 8	0.791 2	0.796 7

表 3 不同算法下运动场景的背景恢复精度 (RMSE)

Table 3 The sports scene background recovery precision under different algorithms (RMSE)

数据集	RPCA	WMPKA	DECOLOR	本文方法
offices	0.308 9	0.283 8	0.174 1	0.154 8
pedestrian	0.203 5	0.196 0	0.178 3	0.165 5
hall	0.316 1	0.284 3	0.236 7	0.219 6
waterSurface	0.218 8	0.164 3	0.086 8	0.081 3
airport	0.167 2	0.179 8	0.138 6	0.136 4

法融合了前景目标的区域连续性信息,避免了一些零散的稀疏大噪声被误检为前景目标。而本文方法综合了 WMPKA 算法和 DECOLOR 算法的优势,实验结果也表明本文方法优于文中对比的其他 3 种低秩目标检测方法。从表 3 可以看出,对于前 3 个数据集,本文方法恢复出的背景精度要明显高于 DECOLOR 方法,而对于后两个数据集本文方法和 DECOLOR 方法背景恢复差异并不是很明显,分析发现其原因在于前 3 个数据集的背景是静态背景而后 2 个数据集的背景具有周期性的动态纹理,如波动的水纹和运行中的电梯。

4 结 论

当前运动目标检测方法非常多,而基于低秩的方法利用运动序列连续帧图像间的强相关性能够有效地解决其他方法所不能解决的许多问题,例如周期性动态纹理背景、光照变化和目标部分遮挡等。另外基于低秩的方法能够同时完成前景目标区域分割和背景图像的恢复。因此基于低秩的运动目标检测方法相对于其他方法具有明显优势,进而吸引了越来越多的关注。

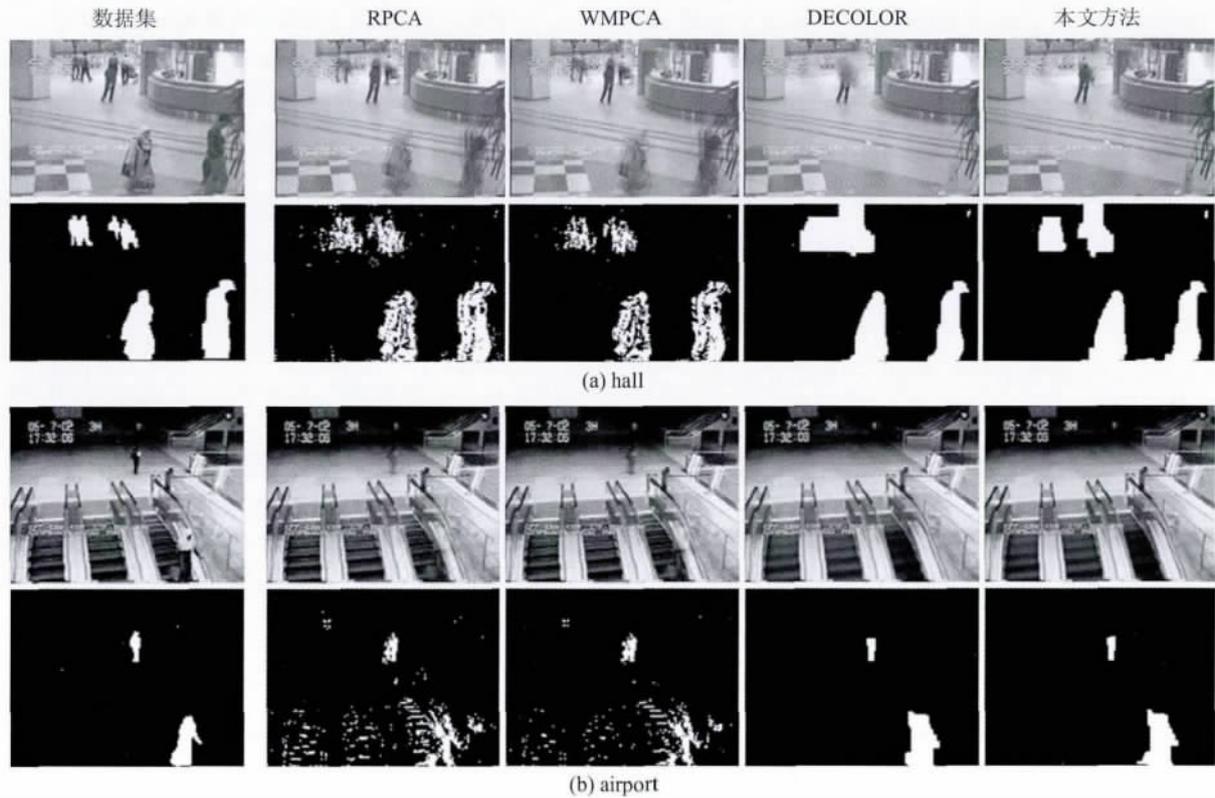


图5 不同算法下不同条件的数据集最后一帧图像对应的前景目标区域和背景恢复结果

Fig. 5 The background restored and foreground objects detected of the last frames of different datasets under different algorithms

本文提出的一种基于非凸加权核范数的运动目标检测方法也是一种低秩类运动目标检测方法。该方法一方面在传统的鲁棒主成分分析(RPCA)方法中,使用更加合理的非凸加权核范数取代核范数去逼近矩阵的秩函数,以期得到更精确的背景恢复图像;另一方面结合了运动过程中区域连续性的先验知识,理论上能够大大降低零散稀疏大噪声的干扰,从而提升前景目标检测的正确率。实验结果也表明,无论是仿真数据集还是真实数据集,本文方法相对于其他3种基于低秩的方法能够取得更好的效果。虽然本文方法在运动序列中前景目标检测和背景图像恢复方面的精度都高于其他方法,但是本文方法依然存在着一些问题。该方法对于一些较大的纹理稀少的目标区域会出现错误分割为前景目标的现象,这是由于较大的纹理稀少区域对于背景的低秩结构影响较大,为了解决该问题,可以在模型中加入运动目标的外观或者形状等先验知识。另外,该方法需要较大的计算量,尤其是当视频帧数不断增加时,计算量急剧增加,因此,该算法难以用于实时系统。为了解决实时性问题,后续工作可以从两方面入手,一方面可以通过增量奇异值分解^[29]方法解

决视频帧数增加导致的计算时间急剧增加问题;另一方面考虑到该方法恢复出的背景图像精度较高,可以使用该方法作为背景建模的一种方法,然后使用恢复出的背景图像和其他的,如背景减除法,较简单的运动目标检测方法实现运动目标的实时性检测。

参考文献(References)

- [1] Yilmaz A, Javed O, Shah M. Object tracking: a survey [J]. *ACM Computing Surveys*, 2006, 38(2): 1-45. [DOI: 10.1145/1177352.1177355]
- [2] Cremers D, Soatto S. Motion competition: a variational approach to piecewise parametric motion segmentation [J]. *International Journal of Computer Vision*, 2005, 62(3): 249-265. [DOI: 10.1007/s11263-005-4882-4]
- [3] Amiaz T, Kirvati N. Piecewise-smooth dense optical flow via level sets [J]. *International Journal of Computer Vision*, 2006, 68(2): 111-124. [DOI: 10.1007/s11263-005-6206-0]
- [4] Beauchemin S S, Barron J L. The computation of optical flow [J]. *ACM Computing Surveys (CSUR)*, 1995, 27(3): 433-466. [DOI: 10.1145/212094.212141]
- [5] Nagel H H, Enkelmann W. An investigation of smoothness con-

- straints for the estimation of displacement vector fields from image sequences[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1986, 8(5): 565-593. [DOI: 10.1109/TPAMI.1986.4767833]
- [6] Meoslund T B, Hilton A, Kruger V. A survey of advances in vision-based human motion capture and analysis[J]. *Computer Vision and Image Understanding*, 2006, 104(2): 90-126. [DOI: 10.1016/J.CVIU.2006.08.002]
- [7] Wren C R, Azarbayejani A, Darrell T, et al. Pfister: real-time tracking of the human body[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997, 19(7): 780-785. [DOI: 10.1109/34.598236]
- [8] Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking[C]//*Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1999*. Fort Collins, CO: IEEE, 1999, 2: 246-252. [DOI: 10.1109/CVPR.1999.784637]
- [9] Mittal A, Paragios N. Motion-based background subtraction using adaptive kernel density estimation[C]//*Proceedings of Computer Society Conference on Computer Vision and Pattern Recognition*. Washington DC: IEEE, 2004, 2(2): II-302-II-309. [DOI: 10.1109/CVPR.2004.1315179]
- [10] Rittscher J, Kato J, Joga S, et al. A Probabilistic Background Model For Tracking[M]//*European Conference on Computer Vision*. Springer Berlin Heidelberg: ECCV, 2000(1843): 336-350. [DOI: 10.1007/3-540-450530X_32]
- [11] Zhong J, Sclaroff S. Segmenting foreground objects from a dynamic textured background via a robust kalman filter[C]//*Proceedings of Ninth IEEE International Conference on Computer Vision*. Nice, France: IEEE, 2003: 44-50. [DOI: 10.1109/ICCV.2003.1238312]
- [12] Kim K, Chalidabhongse T H, Harwood D, et al. Real-time foreground-background segmentation using codebook model[J]. *Real-Time Imaging*, 2005, 11(3): 172-85. [DOI: 10.1016/J.RTI.2004.12.004]
- [13] Wang L, Cheng H, Liu C, et al. A robust elastic net approach for feature learning[J]. *Journal of Visual Communication and Image Representation*, 2014, 25(2): 313-321.
- [14] Candes E, Li X, Ma Y, et al. Robust principal component analysis? [J]. *Journal of the ACM*, 2011, 58(3): 11. [DOI: 10.1145/1970392.1970395]
- [15] Zhou X, Yang C. Moving object detection by detecting contiguous in the low-rank representation[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, 35(3): 597-610. [DOI: 10.1109/TPAMI.2012.132]
- [16] Chen K, Dong H, Chan K. Reduced rank regression via adaptive nuclear norm penalization[J]. *Biometrika*, 2015, 102(2): 439-456. [DOI: 10.1093/BIOMET/AST036]
- [17] Li S Z. Markov Random Field Modeling In Image Analysis[M]. London: Springer, 2009: 1-21.
- [18] Candes E, Tao T. The power of convex relaxation: near-optimal matrix completion[J]. *IEEE Transactions on Information Theory*, 2010, 56(5): 2053-2080. [DOI: 10.1109/TIT.1020.2044061]
- [19] Candes E, Wakin M, Boyd S. Enhancing sparsity by reweighted l_1 minimization[J]. *Journal of Fourier Analysis and Applications* 2008, 14(5-6): 877-905. [DOI: 10.1007/S00041-008-9045-X]
- [20] Lu C Y, Tang J H, Yan S C, et al. Generalized nonconvex nonsmooth low-rank minimization[C]//*Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 2014*. Columbus, Ohio: IEEE, 2014: 4130-4137. [DOI: 10.1109/CVPR.2014.526]
- [21] Lu C, Zhu C, Xu C, et al. Generalized singular value thresholding[C]//*Proceedings of Twenty-Ninth AAAI Conference on Artificial Intelligence*. Austin, Texas: AAAI, 2015: 1805-1811.
- [22] Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2004, 26(9): 1124-1137. [DOI: 10.1109/TPAMI.2004.60]
- [23] Boykov Y, Veksler O, Zabih R. Fast approximate energy minimization via graph cuts[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2001, 23(11): 1222-1239. [DOI: 10.1109/34.969114]
- [24] Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? [J]. *IEEE Transactions Pattern Analysis and Machine Intelligence*, 2004, 26(2): 147-159. [DOI: 10.1109/TPAMI.2004.1262177]
- [25] Davis J, Goadrich M. The relationship between Precision-Recall and ROC curves[C]//*Proceedings of the 23rd International Conference on Machine Learning*. Corvallis: ACM, 2006: 233-240. [DOI: 10.1145/1143844.1143874]
- [26] Wang H, Suter D. A Novel Robust Statistical Method For Background Initialization And Visual Surveillance[M]//*Computer Vision-ACCV 2006*. Berlin Heidelberg: Springer, 2006: 328-337. [DOI: 10.1007/11612032_34]
- [27] Chan A, Vasconcelos N. Layered dynamic textures[J]. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 2009, 31(10): 1862-1879. [DOI: 10.1109/TPAMI.2009.110]
- [28] Li L, Huang W, Gu I, et al. Statistical modeling of complex backgrounds for foreground object detection[J]. *IEEE Transactions on Image Processing*, 2004, 13(11): 1459-1472. [DOI: 10.1109/TIP.2004.836169]
- [29] Ross D A, Lim J, Lin R S, et al. Incremental learning for robust visual tracking[J]. *International Journal of Computer Vision*, 2008, 77(1-3): 125-141. [DOI: 10.1007/S11263-007-0075-7]