

# 基于连续性约束背景模型减除的运动目标检测

祝 轩 王 磊 张 超 梅东锋 薛珈萍 曹晴雯

(西北大学信息科学与技术学院 西安 710127)

**摘 要** 运动目标检测是机器视觉领域中的关键技术之一,其在视频运动目标检测、遥感信息处理和军事侦察等领域有广泛的应用。考虑到视频中相邻视频帧背景相似性高且时间连续性长,而阴影和噪声具有非连续性的特征,文中提出一种时间连续性约束的低秩分解背景更新模型,并将其应用于背景模型减除的视频运动目标检测。该方法首先对视频进行低秩分解,获得低秩分量和稀疏分量;然后基于连续性约束背景更新模型更新低秩分量,构建背景;最后通过背景减除及自适应阈值分割获得运动目标。实验结果表明,无论是 FM 指标还是 ROC 曲线都反映出所提方法相比目前较好的背景减除方法能够有效克服阴影和噪声的影响,避免“空洞”,更准确地提取运动目标,且鲁棒性好。

**关键词** 运动目标检测,连续性约束,背景减除,低秩分解

中图分类号 TN911.73 文献标识码 A DOI 10.11896/j.issn.1002-137X.2019.06.047

## Moving Object Detection Based on Continuous Constraint Background Model Deduction

ZHU Xuan WANG Lei ZHANG Chao MEI Dong-feng XUE Jia-ping CAO Qing-wen

(School of Information Science and Technology, Northwest University, Xi'an 710127, China)

**Abstract** Moving target detection is one of the key technologies in the field of machine vision. Moving object detection is widely used in video moving object detection, remote sensing information processing and military reconnaissance, etc. Considering that the background similarity of adjacent video frames is high, and the shadow and noise are discontinuous, this paper proposed a low-rank decomposition background updating model with time continuity constraint, and applied it to the moving object detection of background subtraction. Firstly, low-rank components and sparse components are obtained by using low-rank decomposition. Then the background is constructed by updating the low-rank components based on time continuity constrained. Finally, moving object is obtained by background subtraction and adaptive threshold segmentation. Experimental results show that both the FM index and the ROC curve reflect that compared with the state-of-the-art background subtraction methods, this method can effectively overcome the influence of shadow and noise, reduce holes, extract moving objects more accurately, and has good robustness.

**Keywords** Moving object detection, Background subtraction, Low rank decomposition, Continuity constraint

## 1 引言

通常视频内容由前景(运动目标)和背景组成,运动目标相比于静止的背景更能表征视频的内容,其是视频最重要的信息<sup>[1]</sup>。运动目标检测技术就是将视频中运动物体所在区域从背景中提取出来的技术,如提取道路监控视频中运动的车辆和行人等。

目前,常用的运动目标检测方法主要分为三大类:帧间差分法<sup>[2]</sup>、光流法<sup>[3]</sup>和背景减除法<sup>[4]</sup>。帧间差分法是利用运动目标在各视频帧中的不同位置关系进行目标检测。该方法的优势在于实时性好<sup>[5]</sup>,但当目标运动速度较快、场景变化较大以及噪声较强时,检测结果往往会出现“空洞”和“阴影”<sup>[6]</sup>。

帧间差分法主要包括三帧间差分法<sup>[7]</sup>和累计帧间差分法<sup>[8]</sup>等。光流法是利用目标运动的光流矢量具有连续性的特点来检测视频中的运动目标<sup>[9]</sup>。光流法的优势在于不易受背景运动影响,检测精度较高,但算法复杂度高、实时性差<sup>[10]</sup>。背景减除法首先建立视频的背景模型,然后计算背景模型与视频的差分以获取运动目标所在的区域。背景减除法的优势在于适合复杂背景下的运动目标检测,且检测精度较高<sup>[11]</sup>。目前,背景减除法主要包括低秩分解<sup>[12]</sup>、IALM<sup>[13]</sup>、PSPG<sup>[14]</sup>、RPCA<sup>[15]</sup>、GoDec<sup>[16]</sup>和 MAMR<sup>[17]</sup>等。背景减除法因良好的检测能力而受到相关学者的广泛关注,成为运动目标检测的热点研究方法。

考虑到视频中相邻视频帧的背景相似性高且时间连续性

到稿日期:2018-05-29 返修日期:2018-10-16 本文受陕西省自然科学基金基础研究计划重点项目(2018JZ6007)资助。

祝 轩(1968—),女,博士,教授,主要研究方向为计算机视觉、稀疏表示、模式识别,E-mail:xuan\_zhu@126.com(通信作者);王 磊(1994—),女,硕士生,主要研究方向为图像/图形处理;张 超(1995—),女,硕士生,主要研究方向为计算机视觉、稀疏表示;梅东锋(1993—),男,硕士生,主要研究方向为机器学习;薛珈萍(1995—),女,硕士生,主要研究方向为图像/图形处理;曹晴雯(1994—),女,硕士生,主要研究方向为图像/图形处理。

长,而阴影和噪声具有非连续性的特征<sup>[22]</sup>,本文提出了基于连续性约束的背景更新模型,并将其应用于背景减除法,以实现运动目标检测。该方法首先对视频进行低秩分解获得低秩分量(背景模型)和稀疏分量;然后用连续性约束更新模型更新低秩分量,构建背景;最后进行背景减除,利用自适应阈值分割出运动目标。实验结果表明,与目前比较先进的背景减除法相比,本文方法能够有效克服阴影和噪声的影响,更准确地分割出运动目标,并且对于大运动目标提取,其能够避免“空洞”。

## 2 连续性约束低秩分解背景更新模型的构建

### 2.1 低秩分解模型

低秩分解模型<sup>[18]</sup>也称为低秩矩阵分解模型:

$$\min_{L,S} \text{rank}(L) + \|S\|_0, \text{ s. t. } M=L+S \quad (1)$$

其中, $M$ 是输入矩阵, $L$ 是低秩矩阵, $S$ 是稀疏矩阵, $\text{rank}(\cdot)$ 表示矩阵的秩, $\|\cdot\|_0$ 表示矩阵的 $\ell_0$ 范数。式(1)是一个高

度非凸的优化求解问题,可以转化为凸优化求解问题<sup>[19]</sup>:

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1, \text{ s. t. } M=L+S \quad (2)$$

其中, $\|\cdot\|_*$ 表示矩阵的核范数; $\|\cdot\|_1$ 表示矩阵的 $\ell_1$ 范数; $\lambda$ 是正则化参数,用于调节稀疏矩阵所占的权重,通常取:

$$\lambda = \frac{1}{\sqrt{\max(p \times q)}} \quad (3)$$

其中, $p$ 和 $q$ 为输入矩阵 $M$ 的行和列的大小。

对于一段 $n$ 帧视频,若每一帧均用列向量表示, $n$ 个列向量组成矩阵 $M$ ,则 $M$ 可以看成式(1)中的输入矩阵,其具有低秩特性<sup>[20]</sup>;视频背景可以通过式(1)中的低秩矩阵 $L$ 构建;式(1)中的稀疏矩阵 $S$ 反映的是视频中的运动目标、阴影和噪声等。如果低秩矩阵 $L$ 能够准确地表示视频背景,那么采用背景减除法就能准确提取视频中的运动目标<sup>[23]</sup>。

### 2.2 连续性约束的低秩分解背景更新模型

连续性约束背景模型减除的运动目标检测如图1所示。

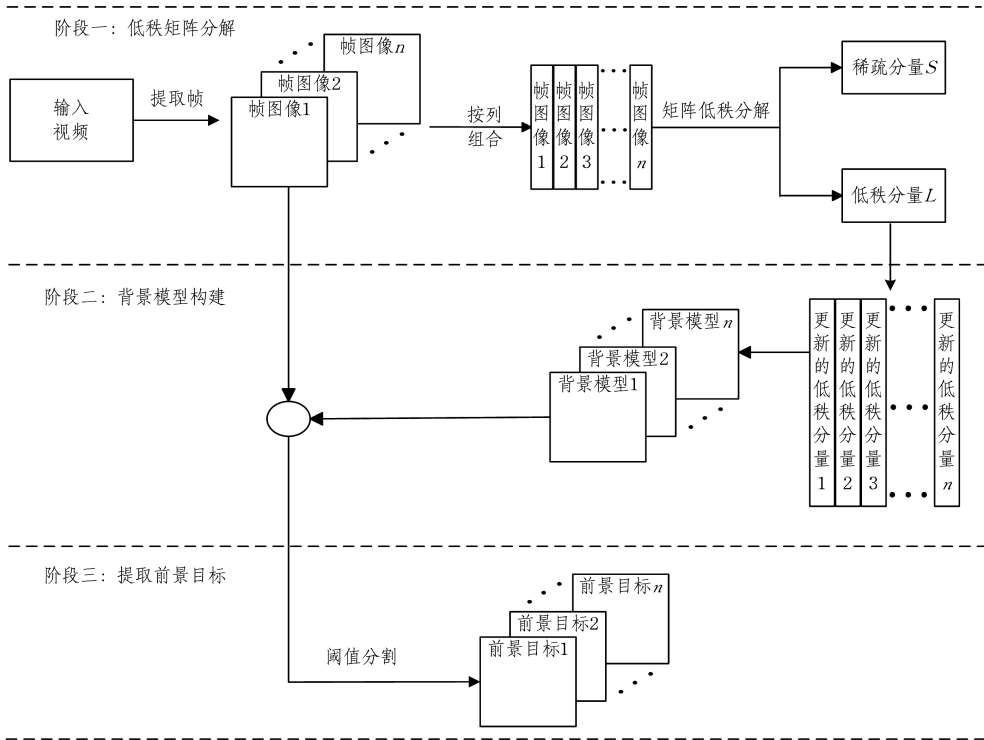


图1 连续性约束背景模型减除的运动目标检测

Fig. 1 Motion object detection by deducing continuous constraint background

根据式(1)中的低秩分量 $L$ 构建背景模型并提取运动目标存在两方面的不足:1)低秩分量 $L$ 构建的背景模型中包含除了运动目标和运动目标阴影之外的背景信息,背景减除的结果是运动目标及其阴影都被提取出来;2)噪声被划分到稀疏分量 $S$ ,背景减除后噪声也被当作运动目标提取出来。也就是说,基于低秩分解模型的背景减除法检测到的运动目标包括运动目标、阴影和噪声。考虑到阴影和噪声在邻近的几帧视频中连续出现,但不会持续较长时间,以及阴影受一个光源的影响,其形状在不同的视频帧中往往各不相同,而运动目标通常具有时间上的持续性,本文基于阴影和噪声的非连续性特点,提出引入时间连续性约束的低秩矩阵更新模型式(4),利用更新后的低秩矩阵 $L'$ 构建准确的背景模型,从而

有效克服阴影和噪声的影响,提高运动目标检测的准确性。

设有 $n$ 帧视频 $M$ ,若记 $M=[M^1, M^2, \dots, M^n]$ ,其中 $M^k=M(:,k)$ ( $k=1,2,\dots,n$ ),则 $M$ 满足低秩分解模型式(1);若记低秩矩阵 $L=[L^1, L^2, \dots, L^n]$ ,稀疏矩阵 $S=[S^1, S^2, \dots, S^n]$ ,则更新后的背景模型 $L'=[L'^1, L'^2, \dots, L'^n]$ ,其中 $L^k=L(:,k)$ ( $k=1,2,\dots,n$ ), $S^k=S(:,k)$ ( $k=1,2,\dots,n$ ), $L'^k=L'(:,k)$ ( $k=1,2,\dots,n$ )。我们定义连续性约束低秩矩阵更新模型式(4),并用其更新背景模型 $L'$ 。

$$F(m_{i,j}, l_{i,j}) = \begin{cases} l'_{i,j} = m_{i,j}, & l_{i,j} = m_{i,j} \\ l'_{i,j} = m_{i,j}, & G(l_{i,j-r}, \dots, l_{i,j}, \dots, l_{i,j+r}) < v \\ l'_{i,j} = m_{i,j}, & G(s_{i,j-r}, \dots, s_{i,j}, \dots, s_{i,j+r}) < u \end{cases} \quad (4)$$

其中,  $F(\cdot)$  是更新函数,  $m_{i,j}, l_{i,j}, s_{i,j}, l'_{i,j}$  分别表示矩阵  $\mathbf{M}, \mathbf{L}, \mathbf{S}, \mathbf{L}'$  中的元素,  $G(\cdot)$  是  $\mathbf{L}$  和  $\mathbf{S}$  矩阵邻域像素相似性判断函数,  $r$  是邻域范围,  $v$  和  $u$  是阈值。

更新函数  $F(\cdot)$  的作用体现在: 1) 当背景矩阵  $\mathbf{L}$  的  $l_{i,j}$  与输入视频矩阵  $\mathbf{M}$  的  $m_{i,j}$  相等时,  $m_{i,j}$  就是真实背景; 2) 当背景矩阵  $\mathbf{L}$  的  $r$  个邻域帧对应  $(i, j)$  位置的像素中有  $v$  个以下不相似, 则可判定  $(i, j)$  位置是阴影, 更新  $l'_{i,j} = m_{i,j}$ ; 3) 当稀疏矩阵  $\mathbf{S}$  的  $r$  个邻域帧对应  $(i, j)$  位置的像素中有  $u$  个以下不相似, 则可判定  $(i, j)$  位置是噪声, 更新  $l'_{i,j} = m_{i,j}$ , 否则不更新。更新后的背景模型  $\mathbf{L}' = [\mathbf{L}^1, \mathbf{L}^2, \dots, \mathbf{L}^n]$  ( $\mathbf{L}^k = \mathbf{L}'(:, k)$  ( $k = 1, 2, \dots, n$ )) 包含了背景、阴影和噪声。

将  $\mathbf{L}'$  代入式(5), 可求得目标  $\mathbf{A}$ :

$$\mathbf{A} = \mathbf{M} - \mathbf{L}' \quad (5)$$

将连续性约束背景更新模型应用于背景减除运动目标检测的优势在于能够准确构建背景模型, 有效提高了运动目标检测的准确性。使用该方法检测运动目标主要包括 3 个阶段: 低秩矩阵分解、背景模型更新和运动目标提取。

### 3 连续性约束背景减除目标检测算法的实现

#### 阶段 1 低秩分解

(1) 提取分辨率为  $p \times q$  的  $n$  帧输入视频中的每一帧, 并按相同的规则组成列矢量,  $n$  个列矢量按列放置于矩阵  $\mathbf{M}$  中,  $\mathbf{M}$  的大小为  $(p \times q) \times n$ ;

(2) 采用 IALM (Inexact Augmented Lagrange Multiplier, 非精确的增广拉格朗日乘子) 算法<sup>[13]</sup> 按式(1)分解矩阵, 可得低秩矩阵  $\mathbf{L} \in R^{(p \times q) \times n}$  和稀疏矩阵  $\mathbf{S} \in R^{(p \times q) \times n}$ 。

#### 阶段 2 背景更新

(1) 背景模型更新: 按式(4)更新低秩矩阵  $\mathbf{L} \in R^{(p \times q) \times n}$ , 可得时间连续性约束背景模型  $\mathbf{L}' \in R^{(p \times q) \times n}$ , 其包含背景、阴影和噪声;

(2) 背景图像获取: 根据  $\mathbf{L}'$  按组成  $\mathbf{M}$  矩阵相同的规则恢复出与  $n$  帧视频  $M^1, M^2, \dots, M^n$  对应的  $n$  帧低秩图像  $L^1, L^2, \dots, L^n$  (大小为  $p \times q$ ), 即背景图像。

#### 阶段 3 运动目标提取

(1) 背景减除:  $n$  帧视频  $M^1, M^2, \dots, M^n$  分别对应减去  $n$  帧背景  $L^1, L^2, \dots, L^n$ , 可得背景减除后的误差矩阵  $\mathbf{E}^1, \mathbf{E}^2, \dots, \mathbf{E}^n$ ;

(2) 阈值分割: 计算每一帧目标与背景之间类间方差的最大值  $T$ , 将  $T$  作为自适应阈值分割误差矩阵  $\mathbf{E}^k$  ( $k = 1, 2, \dots, n$ ), 其像素为:

$$e^k(i, j) = \begin{cases} 0, & e^k(i, j) < T \\ e^k(i, j), & e^k(i, j) \geq T \end{cases} \quad (6)$$

二值化处理  $e^k(i, j)$ , 得:

$$a^k(i, j) = \begin{cases} 255, & e^k(i, j) \neq 0 \\ 0, & e^k(i, j) = 0 \end{cases} \quad (7)$$

元素  $a^k(i, j)$  组成的  $\mathbf{A}^k$  ( $k = 1, 2, \dots, n$ ) 为运动目标。

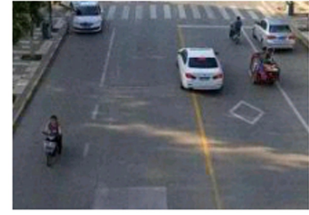
值得指出的是, 背景更新模型已有效去除阴影和噪声对运动目标提取的影响, 最大类间方差法<sup>[21]</sup> 确定的自适应分割

阈值有利于减少“空洞”, 进一步降低了错误分割的概率, 提升了运动目标检测的准确度。

### 4 实验结果及评价

本文仿真实验在 CPU 为 Intel Core i5-2400、主频为 3.1 GHz、内存为 4GB 的主机上使用 Matlab R2014b 完成。

图 2 为本文方法和低秩分解方法检测路面实拍视频的结果。图 2(a) 是加入  $\mu = 0, \sigma^2 = 0.2$  高斯白噪声的视频帧, 图 2(b) 是低秩分解建立的背景模型, 图 2(c) 是低秩分解结果, 图 2(d) 是本文方法建立的背景模型, 图 2(e) 是本文方法结果。实验参数为:  $\mathbf{M} = 225 \times 320, \lambda = 0.056, r = 15, v = 12, u = 12$ 。



(a) 输入视频



(b) 低秩分解背景



(c) 低秩分解结果



(d) 本文方法背景



(e) 本文方法结果

图 2 实拍视频的运动目标检测结果

Fig. 2 Results of moving target detection on real video shoot

图3为本文方法、IALM<sup>[13]</sup>、PSPG<sup>[14]</sup>、RPCA<sup>[15]</sup>、GoDec<sup>[16]</sup>和MAMR<sup>[17]</sup>方法对视频数据集 MovedObject 中某视

频建立的背景模型及提取运动目标的结果。实验参数为： $M=250 \times 320, \lambda=0.056, r=11, v=7, u=7$ 。

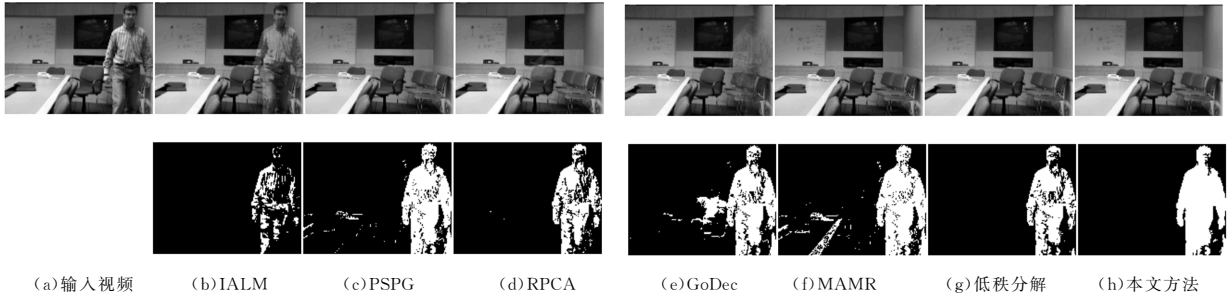


图3 MovedObject 数据库上的运动目标检测结果

Fig. 3 Results of moving object detection on MovedObject dataset

观察图2和图3可以看出,本文方法相比于低秩分解、IALM、PSPG、RPCA、GoDec和MAMR方法能够建立更准确的背景模型,克服阴影和噪声的影响,避免“空洞”。

为了验证本文方法的鲁棒性及效率,表1列出Wallflower数据库中全部7组视频分别采用IALM、PSPG、RPCA、

GoDec、MAMR和本文方法检测运动目标的结果的FM值(F-Measure)及各算法运行时间T。FM指标综合考虑了准确率和召回率两个重要参数,常用于评价运动目标检测质量。由表1可知,本文方法运算时间稍长,但具有更高的FM值,且鲁棒性更好。

表1 6种目标检测方法的FM值及运行时间T

Table 1 FM and T values of six target detection methods

Video	IALM(FM/T)	PSPG(FM/T)	RPCA(FM/T)	GoDec(FM/T)	MAMR(FM/T)	Ours(FM/T)
Bootstrap	0.6079/9.97 s	0.5604/25.45 s	0.5691/19.63 s	0.5806/6.78 s	0.5109/22.66 s	0.6953/34.22 s
Camouflage	0.3027/9.78 s	0.6248/24.09 s	0.2858/19.66 s	0.7051/6.49 s	0.6527/22.86 s	0.7768/33.65 s
Foreground Aperture	0.6729/8.10 s	0.6945/23.48 s	0.3011/17.09 s	0.295/6.76 s	0.7166/21.63 s	0.7206/32.49 s
LightSwitch	0.6480/9.99 s	0.5756/23.85 s	0.6255/15.86 s	0.3855/6.69 s	0.5473/21.76 s	0.7181/34.10 s
MovedObject	0.8744/8.95 s	0.7285/24.19 s	0.8923/15.79 s	0.5354/6.77 s	0.8153/21.61 s	0.8972/31.17 s
TimeofDay	0.6724/9.77 s	0.6715/23.22 s	0.6715/15.28 s	0.5158/6.98 s	0.6712/22.23 s	0.6733/32.41 s
WavingTrees	0.6567/8.83 s	0.5529/23.69 s	0.7177/17.55 s	0.5065/6.66 s	0.6976/24.77 s	0.7629/34.10 s

图4给出Wallflower数据库中全部7组视频分别采用IALM、PSPG、RPCA、GoDec、MAMR和本文方法检测运动目标平均结果的ROC曲线。可用曲线下的面积来评价运动目标检测的准确性。由ROC曲线可见,相比于其他方法,本文方法ROC曲线右下方包围的面积最大,说明其运动目标检测效果最好。

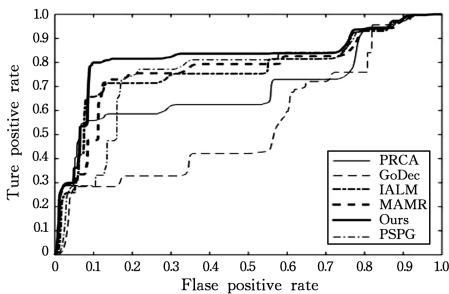


图4 ROC曲线

Fig. 4 ROC curves

**结束语** 综上所述,背景模型的准确构建对于背景减除运动目标检测的质量至关重要。本文提出的时间连续性约束背景更新模型能够建立准确的背景模型,有效克服阴影和噪声对运动目标提取的影响,相比于目前主流的背景减除运动目标检测方法,提取的运动目标更加完整,且鲁棒性更好。

## 参考文献

- [1] ZHAO Y, SHI H, CHEN X, et al. An overview of object detection and tracking[C]//IEEE International Conference on Information and Automation. 2015:280-286.
- [2] XUE L X, LUO Y L, WANG Z C. Detection algorithm of adaptive moving objects based on frame difference method[J]. Application Research of Computers, 2011, 28(4): 1551-274.
- [3] GAO P, SUN X, WANG W. Moving object detection based on kirsch operator combined with Optical Flow[C]// International Conference on Image Analysis and Signal Processing. 2010:620-624.
- [4] BARNICH O, DROOGENBROECK M V. ViBe: A Universal Background Subtraction Algorithm for Video Sequences[J]. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2011, 20(6): 1709-1724.
- [5] WENG M, HUANG G, DA X. A new interframe difference algorithm for moving target detection[C]// International Congress on Image and Signal Processing. 2010: A52-A53.
- [6] HUANG S Q, LI G. A Review of Research on Moving Target Detection Technology in Intelligent Video Surveillance System [J]. Information and Communications, 2012, 120(4): 57-58. (in Chinese)

- 黄斯茜,李光.智能视频监控系统中运动目标检测技术研究综述[J].信息通信,2012,120(4):57-58.
- [7] LI Q L, HE J F. Vehicle Detection Based on Three-Frame Difference Method and Cross Entropy Threshold Method[J]. Computer Engineering, 2011, 37(4): 172-174. (in Chinese)  
李秋林,何家峰.基于三帧差法和交叉熵阈值法的车辆检测[J].计算机工程,2011,37(4):172-174.
- [8] ZUO F Y, GAO S F, HAN J Y. Moving Object Detection and Tracking Based on Weighted Accumulative Difference[J]. Computer Engineering, 2009, 35(22): 159-161.
- [9] KROEGER T, TIMOFTE R, DAI D, et al. Fast Optical Flow Using Dense Inverse Search [M] // Computer Vision-ECCV 2016. Springer International Publishing, 2016: 471-488.
- [10] YANG H, QU S. Real-time vehicle detection and counting in complex traffic scenes using background subtraction model with low-rank decomposition[J]. Iet Intelligent Transport Systems, 2018, 12(1): 75-85.
- [11] LI X. Research on Moving Target Detection Method in Intelligent Video Surveillance Images[J]. Wireless Interconnect Technology, 2013(8): 158-158. (in Chinese)  
李想.智能视频监控图像中运动目标检测方法的研究[J].无线互联科技,2013(8):158-158.
- [12] ZHOU M, SONG Z J. Video background modeling based on sparse and low rank matrix decomposition[J]. Application Research of Computers, 2015, 32(10): 3175-3178. (in Chinese)  
周密,宋占杰.基于稀疏与低秩矩阵分解的视频背景建模[J].计算机应用研究,2015,32(10):3175-3178.
- [13] LIN Z, CHEN M, MA Y. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low-Rank Matrices [J]. arXiv:1009.5055.
- [14] AYBAT N S, GOLDFARB D, MA S. Efficient algorithms for robust and stable principal component pursuit problems [J]. Computational Optimization & Applications, 2014, 58(1): 1-29.
- [15] LIU G, YAN S. Active subspace: toward scalable low-rank learning [J]. Neural Computation, 2012, 24(12): 3371-3394.
- [16] ZHOU T, TAO D. GoDec: Randomized Lowrank & Sparse Matrix Decomposition in Noisy Case [C] // International Conference on Machine Learning. DBLP, 2011: 33-40.
- [17] YE X, YANG J, SUN X, et al. Foreground-Background Separation From Video Clips via Motion-Assisted Matrix Restoration [J]. IEEE Transactions on Circuits & Systems for Video Technology, 2015, 25(11): 1721-1734.
- [18] WRIGHT J, GANESH A, RAO S, et al. Robust Principal Component Analysis: Exact Recovery of Corrupted Low-Rank Matrices [C] // Neural Networks for Signal Processing X, 2000. Proceedings of the 2000 IEEE Signal Processing Society Workshop. IEEE, 2009: 289-298.
- [19] VIDAL R, MA Y, SASTRY S S. Robust Principal Component Analysis [J]. Journal of the Acm, 2016, 58(3): 1-37.
- [20] LIU G, LIN Z, YU Y. Robust Subspace Segmentation by Low-Rank Representation [C] // International Conference on Machine Learning. DBLP, 2010: 663-670.
- [21] LI S S, AN J B, LI C G, et al. Sports Ship Detection Method Based on Background Difference and Visual Saliency [J]. Internet of Things Technology, 2018(1): 17-20.
- [22] GU S, XIE Q, MENG D, et al. Weighted Nuclear Norm Minimization and Its Applications to Low Level Vision [J]. International Journal of Computer Vision, 2017, 121(2): 183-208.
- [23] MEI D, ZHU X, WANG X, et al. Image super-resolution based on structural dissimilarity learning dictionary [C] // International Conference on the Frontiers and Advances in Data Science. IEEE, 2018: 12-17.