

采用低秩与加权稀疏分解的 视频前景检测算法

常 侃^{1,2,3}, 张智勇¹, 陈 诚¹, 覃团发^{1,2,3}

(1. 广西大学计算机与电子信息学院, 广西南宁 530004; 2. 广西大学广西多媒体通信与网络技术重点实验室
培育基地, 广西南宁 530004; 3. 广西大学广西高校多媒体通信与信息处理重点实验室, 广西南宁 530004)

摘 要: 传统的鲁棒主成分分析模型能较好地解决视频前景检测问题. 但是, 若该模型的假设条件不能满足, 算法性能会变差. 针对此问题, 本文提出了一种低秩与加权稀疏分解模型, 通过对前景矩阵加权以增强其稀疏性. 在建立加权矩阵的过程中, 采用光流法获取每帧的运动矢量, 以区分真实运动区域. 其次, 进一步提出一种增强模型, 通过将加权矩阵作用于观测矩阵及背景矩阵, 防止前景与背景的错误分离. 实验结果表明, 在无噪和有噪的情况下, 提出的算法均能有效地分离监控视频中的前景和背景.

关键词: 前景检测; 运动目标检测; 鲁棒主成分分析; 低秩表示; 光流法

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2017)09-2272-09

电子学报 URL: <http://www.ejournal.org.cn> **DOI:** 10.3969/j.issn.0372-2112.2017.09.031

Video Foreground Detection by Low-Rank and Reweighted Sparse Decomposition

CHANG Kan^{1,2,3}, ZHANG Zhi-yong¹, CHEN Cheng¹, QIN Tuan-fa^{1,2,3}

(1. School of Computer and Electronic Information, Guangxi University, Nanning, Guangxi 530004, China;
2. Guangxi Key Laboratory of Multimedia Communications and Network Technology (Cultivating Base),
Guangxi University, Nanning, Guangxi 530004, China; 3. Guangxi Colleges and Universities Key Laboratory
of Multimedia Communications and Information Processing, Guangxi University, Nanning, Guangxi 530004, China)

Abstract: The traditional robust principal component analysis (RPCA) model is able to solve the video foreground detection problem well. However, if the basic assumptions are violated, this model will have poor performance. This paper proposes a low rank and reweighted sparse decomposition model, where the foreground matrix is reweighted so as to enhance its sparsity. When the weighting matrix is established, the optical flow method is used to get the motion vectors in each frame in order that the real moving areas can be recognized. Afterwards, based on the newly proposed model, an enhanced decomposition model is also developed. Since the weighting matrix is applied to both the observation matrix and the background matrix, the enhanced model is able to prevent the foreground and the background from being wrongly separated. Experimental results show that the proposed algorithm can efficiently separate foreground and background components for video clips with or without noises.

Key words: foreground detection; moving object detection; robust principal component analysis (RPCA); low-rank representation; optical flow method

1 引言

视频前景检测也称为运动目标检测, 是智能视频监控系统中的重要组成部分. 前景检测结果的准确度, 对目标跟踪^[1]、分类、识别等一系列应用的性能有着较

大的影响.

已有的前景检测算法大致可以分为两类: 局部算法和非局部算法. 局部算法通常对每个像素点进行单独操作^[2-4], 分离出的前景和背景容易出现空域上的不连续; 此外, 这类方法对噪声、亮度变化等干扰较为

敏感.非局部算法更多地应用空域、时域信息进行前景检测^[2,5-8].近年来,基于鲁棒主成分分析(Robust Principal Component Analysis,RPCA)^[7,8]的算法得到了较好的发展,但是,在复杂的监控场景下,传统RPCA模型中的假设条件不易满足,从而导致该算法性能下降.

为了进一步提高RPCA模型的性能,一些研究人员对其进行了改进.例如,Peng等人^[9]对低秩矩阵的奇异值及稀疏矩阵进行加权,提高了低秩与稀疏分解算法的求解精度.但是,Peng等人的方法并未有效应用前景或背景的特征.根据在实际应用中前景的像素点通常是密集地分布在小块区域的特性,Liu^[10]提出低秩和结构化的稀疏分解算法.Ye等人^[11]将前景的运动信息用于矩阵恢复,对运动区域上前景和背景之和进行加权,以保证前景和背景叠加结果的准确性.在Wen等人^[12]的工作中,同样运用了运动信息,通过将完整的监控视频序列划分为若干组,以确保获取更准确的低秩和稀疏分量.Zhao等人^[13]认为不应使用 l_1 范数或 l_2 范数来对前景进行建模,提出在贝叶斯框架下将混合高斯模型和传统RPCA模型结合,在前景检测效果上有一定的提升.

不同于前述已有算法,本文提出了一种新的低秩与加权稀疏分解模型.根据视频各帧的运动矢量建立加权矩阵并对前景矩阵进行加权,从而提高加权前景的稀疏性.进一步地,通过将加权矩阵作用于观测矩阵及背景矩阵,从而防止前景与背景的错误分离.

2 基于传统RPCA模型的视频前景检测

对于包含 n 帧的监控视频序列,可以将每一帧向量化为一个 m 维的列向量,由此整个视频序列可以视为一个大小为 $m \times n$ 的观测矩阵 D .在传统的RPCA模型^[7]中,通过求解下述最优化问题获取低秩背景矩阵 A 以及稀疏前景矩阵 E :

$$\{\hat{A}, \hat{E}\} = \arg \min_{A, E} \|A\|_* + \lambda \|E\|_1, \text{ s. t. } D = A + E \quad (1)$$

其中, \hat{A}, \hat{E} 分别是重建的背景矩阵以及重建的稀疏前景矩阵, λ 是权重系数, $\|A\|_*$ 表示矩阵 A 的核范数, $\|E\|_1$ 表示矩阵 E 的 l_1 范数.有许多方法可以求解式(1)^[8],例如非精确增广拉格朗日数乘法(Inexact Augmented Lagrange Multipliers, IALM)等.

但是,若监控视频中的运动物体速度过慢或体积过大,采用上述传统RPCA模型得到的结果较差.一方面,若速度缓慢,运动物体在连续帧中出现较多重叠区域,不满足真实运动区域应随机、均匀分布于稀疏矩阵 E 中的假设^[7].若直接应用模型式(1),则重叠区域容易被误判为低秩背景.另一方面,若运动物体的体积较大,此情况下前景矩阵 E 稀疏的条件不满足,容易在分离的背景中出现鬼影和拖尾.

3 采用低秩与加权稀疏分解的前景检测算法

3.1 基本的低秩与加权稀疏分解模型及算法

为了解决第2节所述传统RPCA模型式(1)的问题,受到课题组前期研究工作的启发^[14],本文提出对式(1)中的稀疏矩阵 E 进行加权,得到下述新的低秩与加权稀疏分解模型:

$$\{\hat{A}, \hat{E}\} = \arg \min_{A, E} \|A\|_* + \lambda \|W \circ E\|_1, \text{ s. t. } D = A + E \quad (2)$$

其中 W 是加权矩阵, \circ 代表两矩阵中的元素对位相乘. W 在运动区域取较小的系数,在背景区域取较大的系数.在求解式(2)时,加权矩阵 W 是独立的参数.若 W 设置正确,求解式(2)能提升检测准确性,原因在于:

(1)若视频中包含缓慢运动或占据较大区域的物体,可通过给运动区域分配较小的加权系数,增强加权后前景矩阵的稀疏性,提升检测准确度;

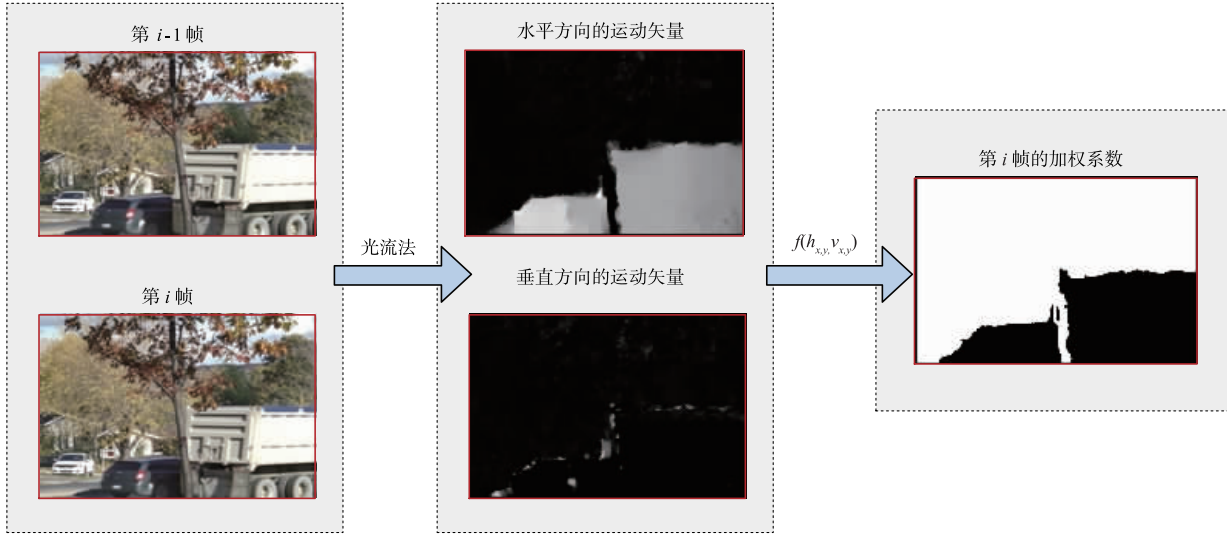
(2)对前景矩阵 E 中的背景区域分配较大的系数,可强调这些区域值的近零特性,以防止一些离散的、不可靠的像素点被错判为运动目标.由此增强了模型对某些特殊监控场景,如包含动态背景的场景的鲁棒性.

为了正确设定加权矩阵 W ,需要先确定每帧中的运动区域.本文采用如图1所示的流程分配第 i 帧的加权系数.对于第 i 帧,首先采用光流法以第 $i-1$ 帧为参考进行运动估计,提取水平和垂直方向的运动矢量矩阵 H_i 及 V_i ;其次,将 H_i 和 V_i 中的相同位置元素 $h_{x,y}$ 及 $v_{x,y}$ 代入权重计算函数 $f(h_{x,y}, v_{x,y})$,得到第 i 帧中坐标为 (x, y) 处的加权系数值;然后,将第 i 帧所有加权系数按列向量形式排列,获得第 i 帧的加权向量 w_i ;最后,将 n 帧图像对应的加权向量合并,构成整个视频序列的加权矩阵 W ,即 $W = [w_1, w_2, \dots, w_i, \dots, w_n]$.

可采用多种方法构造函数 $f(h_{x,y}, v_{x,y})$.基本的构造原则是对运动区域分配小系数,对背景区域分配大系数,以强调加权结果的稀疏性.在本文中, $f(h_{x,y}, v_{x,y})$ 的定义如下:

$$f(h_{x,y}, v_{x,y}) = \begin{cases} 1, & \text{当 } \frac{1}{0.01 + \sqrt{h_{x,y}^2 + v_{x,y}^2}} \geq T \\ 0, & \text{当 } \frac{1}{0.01 + \sqrt{h_{x,y}^2 + v_{x,y}^2}} < T \end{cases} \quad (3)$$

其中 T 为预先设定的门限值.从式(3)可知,当某位置的像素运动较大时,判定其属于运动区域,从而分配最小的系数0;反之,判定其属于背景区域,分配最大的系数1.另外,若 T 值选取合理,能有效区分真实运动区域及动态背景.在图1的例子中,虽包含动态背景(树叶摆动),式(3)仍能准确分配加权系数.本文通过实验设定 T 的经验值,取值在4.1节给出;此外,还将在4.1节讨论为何选取加权系数值为0或1.

图1 第 i 帧加权系数的分配过程

可对 IALM 算法^[15] 略微调整, 得到式(2)的求解方法. 为了与 IALM 算法^[15] 区分, 将求解式(2)的算法命名为重加权 IALM (Reweighted-IALM, R-IALM). R-IALM 与 IALM 的区别在于: 首先, 需要将加权矩阵 \mathbf{W} 作为输入参数; 其次, 更新前景矩阵 \mathbf{E} 时, 收缩算子的比较门限需更改为 $(\lambda/\mu^k) \mathbf{W}$ (μ^k 为传统 IALM 算法的迭代中间变量).

从流程上看, R-IALM 算法与文献[9]的 NSVT 算法相似. 但是, 在 NSVT 算法中, 加权矩阵 \mathbf{W} 的每个系数被设定为矩阵 \mathbf{E} 中对应位置元素的绝对值的倒数, 其缺陷是: 首先, 并未考虑监控视频中前景的特征; 其次, 需要迭代更新 \mathbf{W} . 本文的加权方法考虑真实运动轨迹, 因而能有效辅助真实前景的分离; 另外, 加权矩阵 \mathbf{W} 的建立独立于 R-IALM 算法, 无须进行多次迭代更新.

3.2 增强的低秩与加权稀疏分解模型及算法

无论是传统 RPCA 模型式(1)还是本文提出的模型式(2), 都存在一项重要的先验假设, 即 $\mathbf{D} = \mathbf{A} + \mathbf{E}$. 在观测矩阵 \mathbf{D} 的背景区域, 上述假设是成立的; 但是在 \mathbf{D} 的运动区域, 由于前景遮挡背景, 并不能简单认为前景与背景之和与观测值相等. 因此, 若直接使用 R-IALM 算法, 在运动区域可能出现前景和背景的错误分离. 为解决上述问题, 进一步提出增强模型, 通过顺序求解式(4)与式(5)分离出背景和前景:

$$\hat{\mathbf{A}} = \arg \min_{\mathbf{A}} \frac{1}{2} \|\mathbf{W} \circ (\mathbf{D} - \mathbf{A})\|_{\text{F}}^2 + \alpha \|\mathbf{A}\|_{*} \quad (4)$$

$$\hat{\mathbf{E}} = \arg \min_{\mathbf{E}} \frac{1}{2} \|\mathbf{D} - \hat{\mathbf{A}} - \mathbf{E}\|_{\text{F}}^2 + \lambda \|\mathbf{W} \circ \mathbf{E}\|_{1} \quad (5)$$

其中 λ 和 α 是权重系数, $\|\cdot\|_{\text{F}}$ 表示矩阵的 Frobenius 范数.

在式(4)中, 由于加权矩阵 \mathbf{W} 的作用, 第一项约束仅强调观测矩阵 \mathbf{D} 中的背景区域应与背景矩阵 \mathbf{A} 一致; 为了恢复出被前景遮挡的背景, 通过第二项约束强调 \mathbf{A} 的低秩特性. 与式(2)中的约束 $\mathbf{D} = \mathbf{A} + \mathbf{E}$ 相比, 式(4)去除了前景对背景的影响, 防止在背景中出现鬼影. 另一方面, 式(5)可视为在已知 \mathbf{A} 的前提下式(2)的简化. 将式(4)得到的更可靠的背景 \mathbf{A} 代入式(5), 得到的前景 \mathbf{E} 也会比式(2)更准确.

为了求解式(4), 引入新变量 \mathbf{L} , 并令 $\mathbf{L} = \mathbf{A}$, 将式(4)转化为如下约束优化问题:

$$\{\hat{\mathbf{A}}, \hat{\mathbf{L}}\} = \arg \min_{\mathbf{A}, \mathbf{L}} \frac{1}{2} \|\mathbf{W} \circ (\mathbf{D} - \mathbf{A})\|_{\text{F}}^2 + \alpha \|\mathbf{L}\|_{*} \quad (6)$$

s. t. $\mathbf{L} = \mathbf{A}$

进一步将式(6)转化成无约束优化问题:

$$\{\hat{\mathbf{A}}, \hat{\mathbf{L}}\} = \arg \min_{\mathbf{A}, \mathbf{L}} \frac{1}{2} \|\mathbf{W} \circ (\mathbf{D} - \mathbf{A})\|_{\text{F}}^2 + \alpha \|\mathbf{L}\|_{*} + \frac{\beta}{2} \|\mathbf{L} - \mathbf{A} - \mathbf{d}\|_{\text{F}}^2 \quad (7)$$

其中, \mathbf{d} 为根据 Bregman 迭代方法^[16] 引入的迭代中间变量, β 为权重系数. 将式(7)拆分为 \mathbf{A} 与 \mathbf{L} 子问题并依次迭代求解, 第 $k+1$ 次循环迭代的表达式如下:

$$\mathbf{A}^{k+1} = \arg \min_{\mathbf{A}} \frac{1}{2} \|\mathbf{W} \circ (\mathbf{D} - \mathbf{A})\|_{\text{F}}^2 + \frac{\beta}{2} \|\mathbf{L}^k - \mathbf{A} - \mathbf{d}^k\|_{\text{F}}^2 \quad (8)$$

$$\mathbf{L}^{k+1} = \arg \min_{\mathbf{L}} \alpha \|\mathbf{L}\|_{*} + \frac{\beta}{2} \|\mathbf{L} - \mathbf{A}^{k+1} - \mathbf{d}^k\|_{\text{F}}^2 \quad (9)$$

求解 \mathbf{A} 子问题时, 通过对目标函数求导, 并令导数为 0, 由此可求得解为:

$$\mathbf{A}^{k+1} = (\mathbf{W} \circ \mathbf{D} + \beta(\mathbf{L}^k - \mathbf{d}^k)) ./ (\mathbf{W} + \beta \mathbf{1}) \quad (10)$$

其中, $./$ 代表两矩阵中的元素对位相除, $\mathbf{1}$ 是全 1 矩阵. 求

解 L 子问题时,先按照式(11)进行奇异值分解(Singular Value Decomposition, SVD),然后采用式(12)进行计算^[15],其中 $(x)_+ = \max(x, 0)$.

$$(U, G, V) = \text{SVD}(A^{k+1} + d^k) \quad (11)$$

$$L^{k+1} = U \left(G - \text{diag} \left(\frac{\alpha}{\beta}, \frac{\alpha}{\beta}, \dots, \frac{\alpha}{\beta} \right) \right)_+ V^T \quad (12)$$

最后,根据 Bregman 迭代过程^[16], d^{k+1} 迭代更新式为:

$$d^{k+1} = d^k + A^{k+1} - L^{k+1} \quad (13)$$

当求解式(5)时,可直接写出该问题的解为:

$$\hat{E} = \text{shrink}(D - \hat{A}, \lambda W) \quad (14)$$

$\text{shrink}(\cdot)$ 代表收缩算子,对于输入 x 与门限 τ ,其定义如下:

$$\text{shrink}(x, \tau) = \frac{x}{|x|} \cdot \max(|x| - \tau, 0) \quad (15)$$

将基于式(4)和式(5)的算法命名为 R-SB (Reweighted-Split Bregman),流程总结为算法 1.

算法 1 R-SB 算法

输入: $D, W, \lambda, \alpha, \beta$

输出: $\{\hat{A}, \hat{E}\}$

初始化: $A^0 = D, L^0 = D, d^0 = 0$

For $k = 1, 2, \dots, N$

 分别根据式(10)、式(12)和式(13)更新 A^{k+1}, L^{k+1} 和 d^{k+1}

 If $\|A^{k+1} - A^k\|_F / \|A^k\|_F < 10^{-4}$ Break

End For

令 $\hat{A} = A^{k+1}$ 并根据式(14)计算 \hat{E}

4 实验结果与分析

4.1 实验设置

使用 100 帧 180×120 的 Office、Fall、Snowfall、Busstation, 190×122 的 Peopleinshade 这 5 个视频序

列^[17]进行实验.测试序列包含了缓慢运动、动态背景、恶劣天气、阴影等较有挑战的场景,因而能较全面地验证算法性能.所比较的算法为 GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13].

采用了三种客观指标评价前景检测效果:(1)查全率(recall): $R = a/(a+c)$,其中 a 表示属于前景的像素点被正确检测的数目, c 是属于前景的像素点被错检为背景的数目;(2)查准率(precision): $P = a/(a+b)$,其中 b 代表了属于背景的像素点被错检为前景的数目;(3)加权调和平均值: $F = 2RP/(R+P)$.

采用了两种客观指标评价背景分离效果:(1)相对重建误差(Relative Reconstruction Error, RRE): $RRE = \|\hat{A} - A\|_F / \|A\|_F$; (2)峰值信噪比(Peak Signal-to-Noise Ratio, PSNR): $PSNR = 20 \lg(255 / \sqrt{MSE})$,其中 MSE 是真实背景和重建背景的均方误差.

本文算法参数在实验中采用统一的取值.在 R-IALM 中, λ 在无噪情况下取 $1.2/\sqrt{m}$;有噪情况下为 $1.8/\sqrt{m}$,其中 m 为 D 的行数.在 R-SB 中,设置 $\beta = 4.5$,无噪情况下 $\alpha = 600\beta$, $\lambda = 12$;有噪情况下 $\alpha = 900\beta$, $\lambda = 30$.两算法的迭代次数上限 N 设置为 200;式(3)中的门限值 T 设置为 0.99.在计算 W 时,采用 Liu^[18]实现的光流法计算各帧的运动矢量.在式(3)中,将背景区域加权值设定为 1,运动区域设置为 0,原因在于:

(1) W 可视为二值掩膜(binary mask),能有效区分背景区域及运动区域,保证模型式(4)的准确性.

(2) 由式(2)可知,尽管背景区域加权值设为 1,仍是通过 λ 控制该区域的实际加权值.对于运动区域,测试了不同加权值下 R-IALM 算法的客观结果,如表 1 所示.可见,运动区域的最佳加权值为 0.

表 1 运动区域不同加权值对 R-IALM 算法结果的影响

运动区域 加权值	Office		Fall		Snowfall		Peopleinshade		Busstation	
	F	PSNR	F	PSNR	F	PSNR	F	PSNR	F	PSNR
0.00	0.84	35.11	0.89	22.39	0.90	29.01	0.90	38.45	0.82	36.89
0.05	0.84	35.11	0.88	22.37	0.90	28.89	0.90	38.45	0.82	36.88
0.10	0.84	35.09	0.88	22.14	0.89	28.78	0.90	38.44	0.82	36.87
0.15	0.84	35.04	0.87	21.60	0.88	28.67	0.90	38.44	0.82	36.86
0.20	0.84	34.97	0.86	20.89	0.87	28.57	0.90	38.34	0.81	36.82
0.25	0.84	34.85	0.85	20.01	0.86	28.48	0.89	38.22	0.81	36.77
0.30	0.83	34.67	0.83	18.90	0.85	28.39	0.89	38.03	0.81	36.72

4.2 无噪情况下的实验结果

表 2 列出了 GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13] 以及本文算法在无噪序列上的前景检测客观结果.可见,除了 Busstation 外, R-IALM 及 R-SB 获得的 F 值都明显高于其他算法;在 Office、Fall、Peopleinshade

上, R-IALM 与 R-SB 算法的 F 值相等,但在 Snowfall、Busstation 上, R-SB 算法优于 R-IALM 算法.图 2 展示了 6 种算法所检测的前景的主观质量比较.在 Busstation 上,除了 GMM^[3] 算法外,其余算法都获得了较准确的前景.在 Office、Fall、Snowfall 和 Peopleinshade 上, GMM^[3]、

RPCA^[7]、NSVT^[9]、MoG-RPCA^[13] 均存在不同程度的漏检和错检;相比之下,本文算法 R-IALM 及 R-SB 都取得了最佳检测结果.

表 2 无噪条件下前景检测客观结果

算法	Office			Fall			Snowfall			Peopleinshade			Busstation		
	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
文献[3]	0.39	0.96	0.54	0.46	0.95	0.62	0.25	0.60	0.35	0.51	0.97	0.67	0.36	0.97	0.53
文献[7]	0.65	0.93	0.76	0.57	0.95	0.71	0.36	0.98	0.52	0.72	0.96	0.82	0.62	0.99	0.76
文献[9]	0.67	0.93	0.77	0.59	0.94	0.72	0.48	0.98	0.64	0.71	0.95	0.81	0.65	0.99	0.79
文献[13]	0.69	0.89	0.77	0.59	0.89	0.70	0.53	0.99	0.68	0.74	0.99	0.84	0.69	0.98	0.81
R-IALM	0.75	0.96	0.84	0.81	0.99	0.89	0.84	0.97	0.90	0.82	0.98	0.90	0.71	0.98	0.82
R-SB	0.76	0.95	0.84	0.82	0.98	0.89	0.90	0.98	0.94	0.82	0.99	0.90	0.72	0.98	0.83

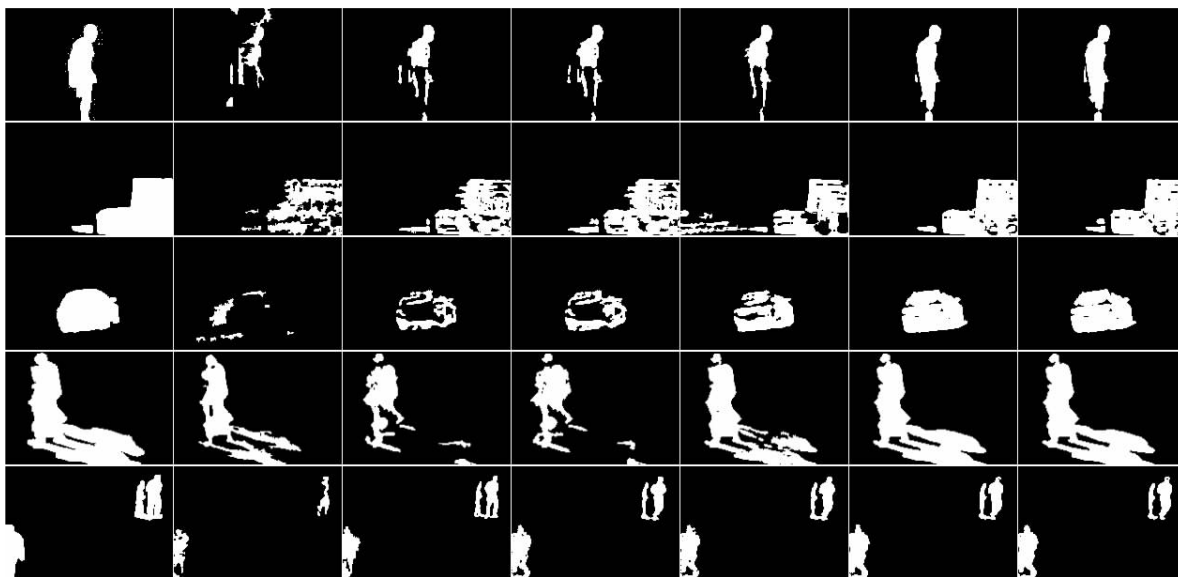


图 2 无噪条件下前景检测主观结果(从上到下依次为Office、Fall、Snowfall、Peopleinshade以及Busstation序列;从左到右依次为真实前景、GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13]、R-IALM及R-SB结果)

表 3 和图 3 分别展示了无噪序列上所分离背景的客观、主观结果.从表 3 可见,本文算法 R-SB 所分离的背景与真实背景最为接近.从图 3 可以观察到,

GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13] 的算法所分离的背景都存在不同程度的鬼影或拖尾. R-IALM 及 R-SB 防止了运动物体被错判为背景,获得了较理想的结果.

表 3 无噪条件下背景分离客观结果

算法	Office		Fall		Snowfall		Peopleinshade		Busstation	
	RRE	PSNR(dB)	RRE	PSNR(dB)	RRE	PSNR(dB)	RRE	PSNR(dB)	RRE	PSNR(dB)
文献[3]	0.147	23.10	0.503	12.48	0.110	25.07	0.167	21.47	0.204	20.23
文献[7]	0.110	26.58	0.416	14.76	0.084	27.66	0.089	30.10	0.042	34.47
文献[9]	0.104	26.98	0.411	14.94	0.080	27.98	0.085	30.18	0.041	34.57
文献[13]	0.096	27.00	0.340	16.86	0.068	29.33	0.027	38.28	0.042	34.18
R-IALM	0.043	35.11	0.154	22.39	0.701	29.01	0.020	38.45	0.031	36.89
R-SB	0.032	36.74	0.133	23.54	0.051	32.04	0.014	41.52	0.029	37.03

4.3 有噪情况下的实验结果

为了检测算法的鲁棒性,进一步测试了混合噪声条件下的算法性能.其中,高斯噪声的标准差设置为 5,脉冲噪声出现的概率设置为 10%.混合噪声条件下前景检测的客观结果列于表 4,主观结果如图 4 所示.可

见,绝大部分情况下 R-IALM 及 R-SB 可取得最佳结果;在背景杂乱的 Snowfall 上,R-SB 优于 R-IALM.

混合噪声下背景分离的客观结果列于表 5,主观结果如图 5 所示.从表 5 可见,R-SB 的结果明显好于其他算法.从图 5 可知,GMM^[3] 所分离背景质量极差;相较

于无噪情况,RPCA^[7]、NSVT^[9]、MoG-RPCA^[13]获取的背景质量均有降低;R-IALM 得到的背景也出现少量鬼影;只有 R-SB 分离出的背景无鬼影、无拖尾。

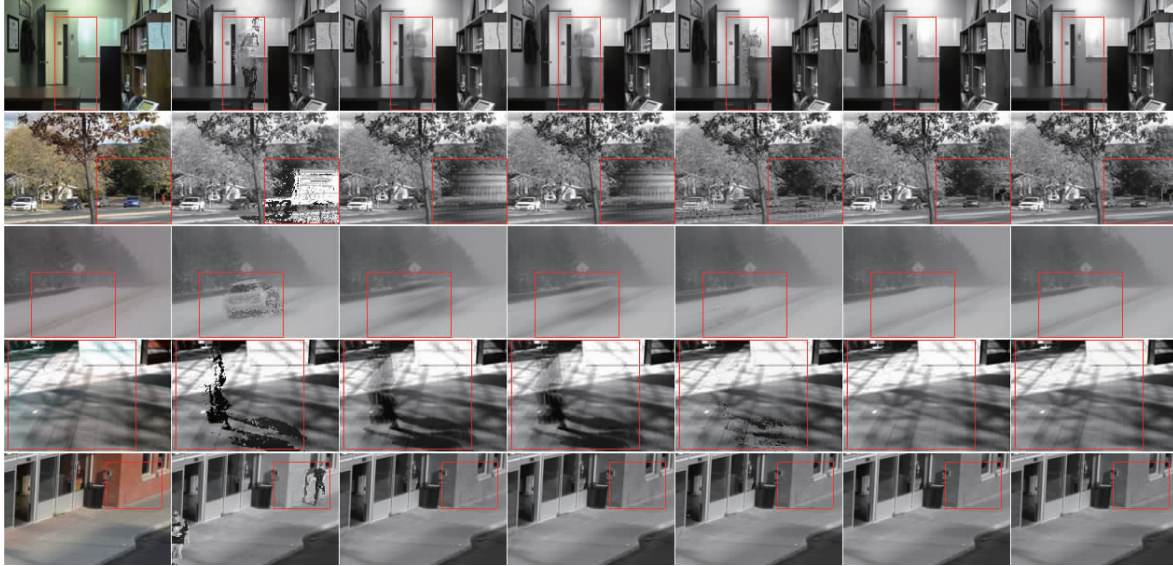


图3 无噪条件下背景分离主观结果(从上到下依次为Office、Fall、Snowfall、Peopleinshade以及Busstation序列;从左到右依次为真实背景、GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13]、R-IALM及R-SB结果)

表 4 混合噪声条件下前景检测客观结果

算法	Office			Fall			Snowfall			Peopleinshade			Busstation		
	R	P	F	R	P	F	R	P	F	R	P	F	R	P	F
文献[3]	0.36	0.94	0.52	0.42	0.87	0.57	0.15	0.50	0.23	0.46	0.94	0.62	0.35	0.92	0.51
文献[7]	0.48	0.91	0.62	0.44	0.95	0.58	0.27	0.98	0.40	0.62	0.98	0.76	0.53	0.99	0.69
文献[9]	0.59	0.89	0.70	0.49	0.95	0.62	0.37	0.98	0.52	0.66	0.97	0.78	0.52	0.99	0.68
文献[13]	0.59	0.84	0.67	0.44	0.94	0.59	0.34	0.96	0.50	0.62	0.99	0.75	0.54	0.94	0.68
R-IALM	0.67	0.96	0.79	0.78	0.96	0.86	0.69	0.96	0.80	0.74	0.99	0.85	0.65	0.98	0.78
R-SB	0.68	0.95	0.79	0.79	0.95	0.86	0.73	0.96	0.83	0.76	0.99	0.86	0.66	0.98	0.79

表 5 混合噪声条件下背景分离客观结果

算法	Office		Fall		Snowfall		Peopleinshade		Busstation	
	RRE	PSNR(dB)	RRE	PSNR(dB)	RRE	PSNR(dB)	RRE	PSNR(dB)	RRE	PSNR(dB)
文献[3]	0.442	13.38	0.614	10.17	0.330	15.25	0.324	14.25	0.428	13.74
文献[7]	0.103	26.50	0.400	15.10	0.079	28.04	0.080	29.84	0.050	32.49
文献[9]	0.098	26.92	0.394	15.28	0.074	28.62	0.074	30.92	0.045	33.46
文献[13]	0.106	25.90	0.333	16.82	0.058	30.84	0.038	32.68	0.045	33.20
R-IALM	0.069	29.69	0.180	21.04	0.075	28.25	0.046	32.84	0.048	32.50
R-SB	0.041	34.03	0.155	21.99	0.046	32.71	0.043	33.79	0.043	33.78

4.4 复杂度分析

R-IALM 与 IALM 算法^[15]相比仅是收缩门限不同,故两者复杂度一致.R-SB 需迭代更新 A 、 L 、 d 共 3 个变量.其中,更新 A 及 d 的复杂度都较低,复杂度主要集中在更新 L 时所需的 SVD 分解.对于单次 thin SVD 分解,其复杂度为 $O(mnr)$ ^[14],其中 m 是每帧像素数, n 是视频序列帧数, r 是待分解矩阵的秩.

表 6 列出了 5 个算法在无噪情况下的平均每帧运行时间.实验用 CPU 为 Intel (R) Pentium (R) G3250 @ 3.20GHz,运算平台为 Matlab.由于文献[3]的源码采用 C++ 实现,故在表 6 中没有比较其运行时间.可以总结出:(1) MoG-RPCA^[13]的平均运算时间最长;(2)由于 NSVT^[9]包含了内、外两层循环,若内循环收敛不快,会拖累整体运算速度,如 Office 序列;(3)借助高效的

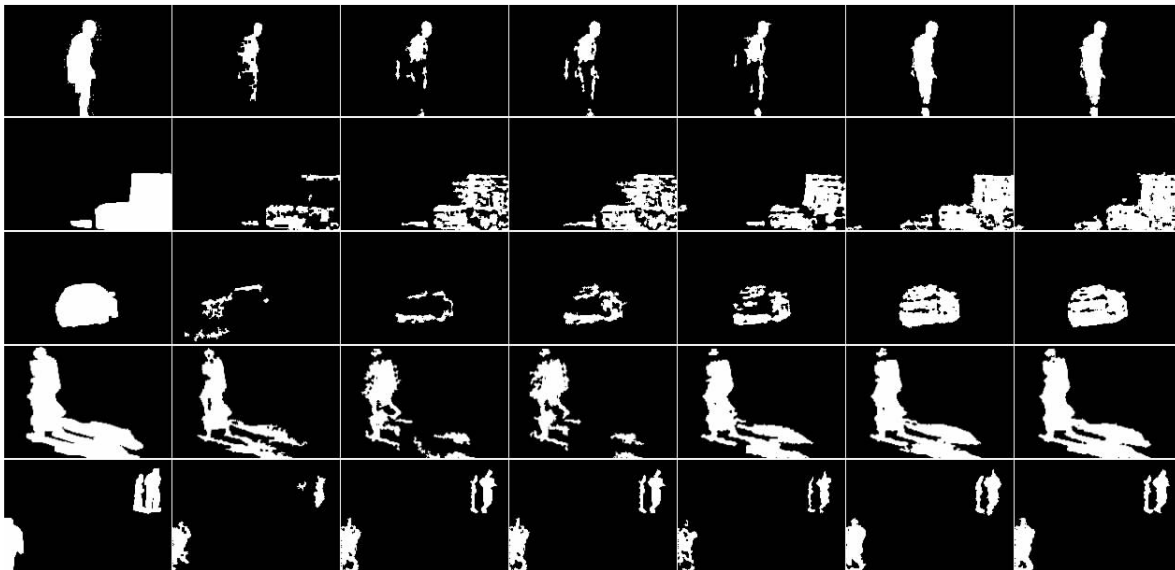


图4 混合噪声条件下前景检测主观结果(从上到下依次为Office、Fall、Snowfall、Peopleinshade以及Busstation序列;从左到右依次为真实前景、GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13]、R-IALM及R-SB结果)

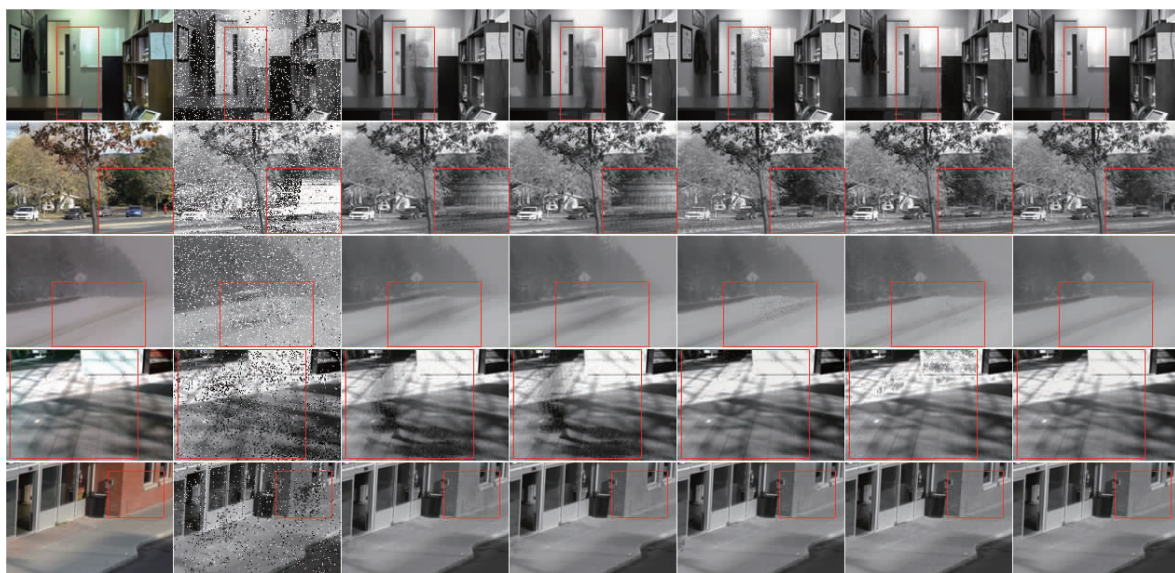


图5 混合噪声条件下背景分离主观结果(从上到下依次为Office、Fall、Snowfall、Peopleinshade以及Busstation序列;从左到右依次为真实背景、GMM^[3]、RPCA^[7]、NSVT^[9]、MoG-RPCA^[13]、R-IALM及R-SB结果)

IALM 算法^[15], R-IALM 收敛很快, 运算时间最短; (4) R-SB 在 Fall 序列上收敛相对较慢, 在其余序列上速度仅次于 R-IALM 算法.

因为 R-IALM 及 R-SB 都需要预先计算加权矩阵 \mathbf{W} , 所以需单独讨论该部分的复杂度. 在 Liu 所实现的光流法中^[18], 采用逐次超松弛迭代法 (Successive Over Relaxation, SOR) 求解文献[19]中的线性模型. 假设有线性问题 $\mathbf{H}\mathbf{x} = \mathbf{z}$, 其中 \mathbf{H} 为矩阵, \mathbf{x} 为未知矢量, \mathbf{z} 为已知矢量, N_n 代表矩阵 \mathbf{H} 中非零元素的个数, 对于单次 SOR 迭代, 其复杂度为 $O(N_n)$ ^[14]. 无噪条件下 5 个测试序列上各帧加权向量 \mathbf{w}_i 的平均生成时间如表 7

所示.

表 6 各算法平均每帧运行时间 (s)

算法	Office	Fall	Snowfall	Peopleinshade	Busstation
文献[7]	1.08	1.11	1.07	1.14	1.06
文献[9]	3.15	0.73	0.72	0.75	1.13
文献[13]	3.01	4.20	1.60	4.01	2.23
R-IALM	0.12	0.15	0.13	0.15	0.19
R-SB	0.41	1.02	0.49	0.55	0.61

表 7 各帧加权向量的平均生成时间(s)

Office	Fall	Snowfall	Peopleinade	Busstation
0.78	0.68	0.34	0.70	0.36

5 结论

为了更好地进行前景检测,本文对前景矩阵进行加权,以保证加权结果的稀疏性,由此提出了两种前景检测算法,即 R-IALM 与 R-SB. 实验结果证明,与已有的算法相比,R-IALM 及 R-SB 能在绝大部分情况下获得最佳的主观及客观结果;在背景较为复杂的情况下,R-SB 所检测的前景准确度高于 R-IALM 算法;在所有情况下,R-SB 能分离出比 R-IALM 质量更高的背景;在收敛速度上,R-IALM 快于 R-SB 算法. 在未来的工作中,需要进一步研究不同监控场景下算法的自适应参数调整方法.

参考文献

- [1] 杨威,付耀文,龙建乾,等. 基于有限集统计学理论的目标跟踪技术研究综述[J]. 电子学报,2012,40(7):1440-1448.
Yang W, Fu Y W, Long J Q, et al. The FISST-based target tracking techniques: a survey[J]. Acta Electronica Sinica, 2012, 40(7):1440-1448. (in Chinese)
- [2] Bouwmans T. Traditional and recent approaches in background modeling for foreground detection: an overview[J]. Computer Science Review, 2014, 11:31-66.
- [3] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction[A]. Proceedings of the 17th International Conference on Pattern Recognition [C]. Cambridge, UK: IEEE, 2004. 28-31.
- [4] Elgammal A, Duraiswami R, Harwood D, et al. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance[J]. Proceedings of the IEEE, 2002, 90(7):1151-1163.
- [5] Zhou Y, Xu W, Tao H, et al. Background segmentation using spatial-temporal multi-resolution MRF[A]. Proceedings of the 7th IEEE Workshop on Applications of Computer Vision/Motion and Video Computing [C]. Breckenridge, USA: IEEE, 2005. 8-13.
- [6] Reddy V, Sanderson C, Lovell B C. Improved foreground detection via block-based classifier cascade with probabilistic decision integration[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(1):83-93.
- [7] Candès E J, Li X, Ma Y, et al. Robust principal component analysis? [J]. Journal of the ACM, 2011, 58(3):1-37.
- [8] 彭义刚,索津莉,戴琼海,等. 从压缩感知到低秩矩阵恢复:理论与应用[J]. 自动化学报,2013,39(7):981-994.
Peng Y G, Suo J L, Dai Q H, et al. From compressed sensing to low-rank matrix recovery: theory and applications [J]. Acta Automatica Sinica, 2013, 39(7):981-994. (in Chinese)
- [9] Peng Y G, Suo J L, Dai Q H, et al. Reweighted low-rank matrix recovery and its application in image restoration [J]. IEEE Transactions on Cybernetics, 2014, 44(12):2418-2430.
- [10] Liu X, Zhao G Y, Yao J W, et al. Background subtraction based on low-rank and structured sparse decomposition [J]. IEEE Transactions on Image Processing, 2015, 24(8):2502-2514.
- [11] Ye X C, Yang J Y, Sun X, et al. Foreground-background separation from video clips via motion-assisted matrix restoration [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2015, 25(11):1721-1734.
- [12] Wen J J, Xu J, Zhan Y W, et al. Joint video frame set division and low-rank decomposition for background subtraction [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2014, 24(12):2034-2048.
- [13] Zhao Q, Meng D, Xu Z, et al. Robust principal component analysis with complex noise[A]. Proceedings of the 31th International Conference on Machine Learning [C]. Beijing, China: IEEE, 2014. 55-63.
- [14] Chang K, Ding P K, Li B. Compressive sensing reconstruction of correlated images using joint regularization [J]. IEEE Signal Processing Letters, 2016, 23(4):449-453.
- [15] Lin Z, Chen M, Wu L, et al. The Augmented Lagrange Multiplier Method for Exact Recovery of Corrupted Low Rank Matrices[R]. Urbana-Champaign: University of Illinois at Urbana-Champaign, 2009.
- [16] Goldstein T, Osher S. The split Bregman method for l_1 regularized problems [J]. SIAM Journal on Image Sciences, 2009, 2(2):322-343.
- [17] Wang Y, Jodoin P M, Porikli F, et al. CDnet 2014: an expanded change detection benchmark dataset[A]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops [C]. Columbus, OH, USA: IEEE, 2014. 387-394.
- [18] Liu C. Beyond Pixels: Exploring New Representations and Applications for Motion Analysis [D]. Cambridge: Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 2009.
- [19] Brox T, Bruhn A, Papenberg N, et al. High accuracy optical flow estimation based on a theory for warping[A]. Proceedings of the European Conference on Computer Vision [C]. Prague, Czech Republic: Springer, 2004. 25-36.

作者简介



常 侃(通信作者) 男,1983年6月生于广西南宁.2010年于北京邮电大学获通信与信息系统专业博士学位,现为广西大学计算机与电子信息学院副教授、硕士生导师.已发表论文四十余篇,主要研究方向为图像处理、视频编码等.

E-mail: pandack0619@163.com



张智勇 男,1991年9月出生于广西扶绥.现为广西大学计算机与电子信息学院计算机技术专业硕士研究生,主要研究方向为稀疏表示、低秩矩阵及应用等.

E-mail: zhangzhiyong1160@163.com



陈 诚 男,1993年10月出生于安徽舒城.现为广西大学计算机与电子信息学院信息处理与通信网络系统学术硕士研究生,主要研究方向为图像处理、稀疏表示等.

E-mail: vigil1993@163.com



覃团发 男,1966年7月出生于广西宾阳.1997年于南京大学获博士学位,现为广西大学计算机与电子信息学院副院长、教授、博士生导师.主要研究方向为多媒体通信、无线传感器网络等.

E-mail: tfqin@gxu.edu.cn