

# 基于稀疏表达残差的自然场景运动目标检测

蒋建国, 金玉龙, 齐美彬, 詹 曙

(合肥工业大学计算机与信息学院, 安徽合肥 230009)

**摘 要:** 本文提出一种基于稀疏表达残差的非参数化运动目标检测算法, 在假设前景变化相对静态背景可以视为稀疏残差的基础上, 采用视频前  $n$  帧初始化稀疏表达字典; 利用字典对后续视频帧进行重构, 提取每帧的重构残差; 结合基于光照强度的全局阈值矩阵, 将残差图像二值化, 提取图像前景; 利用前景区域和边缘点关系剔除 ghost 区域; 采用增量 PCA (Principal Component Analysis) 算法和保守更新的思想对背景模型进行更新. 在 *changedetection.net* 提供的 shadow 数据集上实验表明, 采用全局更新和残差计算的方法, 可以有效的解决由于自然场景光线变化导致的阴影变化, 并且对自然场景中背景的小幅度抖动和相机抖动等问题也具有一定的抵抗能力.

**关键词:** 背景建模; 残差; 运动目标检测; 稀疏表达

**中图分类号:** TP391      **文献标识码:** A      **文章编号:** 0372-2112 (2015)09-1738-07

**电子学报 URL:** <http://www.ejournal.org.cn>      **DOI:** 10.3969/j.issn.0372-2112.2015.09.009

## Moving Target Detection in Natural Scene Based on Sparse Representation of Residuals

JIANG Jian-guo, JIN Yu-long, QI Mei-bin, ZHAN Shu

(Hefei University of Technology, Hefei, Anhui 230009, China)

**Abstract:** The paper proposes a non-parametric moving target detection algorithm based on sparse representation residuals error. In order to achieve precise motion target detection, we assume that the foreground change can be seen as sparse residuals compared with the static background. First of all, we use first  $n$  frames of the video to initialize the sparse representation dictionary. It will be applied to reconstruct the subsequent frame, extract frame residuals of every image, and then extract binary foreground images combining with the pixel-based global threshold value matrix. Furthermore, we remove ghost area on the basis of the foreground and edge regions. Finally, using the incremental PCA (Principal Component Analysis) and the idea of keep and update, we renew the above background model. A set of experiments are conducted on the shadow sets of *changedetection.net* using global update and residual error calculation method, and the result shows that the algorithm is an effective and efficient way to adapt to changes in the shadow of a static scene because of the changes of light. What is more, as to the small amplitude changes of the static scene and camera shake problems, it can also be a good solution.

**Key words:** background modeling; residual error; moving target detection; sparse representation

## 1 引言

监控视频处理中, 运动目标检测是基础性的工作, 其准确性直接影响后续处理的可行性. 在静态场景监控视频中, 运动目标检测主要用于运动目标提取和检测场景中物体的变化. 该领域内, 背景减法是最常用的一种方法<sup>[1~4]</sup>, 其核心思想是通过通过对每个像素位置构建一个背景像素值模型, 利用当前像素值与模型比较来提取运动目标. 对于背景减法的研究有时不仅仅局限于核心算法, 研究人员还常常把注意力放在处理 ghost 区域<sup>[2]</sup>、阴影<sup>[5]</sup>、遮挡和相机抖动等方面, 这些处理需要添加一些

预处理或者后处理, 可能会造成如边缘模糊和过分填充等现象, 使得检测的精度下降. 现有的大部分研究方法<sup>[1,2,4,5]</sup>都是基于像素点的方法, 这些方法可以很好的解决大面积前景和相机剧烈抖动问题, 但算法对全局光照变化很敏感, 特别是在光照发生突变时, 检测的准确度会严重下降. 文献[6,7]采用全局更新的字典学习方法进行运动目标检测, 可有效解决光照敏感问题, 其基本思想是基于场景的不同状态构建不同的字典进行运动目标检测, 但构建多字典时存在计算量大、延迟长等缺点, 并要求已知场景的可能状态及对应帧作为训练集, 在自然场景状态不确定情况下很难实现.

本文提出一种非参数化的基于稀疏表达残差的运动目标检测算法.通过稀疏表达残差进行运动目标检测,对背景的光照变化及其引起的阴影鲁棒性更强;利用对残差的滤波和全局自适应阈值解决小面积的动态背景和相机的小幅抖动问题.此外,本文充分利用了边缘点与前景区域之间的关联性,当某静态前景区域没有边缘点时视为 ghost 区域,使用将前景图像与更新图像分离的方法,在有效地剔除 ghost 区域的前提下,保证了前景的细节部分.

## 2 稀疏表达

稀疏表达在模式识别和机器视觉中被广泛的使用,如:图像修复<sup>[8]</sup>、人脸识别<sup>[9]</sup>和超分辨率重建等.在文献[10]中作者将跟踪问题映射成最相似区域的搜索问题,利用稀疏表达加琐碎模板的方式进行相似区域搜索,目标跟踪阶段根据这部分残差来判断目标的位置:

$$\mathbf{y} = \mathbf{D}\mathbf{z} + \mathbf{e} = [\mathbf{D} \quad \mathbf{I}] \begin{bmatrix} \mathbf{z} \\ \mathbf{e} \end{bmatrix} \quad (1)$$

其中,  $\mathbf{y}$  表示观测目标,  $\mathbf{D}$  表示通过学习构建的字典,  $\mathbf{z}$  表示目标  $\mathbf{y}$  在字典  $\mathbf{D}$  下面的重构系数矩阵,  $\mathbf{e}$  表示重构残差,  $\mathbf{I}$  表示对应残差的琐碎模板部分.

文献[10]中假设每个候选的区域可以使用字典和琐碎模板重构,且重构残差  $\mathbf{e}$  是稀疏的,这种情况下,式(1)可以通过式(2)来求解,

$$\min_{\mathbf{z}, \mathbf{e}} \frac{1}{2} \|\mathbf{y} - \mathbf{D}\mathbf{z} - \mathbf{e}\|_2^2 + \lambda \|\mathbf{e}\|_1 \quad (2)$$

其中  $\|\cdot\|_1$  和  $\|\cdot\|_2$  分别表示 1 范数和 2 范数,  $\lambda$  是一个常量,在本文中取值 0.05.

但是在实际的跟踪问题中,跟踪的字典是由目标区域构建的,因此,当目标外观变化较大时,目标区域会发生很大变化,从而造成残差  $\mathbf{e}$  非稀疏,这将使残差稀疏的假设不成立.但是在静态场景中,背景整体不变,只有目标所在的小部分会发生变化,可以满足残差稀疏的假设.本文从目标区域相对静态场景可以视为稀疏残差的角度出发,利用稀疏表达残差做运动目标检测.

## 3 基于残差的运动目标检测

### 3.1 目标检测模型

字典的构建是稀疏表达的一个关键问题,其构建方式有多种,可以直接使用所检测到的目标作为字典,但这样做会造成字典不断增大,存在很多的冗余信息;本文借鉴文献[10]的方法,采用对场景进行 PCA 提取主分量作为稀疏表达的字典.对于静态场景而言,PCA 可以将作为主分量的背景提取出来,与背景建模中提

取背景的思想相吻合.

本文将视频序列的前 20 帧作为初始化帧,用 PCA 提取主分量来初始化稀疏表达的字典  $\mathbf{D}$ .为了简化模型的求解,我们采用 PCA 主分量的前 2 个分量构建稀疏表达的字典,采用式(3)固定迭代求解的方法,可快速有效的求解出残差.从式(3)中可以看出,当残差  $\mathbf{e}$  稀疏时,对系数  $\mathbf{z}$  的影响很小,所以本文采用两次迭代的方法求解稀疏残差  $\mathbf{e}$ ,  $\mathbf{e}$  的初始值为  $\mathbf{0}$ .

$$\begin{aligned} \mathbf{z} &= \mathbf{D}^{-1} * (\mathbf{y} - \mathbf{e}) \\ \mathbf{e} &= \mathbf{S}_\lambda(\mathbf{y} - \mathbf{D}\mathbf{z}) \\ \mathbf{S}_\tau(\mathbf{x}) &= \text{sgn}(\mathbf{x}) * (|\mathbf{x}| - \tau) \end{aligned} \quad (3)$$

其中,  $\text{sgn}(\mathbf{x})$  为符号函数.我们利用式(3)对后续帧  $\mathbf{y}$  求解残差  $\mathbf{e}$ , 残差  $\mathbf{e}$  中包含运动目标的信息和噪声信息,添加滤波可以去除噪声的干扰.

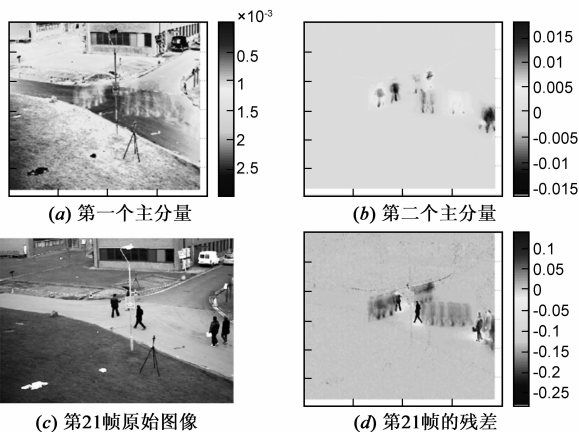


图1 利用PCA提取的静态场景的主分量及其表示的残差

图1(a)是通过静态场景的前 20 帧提取的第一主分量,体现出了静态场景的轮廓,包含了部分前景阴影;图1(b)是场景进行 PCA 提取的第二主分量,这部分及后面的分量体现出来的是前 20 帧中运动目标的信息;为了有效的抵消建立背景时第一主分量引入前景的干扰,采用前两个主分量作为稀疏表达的字典;图1(c)是第 21 帧的原图,图1(d)是利用字典重构第 21 帧场景的残差,从图示可以看出残差是稀疏的,且目标所在的区域残差值明显偏大,便于前景提取.

对图1(d)进行中值滤波滤除离散的前景噪声点,然后采用阈值分割对残差进行二值化.在二值化时,因为图像的不同区域光照强度不同,导致图像上每个像素点的动态范围不一样,采用全局固定阈值会丢掉某些前景点,因此本文采用自适应的阈值图像来对残差图像进行二值化,提取前景.将残差图像  $\mathbf{e}$  取绝对值后映射到区间  $[0 \sim 255]$  之间,然后对残差  $\mathbf{e}$  用改进的大津阈值(otsu)算法<sup>[11]</sup>提取一个自适应全局阈值 level,利用式(4)构建一个阈值图像  $\mathbf{M}_{th}$ ,

$$M_{th} = (1 - |F_N - 128| / 256) * level \quad (4)$$

其中:  $M_{th}$  是阈值图像,  $F_N$  是当前帧图像; 映射的目的是在像素值 128 这个中间值左右图像的动态范围最大, 可以使用最大的阈值, 像素值向 0 (或者 255) 移动时可能是因为光线不足 (或者光线太强) 导致像素的动态范围变小, 所以相对的阈值减小一些. 最后利用  $M_{th}$  对残差图像进行二值化提取前景.

### 3.2 模型更新

外在环境的光照变化和背景物体的改变等不确定因素, 使得仅初始化一次的背景字典不能适应实时变化的自然场景, 本文采用增量 PCA<sup>[12]</sup> 的方式对字典进行更新. 采用增量 PCA 更新的目的是: (1) 利用后续帧的信息, 使背景适应场景的实时变化, 减少初始化的影响, 增加了算法的鲁棒性; (2) 仅需要新样本和原始的基向量, 减小了存储空间和计算量.

更新时采用保守更新法, 检测到的前景一直被视为目标, 不更新到背景中, 所以在更新时剔除前景点更新. 我们将检测到的前景图像  $T$  与需要更新的图像  $U$  分开, 定义如式(5), 对于  $U$  的更新在下一节阐述:

$$T = \begin{cases} 1, & \text{前景} \\ 0, & \text{背景} \end{cases} \quad (5)$$

$$U = \begin{cases} 1, & \text{不更新} \\ 0, & \text{更新} \end{cases}$$

在字典  $D$  更新之前, 我们利用式(6)将当前帧  $F_N$  里面的目标区域每个像素  $x_i$  用背景替换, 去掉当前帧的前景部分, 然后再利用式(7)来更新字典.

$$F_N'(x_i) = \begin{cases} F_N(x_i), & \text{if } U(x_i) = 0 \\ D(x_i) * z, & \text{else} \end{cases} \quad (6)$$

$$D = \text{ICPA}(\alpha * D, F_N') \quad (7)$$

其中,  $F_N'$  是去掉前景后的图像, ICPA 表示增量 PCA 函数,  $\alpha$  表示衰减系数. 图 2 是去掉目标前后的图像. 我们利用式(7)对字典  $D$  进行更新, 更新时对原始的字典进行衰减, 衰减系数  $\alpha$  决定了背景的更新速度 (本文固定其为 0.97), 当实际背景变化慢时调节该系数可以使背景的更新速度和实际背景变化速度相同步.



图2  $F_N$  与  $F'_N$

### 3.3 目标检测的 ghost 现象校正

背景的更新存在更新速率问题, 如果更新过快, 会造成运动速度慢的目标丢失; 更新速度过慢时, 当静止

目标突然运动, 会出现场景中没有目标却一直检测出目标的 ghost 现象, 如图 3 所示, 一些区域原本不存在目标却检测出目标.



图3 静态场景下运动目标检测 Ghost 现象

ghost 区域是由于长期逗留被更新到背景中的目标突然运动, 或者背景中某个静止目标突然运动而形成的; 在这种情况下, 采用保守更新无法去除 ghost 区域. ghost 区域具有两个特征: (1) ghost 区域相对于背景而言是静止不动的; (2) 目标之所以能从背景中提取出来是因为其与背景存在亮度上的差异, 所以对于灰度图像而言, 真正的目标肯定是存在边缘的, 而且与检测到的目标的边缘是重合的. 我们结合上面两点做出以下的假设来去除 ghost 区域, 当某一静止前景区域边缘点数目满足式(8)条件时被设定为前景:

$$N > L/4 \quad (8)$$

式中,  $N$  表示检测到的某一静止前景区域对应的灰度图像中边缘检测的边缘点数目,  $L$  表示该区域二值图像边界长度; 式(8)是利用区域的边缘点数目来判断当前的静态区域是否为目标, 为了确保不误删除前景目标, 本文将边缘点最小数目设定为区域周长的 0.25 倍, 也就是认为如果静止前景是矩形, 至少有一边是有边缘的, 否则该区域被认为是 ghost 区域.

对于 ghost 区域, 不直接在前景图像  $T$  图像中修改前景, 而是修改更新图像, 使其中 ghost 区域更新到背景中, 进而避免盲目的修改前景, 造成前景模糊. 更新图像  $U$  生成方式如式(9)所示:

$$U = T \quad \text{初始化} \quad (9)$$

$$U(a_i) = \begin{cases} 1, & N > L/4 \\ 0, & N \leq L/4 \end{cases} \quad \text{校正}$$

其中,  $a_i$  表示前景中第  $i$  个静止区域,  $N$  和  $L$  分别表示区域  $a_i$  的边缘点数目和前景区域的边长,  $M$  是前景中静止区域的总数目. 根据式(6)在当前帧中去掉前景区域, 更新字典  $D$  时可以对旧的字典增大衰减程度, 使字典中的 ghost 区域尽快消失.

## 4 算法实现步骤

综上所述, 本文算法具体实现步骤如下:

**步骤 1** 对前 20 帧图像做 PCA 变换提取主分量初始字典  $D$ , 初始化字典时选取 2 帧以上即可完成初

始化工作;

**步骤 2** 利用字典  $D$  和当前帧根据式(3)得到残差  $e$ ;将残差映射到区间 $[0 \sim 255]$ 之间,采用 otsu 算法计算出全局阈值 level,利用公式(4)将全局阈值映射到阈值图像  $M_{th}$ ,然后对残差  $e$  进行二值化,得到前景目标  $T$ ;

**步骤 3** 根据前景目标  $T$  和当前帧图像  $F_N$ ,利用式(9)生成  $U$ ,通过式(6)生成更新基底  $D$  的图像  $F'_N$ ;

**步骤 4** 利用增量 PCA 的和图像  $F'_N$  更新字典  $D$ ,回到步骤 2.

## 5 实验及结果分析

本文使用的测试数据集来自网站 <http://www.changedetection.net><sup>[13]</sup>,该网站提供了 31 个视频分为 6 大类,本实验主要针对其中 Shadow 类的 6 个视频.Shadow 类视频主要是由于光照变化产生阴影变化的视频集,平均长度为 3000 帧,每个视频都包含光照等因素变化引起的阴影变化;其中部分视频包含前景目标的间歇性运动,如 VIBE 等算法采用的盲目更新策略会将人物更新到背景中使目标丢失,本文算法会因为人物存在边缘而对间歇性目标不更新,进而保留间歇性运动目标.

实验 1 中,本文算法采用前 20 帧作为初始字典的帧数目.参考网站的评测方法,将运动目标检测看成一个二分类问题来对算法进行评测<sup>[13]</sup>.其中前景对应二分类里面的正样本(positive),背景对应二分类里面的负样本(negative),使用二分类术语来描述检测出来的各个区域:TP 表示正确的前景点数目,TN 表示正确的背景点数目,FP 表示错误的前景点数,FN 表示错误的

背景点数目,在后续的算法评价将以错误点所占比例(PBC)、前景分割精度(Precision)和前景点准确率(Recall)来描述.

$$PBC = 100 \times \frac{FN + FP}{TN + TP + FN + FP} \quad (10)$$

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

$$Recall = \frac{TP}{TP + FN} \quad (12)$$

本文选择与 PBAS<sup>[4]</sup>、VIBE +<sup>[2]</sup>、PSP-MRF<sup>[14]</sup>、SC-SOBS<sup>[15]</sup>、VIBE<sup>[1]</sup>和经典算法 GMM<sup>[16]</sup>等六种方法比较,其中前四种方法是文献[13]的评测综合排名前四的算法,第六种算法是经典的背景建模算法,VIBE<sup>[1]</sup>是当前比较实用的背景建模算法.这六种方法都是基于像素点更新背景的方法.

PBC 表示的是目标检测中错误点所占的比例,这个值越小表示分割的准确性越高,通过表 1 可以看出本文算法分割的 PBC 值在六个视频序列中 4 个是最小,表明本文算法在做运动目标检测时检测错误的像素点少.Precision 表示在检测到的前景点中,正确前景点占总的前景点的比例,这个值越高说明检测出来的前景越准确,通过表 2 可以看出本文算法在六个视频中有五个是排第一的,而且每一个都是 90% 以上,说明检测的前景点的准确率很高.这两个指标说明本文的目标检测算法可以在低前景点检测错误率的前提下检测出高精度的前景目标,证明了本文方法采用的全局背景更新和 ghost 区域剔除的有效性,可以有效的解决光照变化带来的影响.

表 1 错误点所占比例(PBC)

	PBAS	VIBE +	PSP-MRF	SC-SOBS	VIBE	GMM	本文算法
Shadow							
backdoor	0.551491	0.556625	0.609977	0.679521	0.579076	1.92742	0.458769
bungalows	2.24361	2.80029	2.55377	2.80993	3.1974	2.83581	1.083129
busStation	1.05025	1.40227	1.16558	1.17755	1.41705	1.41559	1.396179
copyMachine	1.79759	3.29544	6.81899	6.71429	2.43679	3.91617	1.575413
cubicle	0.718704	0.724207	1.22272	1.31517	0.998293	1.67761	1.182064
peopleInShade	1.29014	1.13055	1.07751	1.10328	1.29977	1.40144	0.179694
均值	1.2753	1.65157	2.24142	2.29996	1.65473	2.19567	0.979208

表 2 检测精度(Precision)

	PBAS	VIBE +	PSP-MRF	SC-SOBS	VIBE	GMM	本文算法
Shadow							
backdoor	0.795689	0.873566	0.797969	0.78754	0.902027	0.509392	0.930222
bungalows	0.746759	0.71674	0.727608	0.718189	0.70258	0.715857	0.997740
busStation	0.842177	0.877513	0.862112	0.856564	0.914483	0.883183	0.947250
copyMachine	0.856445	0.812879	0.505908	0.512186	0.885148	0.838412	0.940927
cubicle	0.798722	0.858931	0.630238	0.61628	0.756284	0.551111	0.693471
peopleInShade	0.845992	0.841779	0.844817	0.846973	0.844546	0.841061	0.999136
均值	0.814297	0.830235	0.728109	0.722955	0.834178	0.723169	0.918124

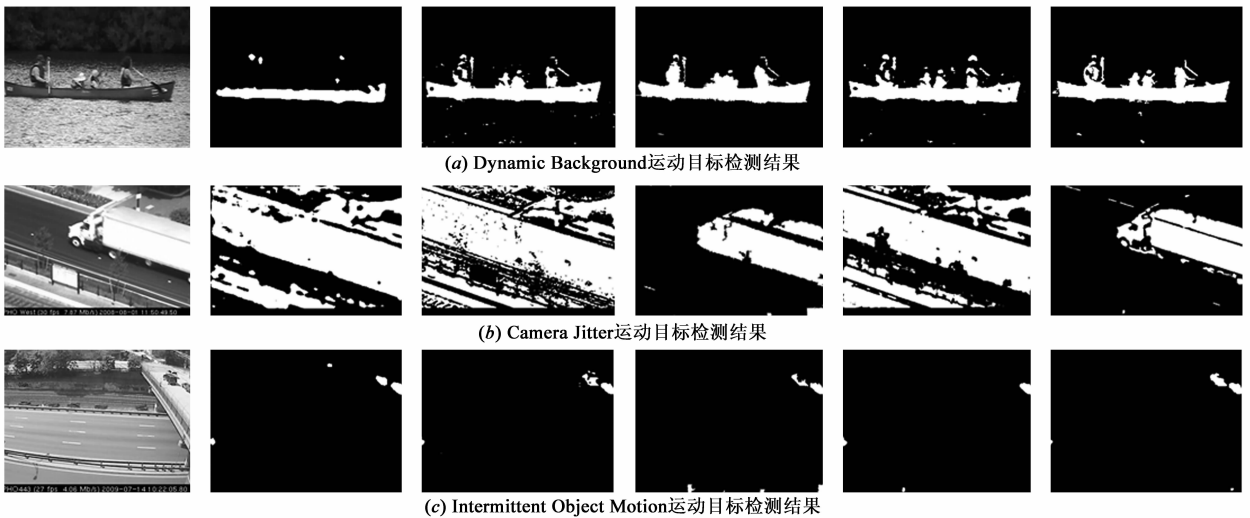


图4 部分算法分割结果(从左至右:原图、PBAS、VIBE、VIBE+、GMM、本文算法)

实验 2 是 PBAS<sup>[4]</sup>算法、VIBE<sup>[1]</sup>算法、VIBE + <sup>[2]</sup>算法、GMM 算法<sup>[16]</sup>和本文算法在 Dynamic Background 数据集、Camera Jitter 和 Intermittent Object Motion 数据集(视频集来源: <http://www.changedetection.net><sup>[13]</sup>)上做的结果比较,检测结果如图 4 所示.图 4(a)是四种算法在具有动态背景(水波纹)的场景对运动目标的检测结果,从结果可以看出本文算法可以准确的检测出前景,并且检测的动态背景很少,同时没有过分的填充;图 4(b)是在具有相机抖动的监控视频上做运动目标检测的结果,从结果可以看出,本文算法对相机的抖动也具有一定的抗干扰能力,并且检测的细节明显;图 4(c)是对间歇性运动目标的检测结果,本文算法也同样具有较好的检测结果,但是也丢失了一部分小的目标信息.

通过该实验可以看出,本文算法对具有动态背景、轻微相机抖动和目标间歇性运动的监控视频都有很好的检测结果,而且没有由于添加预处理和后处理而导致的目标过分填充和目标细节丢失等问题.

实验 3 是与文献[6,7]的方法作比较的实验,文献[6]中采用多字典的方式来建模场景的不同状态,但是

在自然场景没有固定的状态时,检测的准确率就受到影响,文献[7]将鲁棒的字典学习用在运动目标检测,采用的方案与文献[6]一样.表 3 是本文算法的实验结果与文献[7]方法在同一数据集上面实验结果的比较;图 5 是本文算法与文献[6,7]在数据集 RainCar、OceanPerson 和 WaterObject 三个视频上实验结果的比较.视频 RainCar 涉及到雨幕的干扰和前景的突然变化,对于按点更新的背景模型,模型对自然场景的全局变化适应速度会比较慢;但对于基于字典的运动目标检测,都没有这个问题.视频 OceanPerson 和 WaterObject 都存在视频的局部动态的水纹背景.

从结果可以看出几种基于稀疏表达的算法都可以很好的抵抗背景的动态变化,抗背景干扰能力强;同时,从三个视频的测试结果可以看出,本文算法无论是在检测的准确度,还是检测的误差都明显的比文献[6,7]方法好,从而说明本文算法在模型复杂度低于文献[6,7]的情况下,对于没有固定状态的自然场景可以达到更高的检测精度.

表 3 在文献[7]数据集上 Precision 与 Recall 对比

	Ocean man		Water Object		Rain Car	
	Precision	Recall	Precision	Recall	Precision	Recall
ORDL <sup>[7]</sup>	0.8141	0.8148	0.6752	0.7548	0.9236	0.7943
BRDL <sup>[7]</sup>	0.7552	0.7982	0.7009	0.7225	0.9489	0.7920
Mairal et al. <sup>[17]</sup>	0.6367	0.8518	0.6312	0.6149	0.7940	0.8010
本文算法	0.9506	0.9490	0.9953	0.7874	0.9633	0.9135

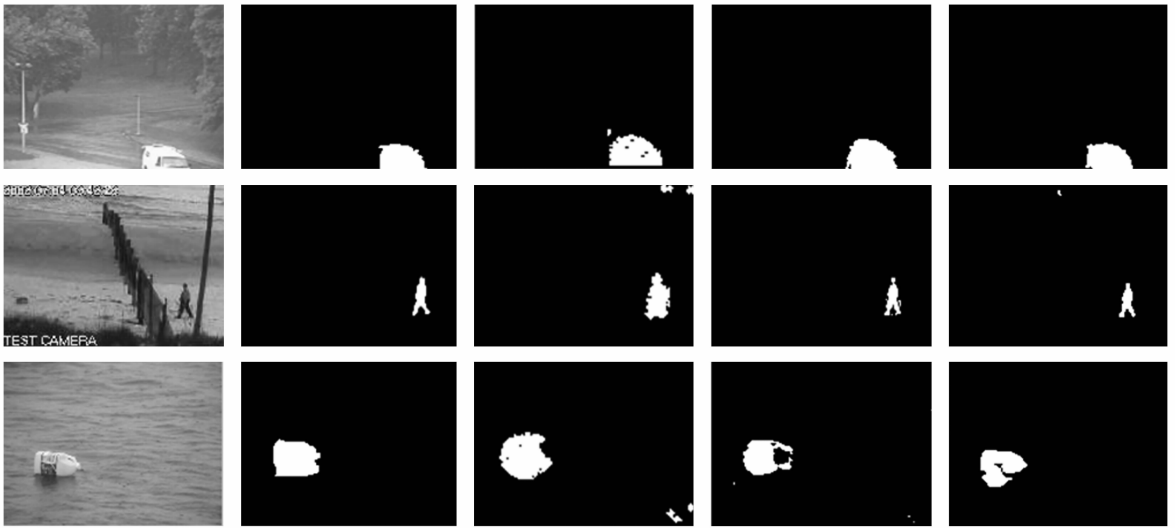


图5 本文算法与文献[6]、文献[7]算法结果比较(从左到右:原图、掩膜、文献[6]结果,文献[7]结果、本文结果)

## 6 总结与展望

本文采用基于残差的运动目标检测算法,是基于全局像素同时更新的背景建模算法,获得的残差是全局最优的,全局的光线变化对结果的影响小;另外采用了基于边缘的 ghost 区域去除方法可以很好的剔除 ghost 区域,使错误目标被提取的可能性大大下降,提高了静态场景下运动目标提取的准确率.并且本文的方法是一个非参数化方法,这一点对实际应用很有意义.本文算法对目标间歇性运动、动态背景和相机的抖动同样具有一定的抗干扰能力;但是在一些视频中,因为视频里面的变化区域过大,导致全局的参数不能很好地适应,本文算法在这些视频的目标检测结果不理想,后续的试验中需要进一步完善.

### 参考文献

- [1] Barnich O, Van Droogenbroeck M. Vibe: a powerful random technique to estimate the background in video sequences[A]. IEEE International Conference on Acoustics, Speech and Signal Processing[C]. Taipei: IEEE, 2009. 945 - 948.
- [2] Van Droogenbroeck M, Paquot O. Background subtraction: Experiments and Improvements for ViBe[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops[C]. Providence, RI: IEEE, 2012. 32 - 37.
- [3] Bouwmans T, El Baf F, Vachon B. Statistical background modeling for foreground detection: A survey[J]. Handbook of Pattern Recognition and Computer Vision, 2010, 4(3): 181 - 199.
- [4] Hofmann M, Tiefenbacher P, Rigoll G. Background segmentation with feedback: The pxel-based adaptive segmenter[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops[C]. Providence, RI: IEEE, 2012. 38 - 43.
- [5] Morde A, Ma X, Guler S. Learning a background model for change detection[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops[C]. Providence, RI: IEEE, 2012. 15 - 20.
- [6] Zhao C, Wang X, Cham W-K. Background subtraction via robust dictionary learning[J]. EURASIP Journal on Image and Video Processing, 2011, 2011(2): 1 - 12.
- [7] Cewu Lu, Jianping Shi, Jiaya Jia. Online robust dictionary learning[A]. IEEE Conference on Computer Vision and Pattern Recognition[C]. Portland, OR: IEEE, 2013. 415 - 422.
- [8] Wright J, Ma Y, Maral J, et al. Sparse representation for computer vision and pattern recognition[J]. Proceedings of the IEEE, 2010, 98(6): 1031 - 1044.
- [9] 胡正平, 李静. 基于低秩子空间恢复的联合稀疏表示人脸识别算法[J]. 电子学报, 2013, 41(5): 987 - 991.  
HU Zheng-ping, LI Jing. Face recognition of joint sparse representation based on low-rank subspace recovery[J]. Acta Electronica Sinica, 2013, 41(5): 987 - 991. (in Chinese)
- [10] Wang D, Lu H, Yang M-H. Online object tracking with sparse prototypes[J]. IEEE Transactions on Image Processing, 2013, 22(1): 314 - 325.
- [11] 何志勇, 孙立宁, 陈立国. Otsu 准则下分割阈值的快速计算[J]. 电子学报, 2013, 41(2): 267 - 272.  
HE Zhi-yong, SUN Li-ning, CHEN Li-guo. Fast computation of threshold based otsu criterion[J]. Acta Electronica Sinica, 2013, 41(2): 267 - 272. (in Chinese)
- [12] Ross D, 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.
- [13] Goyette N, Jodoin, P-M Porikli F, et al. Changedetecti-on: A new change detection benchmark dataset[A]. IEEE

Computer Society Conference on Computer Vision and Pattern Recognition Workshops[C]. Providence, RI: IEEE, 2012. 1 – 8.

- [14] Schick A, Bäuml M, Stiefelhagen R. Improving foreground segmentations with probabilistic superpixel markov random fields[A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops[C]. Providence, RI: IEEE, 2012. 27 – 31.
- [15] Maddalena L, Petrosino A. The SOBS algorithm: What are the limits? [A]. IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops [C]. Providence, RI: IEEE, 2012. 21 – 26.
- [16] Zivkovic Z. Improved adaptive Gaussian mixture model for background subtraction[A]. Proceedings of the IEEE International Conference on Pattern Recognition [C]. Cambridge: IEEE, 2004. 28 – 31.
- [17] Mairal J, Bach F, Ponce J, Sapiro G. Online learning for matrix factorization and sparse coding[J]. The Journal of Machine Learning Research, 2010, 11(3): 19 – 60.

## 作者简介



**蒋建国** 男, 1955 年出生, 教授、博士生导师, 合肥工业大学计算机与信息学院. 主要研究方向: 视频信息分析与处理、多媒体智能监控系统、分布式智能系统、DSP 技术应用.

E-mail: jgjiang@hfut.edu.cn



**金玉龙** 男, 1989 年出生, 硕士研究生, 合肥工业大学计算机与信息学院. 主要研究方向: 数字图像分析与处理.

E-mail: jy119890119@163.com