

Journal of Image
and Graphics

中国图象图形学报



ISSN1006-8961
CN11-3758/TB

2012 **7**
Vol.17 No.

中国科学院遥感应用研究所
中国图象图形学学会主办
北京应用物理与计算数学研究所

中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2012年7月 第17卷 第7期(总第195期)

目次

综述

中国多媒体技术研究: 2011 史元春, 徐光祐, 高原(741)

图像处理和编码

图像引导滤波的局部多尺度 Retinex 算法 方帅, 杨静荣, 曹洋, 武鹏飞, 饶瑞中(748)

基于第一主成分方向稳定性的图像零水印 付剑晶, 王珂(756)

小波变换估计非线性扩散最优停止时间 蒋平, 张建州(770)

图像分析和识别

改进的 Beamlet 与 Canny 相结合提取复杂图像线特征 曾接贤, 周沥沥, 符祥(775)

动静态信息融合及动态贝叶斯网络的步态识别 杨旗, 薛定宇(783)

融合音频单词与视觉特征的成人视频检测 刘毅志, 唐胜, 王向东, 林守勋, 张勇东(791)

基于平面区域跟踪的目标位姿参数自动测量 回丙伟, 文贡坚, 赵竹新, 钟金荣(798)

3 维图像中边界曲面的分类追踪及抽取 丁德福, 程柳航, 王利生(806)

局部时空域模型的核密度估计目标检测方法 王兴宝, 刘纯平, 费兰英, 王朝晖, 季怡(813)

基于相位谱和调谐幅度谱的显著性检测方法 李崇飞, 高颖慧, 卢凯, 曲智国(821)

图像理解和计算机视觉

“目标-场景”语境关联的生成图模型分析 谢昭, 李姍琦, 高隼(828)

融合上下文信息的场景结构恢复 武晖, 于昕, 隋尧, 张利(839)

带有局部控制因子的图割光流估计 路子赟, 唐土生, 高隼, 沈琳, 刘伟(846)

改进的 Harris 亚像素角点快速定位 何海清, 黄声享(853)

计算机图形学

应用最小生成树实现点云分割 孙金虎, 周来水, 安鲁陵(858)

医学图像处理

病毒进化的离散差分进化超声图像特征选择 张巧荣, 朱长明, 倪军, 刘海波(866)

分水岭优化的 Snake 模型肝脏图像分割 兰红, 张璐(873)

遥感图像处理

Normalized Cut 与分水岭变换在高光谱影像混合像元端元提取中的应用 许菡, 李小娟(880)

改进非局部均值滤波的 SAR 图像降噪方法 郑永恒, 程建, 曹宗杰(886)

~~~~~  
第四届国际遥感考古会议 ..... (892)

**中国图象图形学报**

刊名题字: 宋 健

月刊(1996 年创刊)

第 17 卷 第 7 期

2012 年 7 月 16 日出版

**主管单位** 中国科学院  
**主 办** 中国科学院遥感应用研究所  
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  
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  
**印刷装订** 北京北林印刷厂  
**广告经营许可证** 京朝工商广字第 0346 号  
**总 发 行** 北京报刊发行局  
**订 购** 全国各地邮局  
**国外发行** 中国国际图书贸易总公司  
(中国国际书店)  
(北京 399 信箱 邮编 100044)

**Superintended by** Chinese Academy of Sciences  
**Sponsored by** Institute of Remote Sensing Application,  
CAS China Society of Image and Graphics  
Institute of Applied Physics and Computational  
Mathematics  
**Chief editor** LI Xiaowen  
**Editor, Publisher** Editorial and Publishing Board  
of Journal of Image and Graphics  
(P. O. Box 9718, Beijing 100101, China)  
E-mail:jig@irsa.ac.cn  
**Distributed by** Beijing Bureau for Distribution of Newspapers  
and Journals  
**Domestic** All Local Post Offices in China  
**Foreign** China International Book Trading Corporation  
(P. O. Box 399, Beijing 100044, China)  
**Printed by** Beijing Beilin Printing House

ISSN 1006-8961 CN11-3758/TB CODE ZTTFXZ 国内邮发代号: 82-831 国外发行代号: M1406 国内定价: 45.00 元

# Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 17 No. 7 July 2012

## Contents

### Review

Researches on multimedia technology in China, 2011 ..... Shi Yuanchun, Xu Guangyou, Gao Yuan (741)

### Image Processing and Coding

Local multi-scale Retinex algorithm based on guided image filtering  
..... Fang Shuai, Yang Jingrong, Cao Yang, Wu Pengfei, Rao Ruizhong (748)

Image zero-watermark based on direction stability of first principal component vector ..... Fu Jianjing, Wang Ke (756)

Stopping-time estimation for anisotropic diffusion using discrete wavelet transform ..... Jiang Ping, Zhang Jianzhou (770)

### Image Analysis and Recognition

Complex image line feature extraction based on improved Beamlet transform and the Canny operator  
..... Zeng Jiexian, Zhou Lili, Fu Xiang (775)

Gait recognition based on dynamic & static information fusion and dynamic bayesian network ..... Yang Qi, Xue Dingyu (783)

Fusing audio-words with visual features for adult video detection  
..... Liu Yizhi, Tang Sheng, Wang Xiangdong, Lin Shouxun, Zhang Yongdong (791)

Automatic measurement for an object's position and attitude via tracking planar regions  
..... Hui Bingwei, Wen Gongjian, Zhao Zhuxin, Zhong Jinrong (798)

Detection and extraction of boundary surface patches within 3D images ..... Ding Defu, Cheng Liuhang, Wang Lisheng (806)

Foreground object detection method using kernel density estimation of a local spatio-temporal model  
..... Wang Xingbao, Liu Chunping, Fei Lanying, Wang Zhaohui, Ji Yi (813)

Saliency detection method based on phase spectrum and amplitude spectrum tuning  
..... Li Chongfei, Gao Yinghui, Lu Kai, Qu Zhiguo (821)

### Image Understanding and Computer Vision

"Object-Scene" contextual associated generative graph model analysis ..... Xie Zhao, Li Shanqi, Gao Jun (828)

Structure recovery algorithm using contextual information ..... Wu Hui, Yu Xin, Sui Yao, Zhang Li (839)

Graph cut optical flow estimation with a local control factor ..... Lu Ziyun, Tang Tusheng, Gao Jun, Shen Lin, Liu Wei (846)

Improved algorithm for Harris rapid sub-pixel corners detection ..... He Haiqing, Huang Shengxiang (853)

### Computer Graphics

Research on point cloud segmentation using a minimum spanning tree ..... Sun Jinhu, Zhou Laishui, An Luling (858)

### Medical Image Processing

Virus-evolutionary discrete differential evolution algorithm for feature selection of cervical lymph nodes in ultrasound images  
..... Zhang Qiaorong, Zhu Changming, Ni Jun, Liu Haibo (866)

Liver image segmentation algorithm based on the Snake model And optimized by watershed transformation  
..... Lan Hong, Zhang Lu (873)

### Remote Sensing Image Processing

Endmember extraction for hyperspectral image based on normalized cut and watershed transformation  
..... Xu Han, Li Xiaojuan (880)

SAR image denoising via improved non-local means filter ..... Zheng Yongheng, Cheng Jian, Cao Zongjie (886)

中图法分类号: TP391 文献标识码: A 文章编号: 1006-8961(2012)07-0813-08

论文引用格式: 王兴宝, 刘纯平, 费兰英, 王朝晖, 季怡. 局部时空域模型的核密度估计目标检测方法[J]. 中国图象图形学报, 2012, 17(7): 813-820.

## 局部时空域模型的核密度估计目标检测方法

王兴宝, 刘纯平, 费兰英, 王朝晖, 季怡

苏州大学计算机科学与技术学院, 苏州 215006

**摘要:** 针对非参数核密度估计在前期学习阶段信息冗余和计算量大, 在后期背景更新阶段自适应性差需手动调整阈值和检测结果出现阴影等问题, 提出一种基于局部时空域模型的核密度估计目标检测方法。在前期训练学习阶段采用 K 均值聚类选择关键帧, 从而避免信息冗余和计算量大问题; 在后期背景更新阶段, 构建一种局部时空域模型, 在时间域通过历史帧信息自适应调整时间域窗口大小, 在空间域利用颜色和 LBP 描述的纹理特征消除部分阴影问题。在复杂场景下的实验结果表明, 该算法在实时性和检测准确率方面有效得到提高。

**关键词:** 核密度估计; 局部时空域模型; K 均值; LBP 算子

### Foreground object detection method using kernel density estimation of a local spatio-temporal model

Wang Xingbao, Liu Chunping, Fei Lanying, Wang Zhaohui, Ji Yi

Department of Computer Science and Technology, Soochow University, Suzhou 215006, China

**Abstract:** In this paper, we propose a new method for foreground object detection based on the Kernel: Density Estimation of a local spatio-temporal model (LST-KDE), which overcomes information redundancy and the large calculated quantity problem in the training phase as well as the manual adjusting time window size and shadow problem in the detection and updating background phase. The LST-KDE algorithm uses the k-means clustering algorithm to optimize the sample set and to choose the key frames in the training phase. Therefore, it can avoid information redundancy and the large calculated quantity problem. In the detection and updating background phase, the LST-KDE algorithm constructs a local spatio-temporal model. This method can not only adaptively set the time window size by using history frame information in a temporal model, but also uses color and texture features described with the local binary pattern (LBP) algorithm to remove shadows in the spatial model. The experiment in a complex environment demonstrates that the proposed method outperforms recent state-of-the-art methods.

**Key words:** kernel density estimation (KDE); local spatio-temporal pattern; K-means; LBP algorithm

## 0 引言

运动目标检测方法主要有帧间差分法<sup>[1]</sup>、光流

法<sup>[2]</sup>和背景差分法<sup>[3-12]</sup>。帧间差分法应用条件有限, 当目标停止则失效; 光流法计算量大, 实时性差; 背景差分法具有精度高、实时性和反应灵敏等优点, 在摄像机静止条件下, 是最常用的方法。但因背景受光照

收稿日期: 2011-07-06; 修回日期: 2012-02-20

基金项目: 国家自然科学基金项目(61170124, 61170020, 61070223); 江苏省自然科学基金项目(BK2009116); 江苏省科技支撑计划项目(BE2009048); 江苏省主校自然科学研究项目(09KJA520002); 苏州市应用基础研究计划(SYG201116)

第一作者简介: 王兴宝(1988—), 男, 现为苏州大学计算机科学与技术学院硕士研究生, 主要研究方向为图像处理与模式识别。E-mail: wangxingbao@163.com

通讯作者: 刘纯平, E-mail: cpliu@suda.edu.cn

不均、目标移动和摄像机颤动等外界因素影响,构建一个高效、准确和鲁棒的背景模型成为一个难点。国内外许多学者对背景差分做了大量研究。针对复杂多变的场景,Nuria 等人提出背景特征化建模的方法<sup>[3]</sup>,对一段时间内背景图数据进行主成分分析,然后利用主特征矢量张成的子空间描述背景,用当前图像与子空间的距离作为判别前景的依据,检测效果较好,但算法复杂度高实时性差。Stauffer 等人提出了混合高斯模型(GMM)<sup>[4]</sup>,建立参数化的混合高斯分布模型来表示每个像素的概率分布,通过 K 均值近似算法对模型进行自适应更新,一定程度上能够适应多峰分布的背景变化;Jodoin 和 Lee 等人对 GMM 做了改进<sup>[5-9]</sup>,提出利用最大似然估计在混合高斯模型中引入高斯分布个数的自动选择和自适应学习机制,但是 GMM 依赖高斯分布以及参数结构的先验假设,难以准确描述复杂的背景变化;针对 GMM 不足,Elgammal 等人提出基于核密度估计的非参数背景建模算法(KDE)<sup>[10]</sup>,不对背景概率模型做任何假设,直接利用先前获得的像素值估计当前像素为背景的概率,能够适应复杂像素分布密度,克服像素值在短时间内发生频繁变化的问题。针对 KDE 算法通过给定的假正率来选择阈值需要根据场景人工选择问题,徐东彬等人提出自适应前景、背景阈值选择<sup>[11]</sup>,以解决误报率和漏报率间的矛盾;周恩策等人针对 KDE 算法中采样样本包含噪声且非典型运动像素也进行核密度估计所造成计算量大,产生估计错误噪声等问题,提出了一种基于像素时间信息窗的核密度估计前景检测算法<sup>[12]</sup>,该方法针对每个像素维持一个信息窗,利用像素时间窗的信息划定阈值进行背景更新。但文献[10-12]仍存在一些不足:1)算法实时性较差,特别在前期背景训练学习阶段,存在信息冗余和计算量大问题;2)准确率不高和自适应性差,受前期训练时背景包含大量噪声及后期检测时光照、阴影、部分遮挡和背景目标移动的影响,背景不能够及时更新,导致准确率低,不同环境条件需手工调节阈值,自适应性差。

针对以上问题,提出一种基于局部时空域模型的核密度估计目标检测算法(LST-KDE)。该方法在前期背景训练学习阶段通过 K-means 聚类算法得到  $L$  个关键帧为新训练样本集,并选取核密度估计值最大的像素值为背景点,避免重复计算。在背景模型更新中利用历史帧和当前帧背景匹配度自动调整时间窗口,并利用颜色和纹理实时更新背景,消除部分阴影问题,提高检测准确率。通过室内外复杂

视频实验证明该算法在实时性、准确率和鲁棒性方面取得很好效果。

## 1 非参数核密度估计运动检测原理

设视频帧有  $M$  个像素点,每个像素点有  $N$  个背景样本,则时刻  $t$  帧中第  $i$  个像素为  $x(t)_i$ ,对应的第  $j$  个背景样本为  $x(t)_{i,j}$ ,则时刻  $t$  像素  $i$  的概率为

$$P(x(t)_i) = \frac{1}{N} \sum_{j=1}^N K(x(t)_i - x(t)_{i,j}) \quad (1)$$

式中, $K$  为核函数,假设  $K$  服从高斯分布。对于相互独立  $d$  维特征向量,则  $N$  个样本概率之和为

$$P(x(t)_i) = \frac{1}{N} \sum_{j=1}^N \prod_{m=1}^d \frac{1}{\sqrt{2\pi}\sigma_{i,m}} e^{-\frac{(x(t)_{(i,m)} - x(t)_{(i,m),j})^2}{2\sigma_{i,m}^2}} \quad (2)$$

式中, $x(t)_{(i,m)}$  为第  $i$  个像素的第  $m$  个特征值, $\sigma_{i,m}$  为第  $m$  个特征的核宽。

文中采用高斯函数为核函数,若核宽  $\sigma$  过小,则密度估计欠光滑,过大将导致密度估计过度光滑。根据文献[10,12-13],核宽选取为

1) 动态背景区域,若背景像素不满足高斯分布,采用相邻样本绝对差中位数  $m$  (MAD) 计算核宽<sup>[13]</sup>

$$\sigma = \frac{m}{0.68/\sqrt{2}} \quad (3)$$

2) 非动态背景区域,若背景像素满足高斯分布,采用样本方差  $\sigma^2$  来估计核函数核宽。

$$\sigma^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad (4)$$

将核宽代入式(2)可得时刻  $t$  像素  $i$  概率  $P(x(t)_i)$ 。若  $P(x(t)_i)$  满足式(5),则  $x(t)_i$  为前景。

$$P(x(t)_i) < T_f \quad (5)$$

式中, $T_f$  为整幅图像的全局阈值。

## 2 局部时空域模型的 KDE 目标检测

### 2.1 优化训练样本集及构建背景模型

在训练学习阶段,训练样本集  $X_{(h,w)} = \{x_1, x_2, \dots, x_N\}$  为视频序列中某像素点  $N$  个采样值,其中  $1 \leq h \leq H, 1 \leq w \leq W, H$  和  $W$  为图像高度和宽度。

#### 2.1.1 提取关键帧样本集

由于全样本集信息冗余导致后期构建背景模型重复计算的问题,本文采用 K-means<sup>[14]</sup> 聚类方法,从  $N$  个训练样本集中选  $L$  个具有代表性的关键帧

$y_1, y_2, \dots, y_L$ 。关键帧选取为

$$L = \frac{N}{g} \quad (6)$$

式中,  $L$  由样本集个数  $N$  和帧间隔数  $g$  决定, 若视频帧速率大,  $g$  值选取大一点, 若视频帧速率小,  $g$  值选取相对小一点。

### 2.1.2 构建背景模型

背景模型构建以  $Y_{(h,w)} = \{y_1, y_2, \dots, y_L\}$  为样本集。选样本集中核密度估计值最大的样本点, 构建最终背景模型  $B_c(x, y)$ , 步骤如下:

1) 计算样本点的核密度估计值

$$P(y_i) = \frac{1}{L} \sum_{j=1}^L K(y_j - u_i), i \in [1, L] \quad (7)$$

式中,  $y_j$  为新样本集中的任一个样本点, 均值  $u_i = y_i$ 。

2) 构建核密度估计值与样本点的映射关系

$$map(P(y_i), y_i), i \in [1, L] \quad (8)$$

3) 求核密度估计值最大的样本点

$$b_{(h,w)} = \max\{map(P(y_i), y_i)\}, i \in [1, L] \quad (9)$$

图 1 给出了初始训练样本为 50 帧,  $L$  和  $g$  参数构建的背景模型。从图 1 发现, 聚类数目越多, 构建背景模型越精确, 但比较耗时, 相反构建的背景模型就不太精确, 考虑精度和时间平衡, 本文取  $L = 10$ , 间隔  $g = 5$ 。

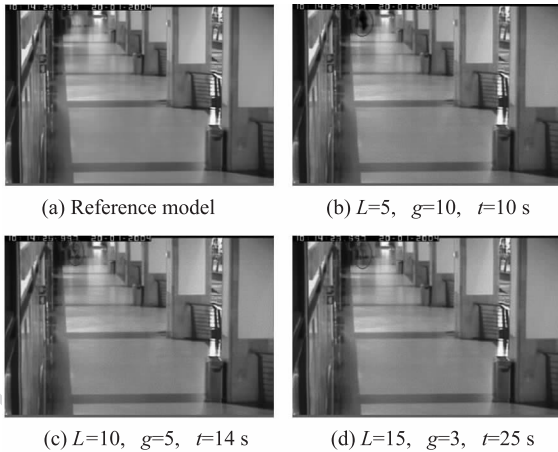


图 1 不同聚类数目  $L$  和间隔  $g$  所构建背景模型及消耗时间  $t$   
Fig. 1 Different clustering number ( $L$ ), time gap ( $g$ ) to construct background model and time-consuming ( $t$ )

## 2.2 局部时空域模型的建立

由于光照变化、树叶摇动或者背景目标移动等外部因素, 会影响运动目标检测效果, 鉴于此, 从时间窗口大小的自适应确定, 纹理和颜色结合的空间域更新提出一种局部时空域模型。

### 2.2.1 构造局部时间域模型

目标检测精度受时间窗口大小影响。通过当前帧背景  $B(x_{c,b})$  和历史帧背景  $B(x_{p,b})$  概率密度函数  $f(x_{c,b}, x_{p,b})$  计算当前帧和历史帧背景匹配度作为调整时间域窗口指标, 自动调整时间域窗口大小。

假设视频帧灰度直方图满足高斯分布 (图 2), 当前帧和历史帧背景分别表示为  $B(x_{c,b}) \sim N(\mu_{c,b}, \sigma_{c,b}^2)$  和  $B(x_{p,b}) \sim N(\mu_{p,b}, \sigma_{p,b}^2)$ , 局部时间域模型步骤如下:

1) 通过当前帧和历史帧背景高斯函数模型, 计算两者所构成的概率密度函数

$$f(x_{c,b}, x_{p,b}) = \frac{1}{2\pi(\det C)^{\frac{1}{2}}} \times \exp\left\{-\frac{1}{2}(X - \mu)^T C^{-1}(X - \mu)\right\} \quad (10)$$

式中,  $C$  为协方差矩阵,  $X$  为像素向量,  $\mu$  为均值向量。

2) 根据概率密度计算前景和背景的匹配度

$$r_{cp} = \frac{c_{ij}}{\sigma_{c,b}\sigma_{p,b}} + \zeta, i \neq j; r_{cp} \in [0, 1] \quad (11)$$

式中,  $c_{ij}$  为协方差,  $\sigma_{c,b}$  和  $\sigma_{p,b}$  分别是前景和背景的方差,  $\zeta$  为松弛因子, 用于弥补帧间隔所产生的误差, 本文取 0.01。

3) 更新背景 通常匹配度  $r_{cp}$  越高说明背景模型的质量越好, 减小更新速度则时间域窗口增大, 反之则说明当前背景变化较大, 需要加快更新速度, 减小时间域窗口, 及时更新背景。窗口更新为

$$G = \begin{cases} T_{old} + m & r_{cp} > T_w \\ T_{old} - m & r_{cp} \leq T_w \end{cases}, T_w \in [0, 1] \quad (12)$$

式中,  $T_{old}$  表示先前的窗口大小,  $m$  表示增减幅度,  $T_w$  表示匹配度阈值, 本文取 0.65。

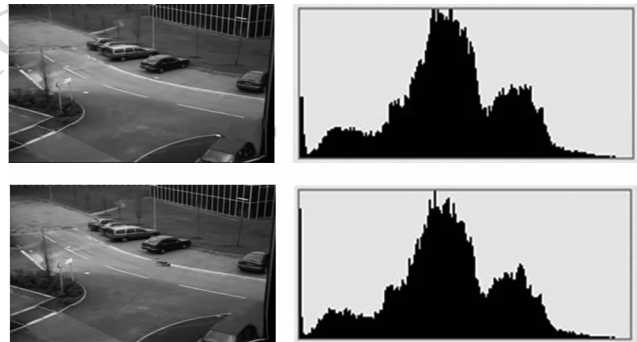


图 2 停车场视频第 450 帧和第 500 帧对应背景颜色直方图  
Fig. 2 The color histogram of frame 450 and 500 background in the park video

### 2.2.2 构造局部空间域模型

仅仅依靠颜色特征在空域更新背景不能很好解决有阴影和部分遮挡等情况的目标检测问题,阴影通常只会导致亮度变化,而纹理特征不会改变,因此,本文采取纹理和颜色特征来更新背景。纹理特征描述采用算子简单,计算量小,且速度快,能够减弱阴影影响的二值差分算子 LBP(local binary pattern)<sup>[15]</sup>。若视频帧与背景参考模型所对应纹理越相似,则该点属于背景概率越大,空间域背景更新如下:

首先,计算  $t$  时刻视频帧和背景参考模型中像素点位置在  $q$  处 LBP 差值

$$D_{LBP}^t(q) = |LBP_{p,R}^t(q) - LBP_{p,R}^b(q)| \quad (13)$$

$$LBP_{p,R}(q) = \sum_{p=0}^{p-1} s(g_p - g_q - n)2^p$$

$$s(u) = \begin{cases} 1 & u \geq 0 \\ 0 & u < 0 \end{cases} \quad (14)$$

式中,  $g_q$  为位置  $q$  处像素灰度值,  $g_p$  为位置  $q$  处周围邻域像素灰度值,这些像素等间隔分布在以  $q$  为中心,以  $R$  为半径的圆环上,  $s(u)$  为二值化值,  $n$  表示克服噪声常数,通常  $n \in [3, 5]$ , 本文取  $n = 4$ 。

然后将前、背景纹理相似点选出并构成集合  $Z$ , 利用 2.2.1 节所述方法选核密度最大像素点为背景。

$$Z = \{map(D_{LBP}^t(q_i), z_i), D_{LBP}^t(q_i) \leq T_{LBP}\} \quad (15)$$

### 2.2.3 局部时空域模型更新背景策略

为有效更新背景模型,分别对非动态背景、动态背景和突变背景 3 种情况采用不同的更新策略,避免全局更新耗时和准确率低等问题。

非动态背景指背景模型没有发生太大变化,只略受光照变化影响。在摄像头固定条件下,大部分背景都属于此种情况。不需要大幅更新背景模型,采用渐进式更新策略:

$$I_b(q) = \begin{cases} (1 - \alpha)I_b(q) + \alpha I_c(q) & T_g(q) = T_g \\ I_b(q) & \text{其他} \end{cases} \quad (16)$$

式中,  $T_g(q)$  为像素点  $q$  未发生变化的帧,  $T_g$  为像素值未发生变化的持续时间,  $\alpha$  为背景模型的更新率,本文取 0.1。

背景模型中有目标移动的动态背景需要删除先前的背景点,再根据时空域模型重新选择像素点。即当  $F_f > \beta$  时,该像素点可能属于动态背景,需要继续观察判断,若该状态持续时间满足  $T_f > T_u$ ,则可以判断背景中已经有新目标移动,采用时空域模型重新选择像素

点。其中  $F_f$  表示前景点频率(指在  $N$  个采样值中被判定为前景点的次数所占的比例),  $\beta$  表示前景点频率阈值,  $T_f$  表示像素点处于观察状态的持续时间,  $T_u$  表示像素点处于观察状态而不需要更新的时间间隔。

因外界因素导致摄像头发生抖动偏移产生的背景突变,根据每帧图像前景点像素统计波动来检测。若前景点比率  $FR$  突然间波动比较大时,可以判断背景存在突变,需要删除背景模型,重新根据局部时空域模型进行背景建模。图 3 给出一个高速公路监控视频,从第 51 帧开始,每隔 10 帧选择一帧作为统计样本,通常前景点比率  $FR$  约为 7%, 如果突然间增加到 70% 以上(第 71 帧),说明背景发生突变,则删除背景模型,重新按照上述算法进行背景建模。

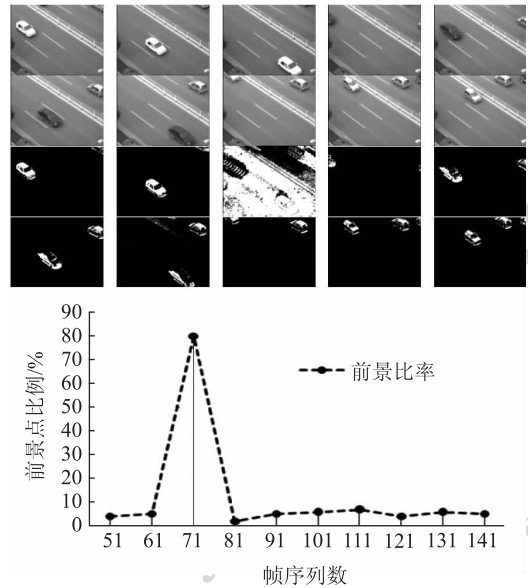


图3 高速公路视频灰度图和前景图以及前景点比率

Fig.3 Gray foreground image and foreground pixel rate

### 2.3 LST-KDE 算法步骤

1) 输入  $N$  帧训练样本集  $X$ , 并用 K-means 聚类算法抽取关键帧建立优化样本集  $Y$ , 选  $Y$  中最大核密度估计值像素点构建参考背景。

2) 第  $N$  帧后的视频帧按照时间域和空间域模型方法,对视频帧每一个像素点构建局部时空域模型,用于后期更新参考背景。

3) 当前视频帧与背景参考模型背景差分,分割前景,并统计前景点所占比例。

4) 更新参考背景:

若前景点比率大于 70%, 背景发生突变,则删除背景参考模型,返回步骤 1) 重新进行背景建模;

否则对像素点进行判断,如是非动态背景采用

渐进式更新;如是动态背景删除当前该背景像素点,且根据局部时空域模型更新该背景像素点。

5) 根据当前和历史背景帧的匹配度以及空间纹理和颜色特征,更新局部时空域模型。

6) 返回步骤3),直到视频全部处理完为止。

### 3 实验结果及分析

实验环境为 IntelCore2@2.0 GHz 的 CPU 1 G 内存的 PC 机器,编程语言是 VC + +,训练采样  $N = 50$ ,背景差全局阈值  $T_c = 45$ ,初始更新时间间隔  $G = 150$ ,

更新幅度  $m = 30$ ,纹理相似度阈值  $T_{LBP} = 20$ ,前景点频率阈值  $\beta = 0.6$ ,LBP 算子半径  $R = 4$ (在文中设定的阈值已由实验验证效果比较好,算法再现时不需要修改,而对于实验分析中所设定的阈值,可依实验视频属性的差异有所改动但调整幅度不大)。本文实验视频为室外停车场监控视频和室内 CAVIAR 视频库中某商场的视频监控,两个视频相关属性是帧宽为 352,帧高为 240,频率为 30 帧/s。

#### 3.1 LST-KDE 算法目标检测性能比较

图 4 和图 5 分别给出了 LST-KDE 方法和文献 [10-12] 算法目标检测结果。

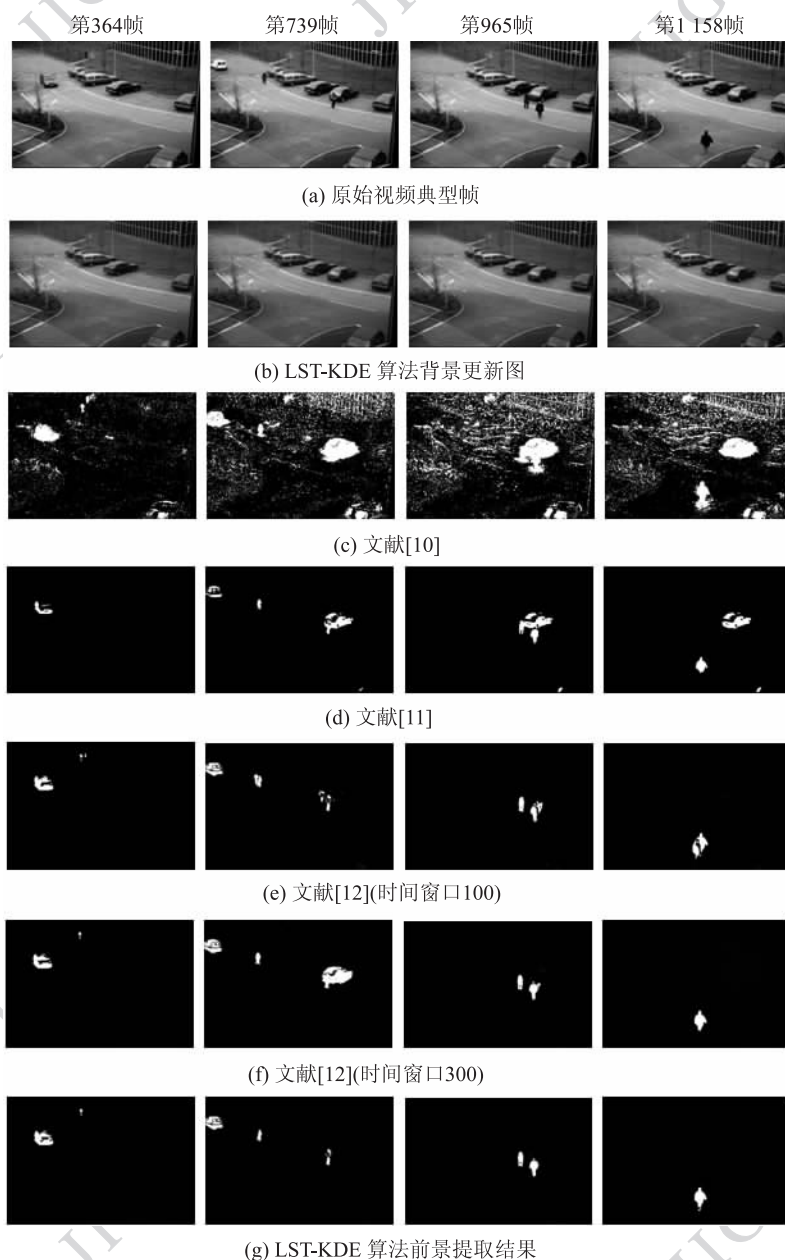


图 4 停车场视频目标检测

Fig.4 Compare LST-KDE with literature [10-12] in park video

图4是室外停车场监控视频,动态背景,需要自适应更新背景。受光照不均影响,文献[10]算法需要一段时间才能将光照变化更新到背景中,导致背景很容易被误检为前景,含有大量噪声,如图4(c)所示;文献[11]采用双阈值方式能够过滤掉很多噪声,但由于背景帧中有物体的移入,如图4(d)第739、965和1158帧所示,造成前景误检,准确率大大降低;文献[12]采用时间窗方法,弥补了文献[11]的不足,但需要在不同的环境下设置阈值,当时间窗取100时,更新时间间隔比较短,由于行人速度比较慢而车辆速度比较快导致前景有拖影

污染,当时间窗取300时,虽然不会产生行人的拖影现象,但由于更新间隔长,若有新目标移入就可能被误检为前景(如图4第6行第739帧所示,在第965帧时背景才开始慢慢更新,直到第1158帧时,才完全更新背景);本文提出的LST-KDE算法弥补了文献[10-12]算法的不足,通过历史帧信息计算当前帧背景和历史帧背景的匹配度,来自适应更新时间域窗口的大小,不会产生拖影和背景被误检为前景现象,如图4(g)所示。

图5是商场室内监控视频,存在光照不均、遮挡和阴影等问题。图5(c)是文献[10]算法所提

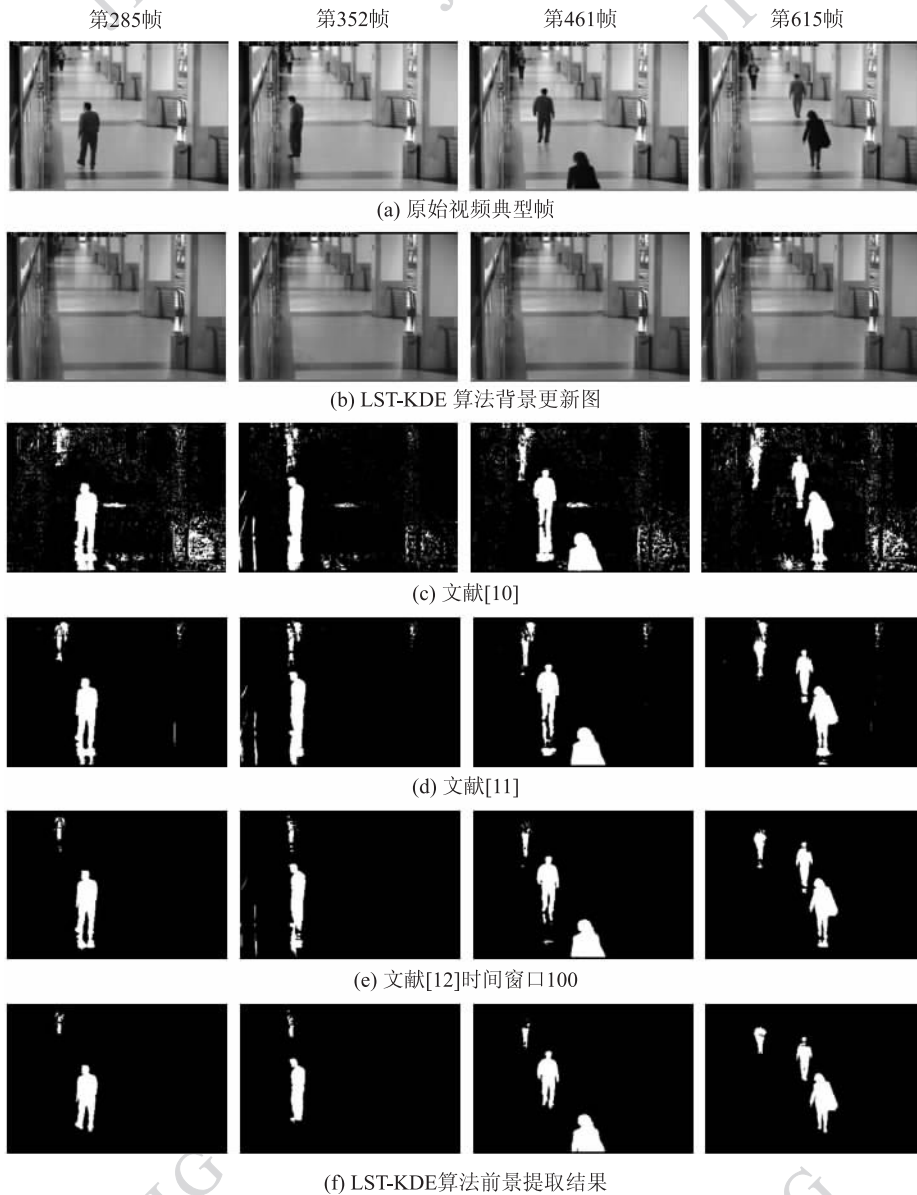


图5 商场视频目标检测

Fig. 5 Compare LST-KDE with literature[10-12] in market video

取的前景目标,由于光照不均背景不能够及时更新,导致大量的噪声和阴影出现,误检率很高;图 5(d)(e)是对文献[10]算法的改进,分别是文献[11-12]提出的算法,这两种算法前景目标提取虽然能够消除大量噪声,但无法消除存在的阴影问题,算法的准确率降低;提出的 LST-KDE 算法利用空间域信息,通过 LBP 描述的纹理特征有效的抑制了阴影,使检测准确率大大提高,如图 5(f)所示。实验证明 LST-KDE 算法能很好分离前、背景目标,去除大量噪声,解决部分阴影问题。

### 3.2 LST-KDE 算法性能分析

为验证提出 LST-KDE 算法在背景建模时间和目标检测误检率与漏检率方面性能进行对比分析。给出室外共 1 161 帧停车场监控视频序列分析结果。

KDE 算法在前期背景建模耗时由视频帧大小和训练样本帧数决定。而视频帧大小由视频属性决定,但可以减少训练样本帧数,解决 KDE 算法在前期背景训练实时性差的问题。基于此,LST-KDE 算法利用 K-means 聚类选出关键帧进行背景建模。图 6 给出了不同训练样本集背景建模时间对比。从图 6 可以看出文献[10]的实时性最差,且随着帧数增加,背景建模时间呈指数增长,文献[11-12]是对文献[10]算法的改进,实时性有所提高,但效果不是很明显,而提出的 LST-KDE 算法对训练样本进行筛选,使训练时间大大减少,建模时间和训练样本集合大小基本保持线性增长,实时性有较大提高。

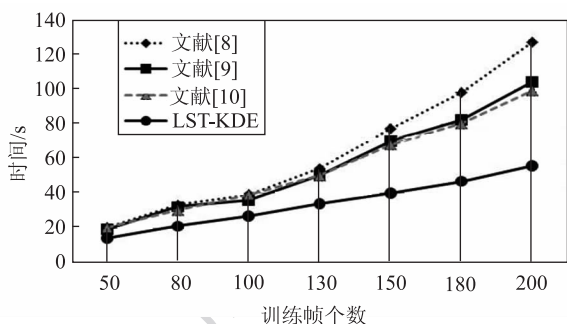


图 6 LST-KDE 算法与其他算法训练时间对比

Fig. 6 LST-KDE and other algorithm time

前景目标检测准确率用误检率和漏检率来分析。误检率(MDR)指背景点被误检为前景点的比例,漏检率(OR)指前景点被误检为背景点的个数。

$$MDR = \frac{MP}{TP + MP} \quad (17)$$

$$OR = \frac{OP}{TP + OP} \quad (18)$$

式中,  $TP$  为实际前景像素,同时正确检测为前景像素的个数,  $MP$  为误检为前景点个数;  $OP$  为前景点被误检为背景点的个数。

对比分析实验从室外停车场视频第 60 帧开始,每隔 100 帧选一个测试样本,对目标检测结果分别计算误检率和漏检率,文献[12]算法时间窗取 300。图 7 给出了 4 种算法误检率对比分析。从图 7 中可以看出,文献[10]由于受到光照变化和背景中目标移动的影响,背景没有自适应更新,误检率在对比方法中最高,文献[11-12]误检率相对较低,但文献[11]在第 560 帧突然上升到 80% 左右,和文献[10]基本持平,主要是因为背景帧中有新目标移入,文献[11]没有自适应更新,导致误检率增加;而文献[12]在第 660 帧误检率突然升到 70%,主要是由于时间窗口太大不能及时更新背景导致背景被误检为前景,直到第 960 帧,背景才慢慢被更新,误检率慢慢处于正常水平,而 LST-KDE 算法可以快速更新背景模型,误检率基本保持在同一水平,误检率为 4 种方法中最小。

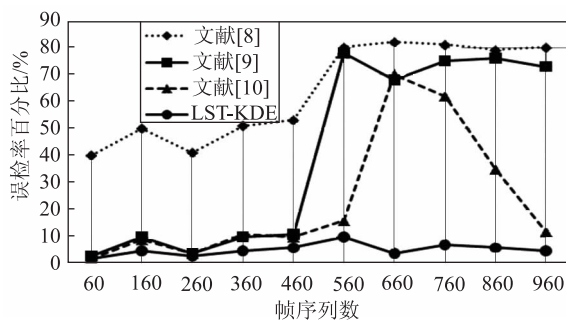


图 7 LST-KDE 算法与其他算法误检率对比

Fig. 7 LST-KDE and other algorithm MDR

误检率和漏检率是一种对立关系,图 8 给出了 4 种算法漏检率对比分析结果。从图 8 可以看出,文献[10]漏检率最低,文献[11]和[12]由于没有能够处理部分遮挡,导致漏检率比较高,LST-KDE 算法由于采用局部时空域模型很好处理了部分遮挡和阴影问题,得到完整的前景目标,漏检率相对也比较低,和文献[10]的漏检率基本持平。综合图 7 和图 8 的比较分析,LST-KDE 算法的误检率和漏检率都比较低,相比其他几个算法具有较好的检测准确性。

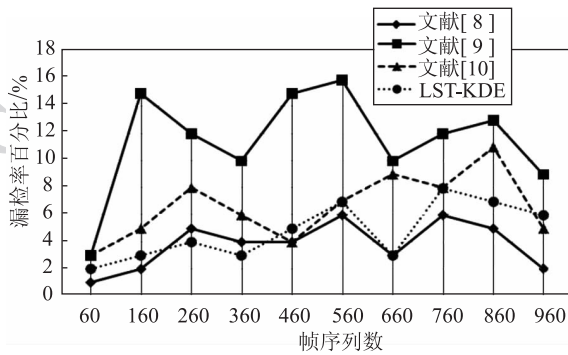


图 8 LST-KDE 算法与其他算法漏检率对比

Fig. 8 LST-KDE and other algorithm OR

## 4 结 论

提出一种局部时空域模型的核密度估计目标检测方法。该方法利用 K-means 聚类选择关键帧进行背景建模,解决了 KDE 训练时计算量大、数据冗余等问题;在后期背景更新阶段采用局部时空域模型自适应调整阈值,并且能够有效抑制阴影,检测效果较好;最后通过大量室内和室外复杂场景的实验,表明该算法在实时性和准确率方面明显优于其他算法。

### 参考文献 (References)

- [ 1 ] Jong B K, Hang J. Efficient region-based motion segmentation for a video monitoring system [ J ]. Pattern Recognition Letters, 2003, 24(3): 113-128.
- [ 2 ] Barron J, Fleet D, Beauchemin S. Performance of optical flow techniques[ J ]. International Journal of Computer Vision, 1994, 12(1):42-77.
- [ 3 ] Nuria M O, Barbara R, Alex P P. A Bayesian computer vision system for modeling human interactions[ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(8): 831-843.
- [ 4 ] Stauffer C, Grimson W E L. Learning patterns of activity using real-time tracking[ J ]. IEEE Transactions on Pattern Recognition Machine Intelligence, 2000, 22(8):747-757.
- [ 5 ] Lee D. Effective gaussian mixture learning for video background subtraction[ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(5):827-832.
- [ 6 ] Bai X F, Li A H, Li X L, et al. A novel background Gaussian mixture model [ J ]. Journal of Image and Graphics, 2011, 16(6):983-988. [ 白向峰,李艾华,李喜来,等. 新型背景混合高斯模型[ J ]. 中国图象图形学报, 2011, 16(6):983-988. ]
- [ 7 ] Esfandian N, Razzazi F, Behrad A, et al. A feature selection method in spectro-temporal domain based on gaussian mixture models[ C ]// Proceedings of International Conference on Signal Processing. Beijing, China: IEEE, 2010:522-525.
- [ 8 ] Jodoin P, Mignotte M, Konrad J. Statistical background subtraction using spatial cues[ J ]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(12):1758-1763.
- [ 9 ] Caseiro R, Henriques J F, Batista J. Foreground segmentation via background modeling on riemannian manifolds [ C ]// Proceedings of International Conference on Pattern Recognition. Istanbul, Turkey: IEEE, 2010:3570-3574.
- [ 10 ] Elgammal A, Duraisvami R, Hawood D. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance[ J ]. Proceedings of the IEEE, 2002, 90(7): 1151-1163.
- [ 11 ] Xu D B, Huang L, Liu C P. Adaptive kernel density estimation for motion detection[ J ]. Acta Automatica Sinica, 2009, 35(4): 379-385. [ 徐东彬,黄磊,刘昌平. 自适应核密度估计运动检测方法[ J ]. 自动化学报, 2009, 35(4):379-385. ]
- [ 12 ] Zhou E C, Liu C P, Zhang L Y, et al. Foreground object detection based on time information window adaptive kernel density estimation[ J ]. Journal on Communications, 2011, 32(3):106-115. [ 周恩策,刘纯平,张玲燕,等. 基于时间窗的自适应核密度估计运动检测方法[ J ]. 通信学报, 2011, 32(3):106-115. ]
- [ 13 ] Bors A G, Nasios N. Kernel bandwidth estimation for nonparametric modeling[ J ]. IEEE Systems Man and Cybernetics Society, 2009, 39(6):1543-1555.
- [ 14 ] Asgharbeygi N, Maleki A. Geodesic K-means clustering [ C ]// Proceedings of International conference on Pattern Recognition. Tampa, FL, USA: IEEE, 2008:1-4.
- [ 15 ] Heikkila M, Pietikainen M. A texture-based method for modeling the background and detecting moving objects[ J ]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2006, 28(4):657-662.