第38卷 第24期
Vol.38 No.24

计 算 机 工 程
Computer Engineering

2012年12月
December 2012

·开发研究与设计技术· 文章编号：1000—3428(2012)24—0291—04 文献标识码：A 中图分类号：TP18

基于自适应时空码书的粒子滤波多目标跟踪

徐 悦，肖 刚，张 冉

(上海交通大学航空航天学院，上海 200240)

摘 要：提出一种基于自适应时空码书检测模型的粒子滤波多目标跟踪算法。使用时空码书模型进行前景背景分割，检测出前景目标，在该模型上加入目标自适应过程。将自适应时空码书检测的结果作为粒子滤波跟踪算法的初始目标状态，通过关联算法和粒子滤波实现多目标跟踪。自适应时空码书模型能明显降低对前景目标的误检率，抑制噪声干扰。实验结果表明，该算法能够在有干扰的复杂背景下实现对运动多目标的快速捕获，并有效提高跟踪的可靠性和精度。

关键词：时空码书；自适应；粒子滤波；跟踪；多目标

Particle Filtering Multi-objective Tracking Based on Adaptive Spatio-temporal Codebook

XU Yue, XIAO Gang, ZHANG Ran

(School of Aeronautics and Astronautics, Shanghai Jiaotong University, Shanghai 200240, China)

【Abstract】A particle filtering tracking algorithm based on adaptive spatio-temporal Codebook detection model is designed to realize multi-objective tracking. Spatio-temporal codebook model is used for foreground-background segmentation as well as detecting foreground targets, and an adaptive phrase of the targets is added to this model. The prior target state distribution of particle filter is generated by detected foreground targets, using association algorithm and particle filter to realize multi-objective tracking. Compared with traditional codebook and spatio-temporal codebook model, the proposed adaptive spatio-temporal codebook model has robustness to interference and noise. Experimental results show that the particle filter tracking algorithm can capture moving targets rapidly with higher validity in moving background with illumination change in existence of interference and noise.

【Key words】spatio-temporal codebook; adaptive; particle filtering; tracking; multi-objective

DOI: 10.3969/j.issn.1000-3428.2012.24.069

1 概述

运动目标跟踪是各类视频处理和理解的基础，在智能视频监控、基于模型的识别、人机交互、视觉伺服等领域都有重要应用。粒子滤波算法是广泛应用于各领域的后验概率逼近技术，能够处理任意的状态空间模型。基于检测的跟踪能为粒子滤波算法提供更准确的目标初始状态，从而获得更可靠的目标状态的后验概率分布，有效提高跟踪结果的准确性。

由于粒子滤波对模型和噪声分布没有限制，因此在目标跟踪领域得到了广泛研究。当前很多的多目标跟踪算法都是在检测的基础上实现的[1]。文献[2]提出了一个多目标跟踪和判断目标存在性的通用框架，将检测结果作为观测数据，算法鲁棒性较好，但目标出现时关联存在滞后，且对跟踪目标的数目有限制。文献[3]利用对象互动机制等检测后验模式辅助目标跟踪，能得到高效的粒子配置，但当一组内目标存在较强机动时会出现关联算法失效、跟踪失败。文献[4]将混合粒子滤波跟踪器[5]与 AdaBoost 检测算法结合，能跟踪数目可变的多目标，并较好地处理视频中目标的进入和移出。但 AdaBoost 算法需要样本量大，训练过程复杂。

与仅依靠保存目标信息相比，利用检测结果跟踪难度较低，且准确度更高。码书模型是常用的背景建模方法[6]。使用该模型的实时前景检测技术，能够处理运动背景或光照变化的场景，可以较好地抑制背景中的准周期运动的干扰，但对大幅值干扰运动的抑制能力较弱。文献[7]在标准码书基础上提出了利用时间和空间的上下文信息，进一步改善了算法的抗干扰性。

为了使跟踪算法在时间和空间上都有尽量低的误检率，本文提出了基于自适应时空码书的粒子滤波算法，通过自适应时空码书算法检测出前景目标；由检测结果给出

基金项目：国家自然科学基金资助项目(60904096)；航空科学基金资助项目(20095557010)

作者简介：徐 悦(1988－)，女，硕士研究生，主研方向：图像处理，视频跟踪；肖 刚，副教授；张 冉，硕士研究生

收稿日期：2011-09-20 **修回日期：**2011-10-20 **E-mail：**xuyue1988mail@gmail.com

跟踪算法中目标状态的先验分布，利用粒子滤波跟踪算法实现多目标跟踪。

2 自适应时空码书前景目标检测算法

2.1 标准码书和时空码书的前景检测

Kim K 提出的码书前景检测算法使用长观测序列建立背景模型，对每个像素创建码书，根据背景变化确定码字的数量。Mingjun W 提出的时空码书根据局部空间依赖关系和先验检测结果，将时空上下文信息引入由马尔可夫随机场理论构建的空间模型中。

$X=\{\boldsymbol{x}_1,\boldsymbol{x}_2,\cdots,\boldsymbol{x}_N\}$ 表示了含 N 维 RGB 向量的像素训练序列；$C=\{c_1,c_2,\cdots,c_L\}$ 为像素包含 L 个码字的码书。每个码字 c_i，$i=1,2,\cdots,L$ 包含一个匹配该码字的 RGB 向量平均值 $\boldsymbol{v}_i=(\overline{R}_i,\overline{G}_i,\overline{B}_i)$ 和数组 $\text{aux}_i=\langle\check{I}_i,\hat{I}_i,f_i,\lambda_i,p_i,q_i\rangle$，$\check{I},\hat{I}$ 表示匹配该码字的全部像素的亮度最小值和最大值；f 表示当前码字被匹配的次数；λ 表示当前码字未更新的最长等待时间；p,q 表示码字首、末次的成功匹配时刻。

对于 t 时刻像素采样值 $\boldsymbol{x}_t$ 是否与当前码书中的码字 c_m 匹配需根据以下准则确定：(1)与该码字的颜色失真度小于匹配阈值。(2)在码书的亮度范围内。成功匹配的像素标记成背景，否则标记成前景。观测像素 $\boldsymbol{x}_t$ 和相应码字矢量 $\boldsymbol{v}_i$ 的颜色失真度定义为：

$$d(\boldsymbol{x}_t,\boldsymbol{v}_i)=\sqrt{\|\boldsymbol{x}_t\|^2-\langle\boldsymbol{x}_t,\boldsymbol{v}_i\rangle^2/\|\boldsymbol{v}_i\|^2} \tag{1}$$

观测像素和相应码字的逻辑亮度函数为：

$$bright\left(\|\boldsymbol{x}_t\|,\langle\check{I}_i,\hat{I}_i\rangle\right)=\begin{cases}\text{true} & \text{if } I_{\text{low}}\leqslant\|\boldsymbol{x}_t\| \quad I_{\text{hi}}\\ \text{false} & \text{otherwise}\end{cases} \tag{2}$$

其中，$I_{\text{low}}=\alpha\hat{I}_i$；$I_{hi}=\min\{\beta\hat{I}_i,\check{I}_i/\alpha\}$。$\alpha$ 和 β 是预先给定的参数：$\alpha<1$，一般取 0.4~0.7；$\beta>1$，一般取 1.1~1.5。$[I_{\text{low}},I_{\text{hi}}]$ 表示的范围在更新码书的过程中相对稳定。前景检测应先选择不含前景或前景短暂出现的一段图像序列训练码书，得到初始码书模型。在空间码书中，对于未被标准码书匹配为前景的像素 $\boldsymbol{x}$，将其与相邻像素的码书进行二次匹配。

2.2 自适应时空码书

为获得更平稳、噪声少的检测效果，本文在空间码书基础上，在变化光照条件下自适应地更新码字，以减少全局光照时的误检率、提高了算法的鲁棒性。同时，选用合适学习率的码字矢量和方差的指数平滑法，通过更新 $\boldsymbol{v}_m$ 和增加 σ_m^2 更新颜色失真度：

$$\begin{cases}\boldsymbol{v}_m\leftarrow\gamma\boldsymbol{x}_t+(1-\gamma)\boldsymbol{v}_m\\ \sigma_m^2\leftarrow\rho\delta^2+(1-\rho)\sigma_m^2\end{cases} \tag{3}$$

其中，γ 和 ρ 代表学习率；σ_m^2 是根据第 1 帧初始化的颜色模型的方差。γ 基于检测时间和相似性测度进行调整，可由多次实验选定：

$$\gamma=\begin{cases}\dfrac{1}{t}\cdot\left[1+\dfrac{M(t)}{1-\alpha}\right] & \text{if } 0.5\leqslant M(t)\quad 0.9\\ \dfrac{1}{t}\cdot\dfrac{M(t)}{1-\alpha} & \text{otherwise}\end{cases} \tag{4}$$

其中，t 为检测时间；$M(t)$ 为 2 帧图像的相似性测度，定义如下：

$$M(t)=\frac{2\cdot\sum_{(x,y)\in t,t-1}I_{t-1}(x,y)\cdot I_t(x,y)}{\sum_{(x,y)\in t-1}\left[I_{t-1}(x,y)\right]^2+\sum_{(x,y)\in t}\left[I_t(x,y)\right]^2} \tag{5}$$

改进的颜色失真度表达式为：

$$d(\boldsymbol{x}_t,\boldsymbol{v}_i)=\delta/\sigma_i$$

为保证码书的可靠性，当搜索空间邻域码书时，如找到匹配的码字，不对其进行更新，只更新自适应全局方差 σ_m^2。该算法的描述如下：

算法 自适应时空码书算法

$L\leftarrow 0^1$，$C\leftarrow\varnothing$

对于 $t=1,2,\cdots,N$ 执行

(1) $\boldsymbol{x}_t=(R,G,B)$；$I\leftarrow\sqrt{R^2+G^2+B^2}$

$\sigma_m^2\leftarrow\rho\delta^2+(1-\rho)\sigma_m^2$；$d(\boldsymbol{x}_t,\boldsymbol{v}_i)=\delta/\sigma_i$

(2)搜索码书 $C=\{c_i|1\leqslant i\quad L\}$ 中 $\boldsymbol{x}_t$ 满足条件 1)、条件 2)的码字 c_m：

1) $d(\boldsymbol{x}_t,\boldsymbol{v}_m)\leqslant\varepsilon_1$

2) $brightness(I,\langle\check{I}_m,\hat{I}_m\rangle)=\text{true}$

(3)若 $C=\varnothing$ 则 $L\leftarrow L+1$ 新建码字 c_L：

$v_L\leftarrow(R,G,B)$；$\text{aux}_L\leftarrow\langle I,I,1,t-1,t,t\rangle$

(4)若搜索到码字：$l_x=b$，更新该码字：

$\boldsymbol{v}_m\leftarrow\gamma(R,G,B)+(1-\gamma)(\overline{R}_m,\overline{G}_m,\overline{B}_m)$

$\text{aux}_m\leftarrow\langle\min\{I,\check{I}_m\},\max\{I,\hat{I}_m\},f_m+1,$

$\max\{\lambda_m,t-q_m\},p_m,t\rangle$

否则搜索结果为空，则：

对于每个 $Uy\in\{Uy|y\in Nx\}$ 搜索符合条件 1)、条件 2)的码字：

若搜索到码字：

更新 σ_m^2，$\sigma_m^2\leftarrow\rho\delta^2+(1-\rho)\sigma_m^2$

否则 $L\leftarrow L+1$ 新建码字 c_L：

$\boldsymbol{v}_L\leftarrow(R,G,B)$；$\text{aux}_L\leftarrow\langle I,I,1,t-1,t,t\rangle$

结束 Uy 搜索

结束 C 搜索

计算 $M(t)$，更新学习率 γ

在训练阶段对于每个码字 $c_i,i=1,2,\cdots,L$ 设置每个 λ_i 为：

$\lambda_i\leftarrow\max\{\lambda_i,(N-q_i+p_i-1)\}$

将得到的自适应空间码书结果放到 Mingjun W 提出的三层 Markov 模型中。根据贝叶斯观点，前景背景分割就是建立最准确的像素标签集合 L_t 的构形，即推断 MRF 的最大后验解。利用改进的 MMD[8]最小化后验能量函数，将时间码书的检测结果作为 MMD 的初始化配置，最终得到自适应时空码书的检测结果。

3 基于检测的粒子滤波跟踪算法

状态先验分布可通过上文自适应时空码书获得，再利用粒子滤波方法实现目标跟踪。在序列图像多目标跟踪时可认为多个目标的运动相互独立，则可定义模型为单目标运动模型的乘积。定义 t 时刻状态变量为：

$$X_t=[(x_{1,t},e_{1,t}),(x_{2,t},e_{2,t}),\cdots,(x_{M_t,t},e_{M_t,t})]$$

其中，M_t 表示 t 时刻的目标状态数；$x_{i,t}$，$i=1,2,\cdots,M_t$ 表示

t 时刻第 i 个目标的状态；$e_{i,t}$ 表示 $x_{i,t}$ 在 t 时刻的存在状态，若存在则 $e_{i,t}$ 为 1。在实验中目标用矩形框标出 $x_{i,t}=[x_i,y_i,w_i,h_i]$，其中，$x_i$ 和 y_i 表示矩形框中心在图像中的坐标；w 和 h 为长和宽。多目标运动模型为：

$$p(X_t\mid X_{t-1})=\prod_{i=1}^{M}p(x_{i,t},e_{i,t}\mid x_{i,t-1},e_{i,t-1})=\prod_{i=1}^{M}p(x_{i,t}\mid x_{i,t-1},e_{i,t},e_{i,t-1})p(e_{i,t}\mid e_{i,t-1}) \tag{6}$$

其中，M 为大于等于 t 与 $t-1$ 时刻的目标状态数。变量 $e_{i,t}$ 用于处理跟踪期间目标短时出现和消失的情况。若连续多帧图像搜索后未发现与该状态对应的观测值，则目标可能已经消失；若在连续多帧图像搜索内都有观测目标与某一新出现目标的状态对应，则该目标很可能是真实目标的[9]。新出现的目标的初始状态由自适应码书检测结果进行初始化。

在确定新目标为真实目标前，其观测量仅由检测结果描述；当确定为真实目标后，保存其颜色模板信息，此时可以用检测结果和颜色特征信息共同表示观测。目标观测可认为相互独立，假设目标观测与状态已数据关联，多目标的观测模型如下：

$$p(Y_t\mid X_t)=\prod_{i\in I_f}p_{\text{false}}(y_{i,t}^{d})\prod_{k\in\{1,2,\cdots,M_t\}}p_{\text{tar}}(y_{r_k,t}^{d},y_{k,t}^{c}\mid x_{k,t},e_{k,t}) \tag{7}$$

其中，$y_{i,t}^{d}$ 为 t 时刻基于检测结果的观测值中第 i 个观测对应的是误检目标，其观测仅由检测结果描述；I_f 为基于检测结果的观测值中属于误检的集合；$\{1,2,\cdots,M_t\}$ 为 t 时刻目标状态数的集合；$y_{r_k,t}^{d}$ 为 t 时刻基于检测结果的观测值中第 r_k 个观测值；$y_{k,t}^{c}$ 是状态为 $x_{k,t}$ 的目标颜色观测信息。

用于对确定为真实目标的观测描述的检测和颜色观测可认为是相互独立的，则式(7)中的 $p_{\text{tar}}(y_{r_k,t}^{d},y_{k,t}^{c}\mid x_{k,t},e_{k,t})$ 可以表示为两者的乘积：

$$p_{\text{tar}}(y_{r_k,t}^{d},y_{k,t}^{c}\mid x_{k,t},e_{k,t})=p_{\text{tar}}(y_{r_k,t}^{d}\mid x_{k,t},e_{k,t})p_{\text{tar}}(y_{k,t}^{c}\mid x_{k,t},e_{k,t}) \tag{8}$$

本文进一步通过简单有效的拍卖分配数据关联算法将多目标跟踪问题转化成单目标跟踪。关联矩阵的行数等于观测数、列数等于状态数，而矩阵元素则由当前的观测值和状态粒子集计算。该矩阵用来判断关联度：如果矩阵的第 (k,l) 项元素近似等于 0，则认为该时刻第 k 个观测与第 l 个目标状态集不关联；(k,l) 项元素的取值越大，则第 k 个观测与第 l 个目标状态集的关联性越大。得到该矩阵后使用拍卖分配算法可以将观测与各目标状态粒子集的进行关联。

执行数据关联后用一组带权重的粒子 $\{(x_{i,t}^{(j)},e_{i,t}^{(j)}),w_{i,t}^{(j)}\}_{j=1}^{N}$，$i=1,2,\cdots,M_t$ 逼近各目标的状态后验分布：

$$p(X_{i,0t}\mid y_{r_i,1t})\approx\sum_{j=1}^{N}w_{i,t}^{(j)}\delta(X_{i,0t}-X_{i,0t}^{(j)}),\ i=1,2,\cdots,M_t \tag{9}$$

其中，$X_{i,0t}=\{X_{i,0},X_{i,1},\cdots,X_{i,t}\}$；$X_{i,t}^{(j)}=\left(x_{i,t}^{(j)},e_{i,t}^{(j)}\right)$；$y_{r_i,1t}=\{y_{r_i,1},y_{r_i,2},\cdots,y_{r_i,t}\}$。$y_{r_i,t}$ 为 t 时刻第 r_i 个观测值与状态 $X_{i,t}$ 的关联。重要性函数选为状态的转移方程，$w_{i,t}^{(j)}$ 可以通过下式计算：

$$w_{i,t}^{(j)}=w_{i,t-1}^{(j)}p(Y_{i,t}\mid X_{i,t}^{(j)}) \tag{10}$$

$p(Y_{i,t}\mid X_{i,t}^{(j)})$ 为式(8)表示的真实目标观测。每个目标的粒子采样过程要考虑该目标出现和消失。对辅助变量 $e_{j,t}$ 的采样定义了 4 种情况：新目标出现，目标持续存在，目标消失，目标未确定。目标状态的采样定义了 3 种情况：目标可能消失，目标持续存在，新目标出现。

对归一化后的权值进行重采样可以获取各目标状态的后验分布，再利用最小均方根误差法估计目标状态。新生成的目标只作为粒子状态集保留，不计入状态数，经过几帧的判断确定其为目标后再对其进行估计。

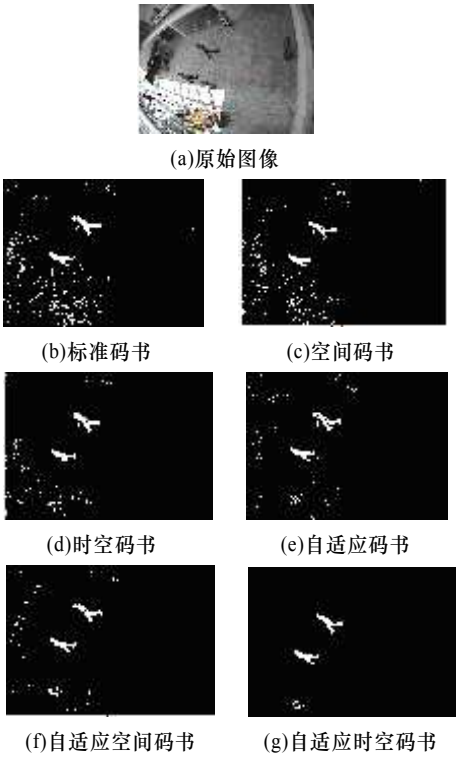
可由关联矩阵中的元素值判断目标的出现和消失。经过连续几帧图像搜索判断目标消失的可能性较大后，如果目标的存在概率小于设定阈值，则认为目标消失，需要删除与该目标相应的状态粒子。如果多帧图像搜索发现某观测没有与之关联的航迹，则认为是新目标或杂波。在连续几帧图像搜索判断新目标为真实目标的可能性较大后，如果目标存在概率大于设定阈值，则认为是真实目标；否则视为杂波，予以剔除。

4 仿真结果与分析

为检验本文算法的有效性，通过 2 组实验将本文算法的检测、跟踪结果分别与标准码书、空间码书、时空码书三者的结果进行对比。

4.1 仿真结果

实验 1 采用 Meet_WalkTogether1 视频(288×384 像素)，数据来源和对应的检测结果如图 1 所示。



(a)原始图像 (b)标准码书 (c)空间码书 (d)时空码书 (e)自适应码书 (f)自适应空间码书 (g)自适应时空码书

图 1 Meet_WalkTogether1 检测结果

从图 1 可以看出，传统算法加入本文的自适应过程后，明显提高了对噪声的抑制能力。自适应时空码书能够准确

地检测出运动目标，并在栅栏状阴影下噪声严重的区域，能够检测到较小的运动目标。序列中第 56、101、153、222、261、302 帧的跟踪结果如图 2 所示。

图 2 Meet_WalkTogether1 检测跟踪结果

由图 2 可知，本文算法可以快速地捕获第 153 帧出现的新目标，2 个目标在 302 帧重叠后也能被正确地关联。

实验 2 采用 Pets 视频(240×320 像素)，时空码书与自适应时空码书的颜色失真均方根误差如图 3 所示。实验数据的来源和第 51、92、137、168、191、317 帧的实验结果如图 4 所示。

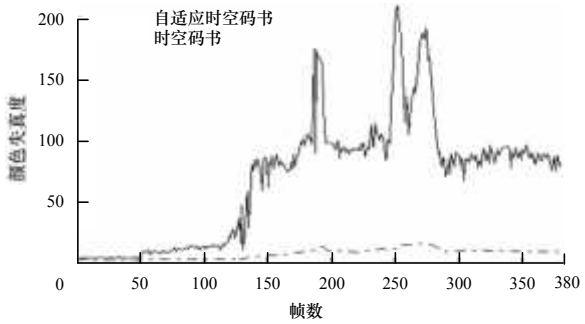
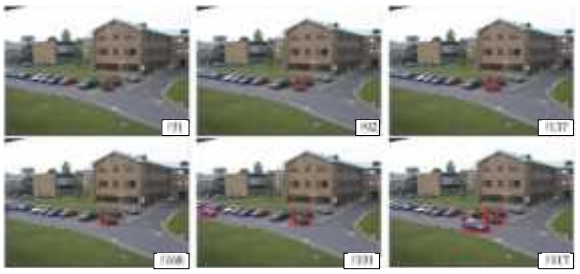


图 3 颜色失真均方根误差对比

(a)标准码书



(b)自适应时空码书

图 4 Pets 检测跟踪结果

从图 3 可以看出，自适应时空码书检测颜色失真均方根误差更加均匀，且误差范围更稳定。视频左侧中间边缘处是会晃动的树枝及树叶，从图 4(a)可以看出，树叶被误检为前景，第 317 帧图像中也有其他干扰被误检为新目标；图 4(b)中对干扰都具有抑制效果，具有较强的鲁棒性。

4.2 实验分析

精确度：即检测出的目标像素的个数与检测出的所有像素的比值[10]。随机抽取 10 帧图像的检测结果其平均值为 99.7%。

跟踪目标中心与真实目标中心的距离：随机抽取 10 帧图像的跟踪结果平均值为 3.2 个像素。

计算复杂度：对于视频 1 和视频 2 均为 $O(10^{11})$，使用时空码书的计算复杂度也为 $O(10^{11})$，在计算复杂度不变的情况下，本文算法提高了检测的准确性，对干扰有更好的抑制效果。

5 结束语

针对带噪声干扰的复杂背景下运动多目标快速检测与跟踪问题，本文通过加入自适应过程对时空码书算法进行了改进，在不增加算法复杂度的情况下对干扰和噪声有较强的抑制；并将自适应时空码书与粒子滤波相结合，使得跟踪算法能够快速捕获新目标，并提高了跟踪的可靠性。实验结果表明，本文提出的算法对不同视频的多目标跟踪都有良好的效果。

参考文献

[1] 徐 悦. 民机 ISS 系统中红外/可见光的跑道监视算法研究[D]. 上海: 上海交通大学, 2012.

[2] Vermaak J, Maskell S. A Unifying Framework for Multi-target Tracking and Existence[C]//Proc. of International Conference on Information Fusion. Cambridge, UK: [s. n.], 2005: 250-258.

[3] Cheng C, Ansari R, Khokhar A. Multiple Object Tracking with Kernel Particle Filter[C]//Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S. 1.]: IEEE Press, 2005: 20-25.

[4] Okuma K, Taleghani A, Freitas N D, et al. A Boosted Particle Filter: Multitarget Detection and Tracking[C]//Proc. of European Conference on Computer Vision. [S. 1]: Springer, 2004: 28-39.

[5] Vermaak J, Doucet A, Perez P. Maintaining Multi-modality Through Mixture Tracking[C]//Proc. of International Conference on Computer Vision. Cambridge, UK: [s. n.], 2003: 1110-1116.

[6] Kyungnam K, Chalidabhongse T, Harwood D, et al. Real-time Foreground-background Segmentation Using Codebook Model[J]. Real-time Imaging, 2005, 11(3): 172-185.

[7] Wu Mingjun, Peng Xianrong. Spatio-temporal Context for Codebook-based Dynamic Background Subtraction[J]. International Journal of Electronics and Communications, 2010, 64(8): 739-747.

[8] Berthod M. Bayesian Image Classification Using Markov Random Fields[J]. Image and Vision Computing, 1996, 14(4): 285-295.

[9] 李安平. 复杂环境下的视频目标跟踪算法研究[D]. 上海: 上海交通大学, 2006.

[10] Ulusoy I, Yuruk H. New Method for the Fusion of Complementary Information from Infrared and Visual Images for Object Detection[J]. IET Image Processing, 2011, 5(1): 36-48.

编辑 索书志