

复杂背景下多目标彩色分割算法

李红波,王 玮,吴 渝

(重庆邮电大学 网络智能研究所,重庆 400065)

(dwtianya2008@163.com)

摘 要:运动视频多目标分割中的背景建模对环境变化有较大的依赖性,直接运用背景差分法会产生不理想甚至是错误的分割。提出了一种基于 Kalman 滤波理论的改进码书背景建模算法。根据码书为每个像素建立一个彩色模型,用来区分前景和背景像素,并利用 Kalman 滤波器的时域递归低通滤波特性对码书背景更新模型进行了校正。实验结果表明,该算法可以有效地更新背景模型,抗干扰能力强,在复杂背景条件下可精确分割出运动目标并满足实时性要求。

关键词:背景模型;彩色模型;码书;运动目标

中图分类号: TP391.9 **文献标志码:** A

Multi-objects colorful segmentation algorithm for complex background

LI Hong-bo, WANG Wei, WU Yu

(Institute of Web Intelligence, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: The background modeling of multi-object segmentation of moving video relies much on the change of environment. Making use of the background subtraction directly may lead to unsatisfactory even wrong segmentation. An improved algorithm of the codebook background modeling based on Kalman filtering theory was proposed. According to the codebook, a colorful model for each pixel was built to distinguish the foreground and background pixels, and according to the characteristics of temporal recursive low passed of Kalman filtering, it was used to rectify the codebook background modeling. The experimental results show that the improved algorithm can update background model efficiently, has the strong anti-jamming ability and segregate the moving objects accurately to meet the real-time requirements.

Key words: background model; colorful model; codebook; moving object

0 引言

视频监控系统得到越来越广泛的应用。如何完整地分割出运动目标是监控视频中一个重要的问题,特别是对于复杂背景条件下的目标分割更是一种重要的研究方向。目前运动目标分割方法可以归纳为三种:背景减法、帧差法、光流法^[1-2]。但远没有完全解决复杂环境中的运动目标分割问题,这主要是由于不论在室内还是在室外环境,随时都有很多引起背景变化的情况产生。因此,建立一个好的自适应的背景模型直接影响运动目标分割的效果。

目前已经提出了许多建立自适应背景模型的方法,文献[3]利用统计的背景相减方法,采用将颜色分解为亮度和色度的模型,克服了对光照强度敏感带来的问题,提高了检测率,但它只适用于静止背景下无运动前景的背景建模。文献[4]利用混合高斯模型建立背景模型,利用多个多维高斯分布来混合模拟背景值,虽然克服光照变化、树叶摇晃等造成的影响,但对运动物体长期静止时检测失败,而且,由于需要计算概率分布,很难达到实时处理的效果。文献[5]在 Stauffer 模型^[4]的框架上提出了不同的模型更新方式,即在初始学习阶段使用足够的统计信息更新方程,之后转入基本的迭代更

新方式,解决了混合高斯模型初始学习速度慢的问题,但这样不利于后面阶段学习新的前景物体。文献[6-8]提出的码书模型为每个像素点的连续采样值根据颜色相似度及其亮度范围生成背景模型。该算法计算复杂度小,占内存少。但背景发生剧烈的变化时,背景模型更新速度较慢。

为了克服上述问题提高分割的准确度,本文利用 Kalman 滤波器^[9]的时域递归低通滤波特点,结合码书背景建模方法,提出了一种基于 Kalman 滤波器的改进码书背景建模,在一定程度上提高了码书背景建模的可靠性和鲁棒性。

1 相关技术基础

1.1 码书背景建模

码书模型为每个像素点的连续采样值根据其颜色相似度及亮度范围生成背景模型。假设 $x = \{x_1, x_2, \dots, x_N\}$ 是一个像素的序列采样值,其中 $x_i (i = 1, \dots, N)$ 是当前像素 t 时刻的 RGB 颜色值。并设 $C = \{c_1, c_2, \dots, c_L\}$ 是该像素的码书,每个码书 $c_i (i = 1, \dots, L)$ 定义为二元组结构:

$$v_i = (\overline{R_i}, \overline{G_i}, \overline{B_i})$$

$$u_i = \langle I_i^{\min}, I_i^{\max}, f_i, \lambda_i, p_i, q_i \rangle$$

其中: I_i^{\min} 和 I_i^{\max} 是码字对于像素的最小和最大亮度值; f 是

收稿日期:2009-06-10;修回日期:2009-08-29。 基金项目:教育部新世纪人才支持计划;重庆市科技攻关项目(7818);重庆市自然科学基金资助项目(2005BB2063);重庆市教委科学技术项目(050509;060504;060517)。

作者简介:李红波(1970-),男,甘肃兰州人,高级工程师,主要研究方向:数字媒体、机器视觉;王玮(1983-),男,安徽安庆人,硕士研究生,主要研究方向:机器视觉;吴渝(1970-),女,重庆人,教授,博士生导师,博士,主要研究方向:网络智能、数据挖掘。

码字出现的频率,记录该码字出现的次数; λ 表示码字在训练中没有再次出现的最大时间间隔; p 和 q 分别表示该码字出现后的第一次和最后一次匹配时间,可简单设为帧序号。码书背景建模的详细过程如下。

第 1 步 首先将每个像素的码书置空,即 $L = 0$ 。

第 2 步 对于训练视频的每个像素的序列值 $x = \{x_1, x_2, \dots, x_N\}$, $x_t = (R_t, G_t, B_t)$, $t = 1, \dots, N$ 。

1) 如果码书为空, $L = 0$, 则创建一个码字:

$$\begin{cases} L = L + 1 \\ I = \sqrt{R_t^2 + G_t^2 + B_t^2} \\ v_L = (R_t, G_t, B_t) \\ U_L = \langle I, I, 1, t-1, t, t \rangle \end{cases} \quad (1)$$

2) 如果码书不为空,则根据下面两个条件从码书中找出与匹配的码字 c_m :

$$\begin{cases} \text{colordist}(x_t, v_m) \leq \varepsilon_1 \\ \text{birghtness}(I, \langle I_m^{\min}, I_m^{\max} \rangle) = \text{true} \end{cases} \quad (2)$$

其中: ε_1 是全局阈值变量,需针对具体应用做合适调整。

① 如果找不到,即码书中没有码字满足上述条件,则利用式(1)为该像素创建一个码字;

② 如果有码字 c_m 满足上述条件,其内容是:

$$\begin{cases} v_m = (R_m, G_m, B_m) \\ u_m = \langle I_m^{\min}, I_m^{\max}, f_m, \lambda_m, p_m, q_m \rangle \end{cases} \quad (3)$$

那么将该码字做如下更新:

$$v_m = \left(\frac{f_m R_m + R_t}{f_m + 1}, \frac{f_m G_m + G_t}{f_m + 1}, \frac{f_m B_m + B_t}{f_m + 1} \right) \quad (4)$$

$$u_m = \langle \min\{I, I_m^{\min}\}, \max\{I, I_m^{\max}\}, f_m + 1, \max\{\lambda_m, t - q_m\}, p_m, t \rangle \quad (5)$$

第 3 步 训练结束后,计算该像素每个码书没有再次出现的最大时间间隔,以消除冗余码字,精炼出初始码书 $M(k)$ 是码字的索引:

$$\begin{cases} \lambda_i = \max\{\lambda_i, (N - q_i + p_i - 1)\} \\ M = \{c_k \mid c_k \in C, \lambda_k \leq T_M\} \end{cases} \quad (6)$$

其中:阈值 T_M 通常取训练帧数的一半,即 $N/2$,表示所有代表背景的码字必须至少在 $N/2$ 帧中出现。

1.2 Kalman 滤波器简介

Kalman 滤波^[10]是最优化自回归数据处理算法,它根据过去的信号,利用统计估计理论,使用线性最小均方误差作为最优准则,预测将来某个时刻的值。Kalman 滤波从与被提取信号有关的测量值中通过算法估计出所需的信号,它实际上是对随时间改变参数估计的一种最小二乘逼近,它考虑一个随时间变化的状态矢量,并通过一个观察模型测得组与参数相对应的不同时刻的参数值。假如能够建立参数的确定性和不确定性影响随时间变化的模型,Kalman 滤波就可以提供任何时刻对状态矢量进行估计的一套算法。

由于每帧时间间隔比较小,相邻每帧运动目标状态变化比较小,可以近似看作匀速运动,设每帧时间间隔为 T ,并设一个随机过程的系统运动方程和系统观测方程分别为:

$$\begin{cases} x(k) = \varphi(k, k-1)x(k-1) + \Gamma(k-1)w(k) \\ z(k) = c(k)x(k) + v(k) \end{cases} \quad (7)$$

其中: $x(k)$ 为系统状态真实值, $w(k)$ 为系统噪声矢量值。 $\varphi(k, k-1)$ 和 $\Gamma(k-1)$ 为状态转移矩阵和测量矩阵。则其滤波运算过程为:

$$\begin{cases} \hat{x}(k|k) = \varphi(k, k-1)\hat{x}(k-1|k-1) + K(k)v(k) \\ \hat{x}(k|k-1) = \varphi(k, k-1)\hat{x}(k-1|k-1) \end{cases} \quad (8)$$

$$\varphi(k, k-1) = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (9)$$

$$\Gamma(k-1) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad (10)$$

观测噪声方差为:

$$v(k) = Z(k) - C(k)x(k|k-1) \quad (11)$$

预测误差的协方差阵:

$$\begin{aligned} p(k|k-1) &= \varphi(k, k-1)p(k-1|k-1) \times \\ &\quad \varphi^T(k, k-1) + \\ &\quad \Gamma(k-1)Q(k-1)\Gamma^T(k-1) \end{aligned} \quad (12)$$

增益算法:

$$\begin{aligned} K(k) &= p(k|k-1) \times C^T(k) \times \\ &\quad [C(k)p(k|k-1)C^T(k) + R(k)]^{-1} \end{aligned} \quad (13)$$

滤波误差方差为:

$$p(k|k) = [I - K(k)C(k)]p(k|k-1) \quad (14)$$

其中: I 是单位矩阵, $R(k)$ 是 $v(k)$ 的方差阵, $Q(k-1)$ 是 $w(k-1)$ 的方差阵。

2 改进的码书背景建模算法

本文利用混合高斯模型的像素值分布规律和先验知识对码书模型的背景更新方法和目标决策机制进行了深入分析,提出了一种基于 Kalman 滤波器的改进码书背景建模算法,改进算法更能精确反映背景的更新程度,并能根据像素值的分布,重新排序码书中的码字,提高码字匹配速度。图 1 是该算法的系统流程。

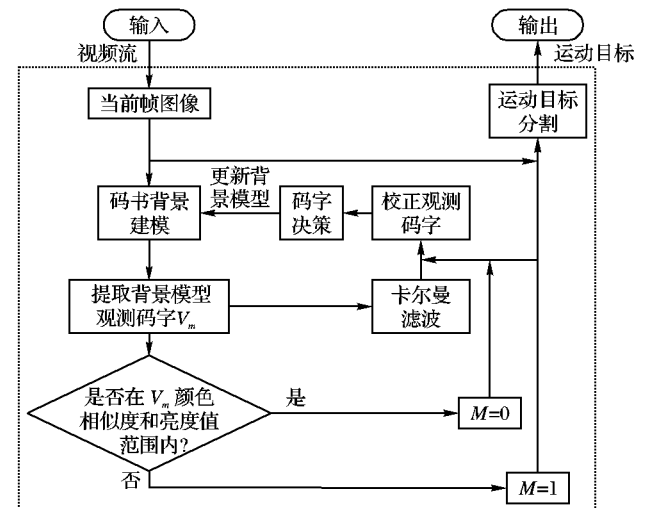


图 1 系统整体流程

2.1 背景更新策略

在进行码书背景建模时,传统的目标分割方法就是将时刻的输入像素值与码书背景模型相比较。然而,当这一时刻背景变化很大,码书背景模型就会把这一时刻的一部分背景也会误判为运动目标,从而使模型学习到的样本码字不准确,随时间推移陷入“恶性循环”,使以后背景区出现重噪声。本文利用文献[9]提出的根据 Kalman 滤波器的时域递归低通滤波特性,当图像序列通过这个 Kalman 时域低通滤波器时,图像序列的缓变部分可以从图像的快速变化过程中分离出来,将码书背景模型提取出来的码字作为 Kalman 滤波器的观测值,利用 Kalman 滤波器对观测值进行修正,从而得到背景模型中校正的码字。其背景模型更新方程表达为:

$$\begin{cases} v_m = (\bar{R}_m + g(t) \times (R_t - \bar{R}_m), \\ \quad \bar{G}_m + g(t) \times (G_t - \bar{G}_m), \bar{B}_m + g(t) \times (B_t - \bar{B}_m)) \\ g(t) = \beta \times (1 - M(t)) + \delta \times M(t) \end{cases} \quad (15)$$

其中: $g(t)$ 是适当的增益因子。 $M(t)$ 是当前像素点是否为运动目标的表示形式,当 $M(t) = 1$ 时,表示该点是运动目标像素点,否则为背景像素点。 β 为背景因子, δ 为运动因子,可以取一些合适的经验值。 β, δ 决定了背景提取过程的自适应特征,所以它们的取值应介于 0 和 1 之间。 δ 必须足够小,以便运动物体能够从背景序列中有效地分离出来。同时,为了使算法运行正常, β 的取值必须大于 δ 。典型的 β 取值要大于 δ 取值的一个数量级。如果 β 值选取得太小,系统会失去自适应特性。反之,如果 β 取值近似于 1,那么越来越多的快速变化会转化到图像背景上,而且这种算法减小噪声的特性就会失去。

2.2 码字决策

在实际运用中,得到精确码字背景模型之后,为探求某一时刻当前帧的每个像素是处于前景状态还是背景状态,从贝叶斯准则观点来看,这种探求是基于后验概率的。因此本文借鉴混合高斯混合模型中提到的像素值分布特点,引用权值思想和先验知识,精炼出初始码书中的码字后,按优先级 f_m 由大到小重新对码字进行排序,每隔 5 帧进行一次码字调整,从而加快之后的像素与码字匹配速度。通过对先前视频帧的训练,计算以 x_t 为背景的后验概率(式(16))。

$$P(x_t | c_m) = \frac{P(c_t | x_m)P(x_m)}{\sum_{i=1}^L P(c_i | x_t)P(x_i)} \quad (16)$$

2.3 算法描述

在训练背景,得到背景模型后,运用减背景运动目标分割思想,本文根据像素采样值与其码书中的各个码字是否匹配来判断。

输入:从视频帧中获取新输入像素 $x_t = (R_t, G_t, B_t)$ 及其对应码书 M 。

输出:分割运动目标像素点 $x_{(i,j)} = (R_{(i,j)}, G_{(i,j)}, B_{(i,j)})$ 。

1) 判断当前像素 $x_t = (R_t, G_t, B_t)$ 的亮度 I_t 范围,为了界定阴暗和高亮,运动目标分割中的亮度变化有一个范围,对于每个码字,其范围定义为 $[I_{low}, I_{hi}]$,并定义布尔变量 $matched = 0$ 。

$$\begin{cases} I_{low} = \gamma I^{max} \\ I_{hi} = \min\{\beta I^{max}, \frac{I^{min}}{\gamma}\} \end{cases} \quad (17)$$

其中: $\gamma < 1, \beta > 1$ 。定义 $brightness(I_t, \langle I^{min}, I^{max} \rangle)$ 如下:

$$brightness(I_t, \langle I^{min}, I^{max} \rangle) = \begin{cases} \text{true}, & I_{low} \leq I_t \leq I_{hi} \\ \text{false}, & \text{其他} \end{cases} \quad (18)$$

2) 计算当前像素 $x_t = (R_t, G_t, B_t)$ 、码字 c_i 和 $v_i = (R_i, G_i, B_i)$ 的颜色相似度,并给阈值变量 ε_2 赋值:

$$\begin{cases} \|x_t\|^2 = R^2 + G^2 + B^2 \\ \|v_i\|^2 = R_i^2 + G_i^2 + B_i^2 \\ \langle x_t, v_i \rangle^2 = (\bar{R}_i R + \bar{G}_i G + \bar{B}_i B)^2 \end{cases} \quad (19)$$

两者颜色相似度判断如下:

$$\begin{cases} p^2 = \|x_t\|^2 \cos^2 \theta = \\ \quad \|x_t\|^2 \cdot \frac{\langle x_t, v_i \rangle^2}{\|x_t\|^2 \cdot \|v_i\|^2} = \\ \quad \frac{\langle x_t, v_i \rangle^2}{\|v_i\|^2} \\ \text{colordist}(x_t, v_i) = \delta = \sqrt{\|x_t\|^2 - p^2} \end{cases} \quad (20)$$

3) 判断前景运动目标像素,根据式(19)~(20)两个条件从其码书 M 中找出与当前像素相匹配的码字 C_m ,如果找到,则 $matched = 1$:

$$\begin{cases} \text{colordist}(x_t, v_m) \leq \varepsilon_2 \\ \text{brightness}(I_t, \langle I^{min}, I^{max} \rangle) = \text{true} \end{cases} \quad (21)$$

$$BGS(x_{(i,j)}) = \begin{cases} x_{(i,j)} \text{ 为前景像素,} & matched = 0 \\ x_{(i,j)} \text{ 为背景像素,} & matched = 1 \end{cases} \quad (22)$$

其中:检测门限 ε_2 比 ε_1 稍大,并需要针对具体应用作适当调整。

3 试验与结果分析

为验证改进算法的有效性,本文采集了四段复杂室外场景的视频进行测试。测试视频的分辨率为 320×240 ,采集频率为 26 fps。实验硬件平台采用处理器英特尔 P4 1.8 GHz,内存为 1 GB 的台式机,软件测试环境为 VC++ 6.0。四段视频中不仅包含运动的目标还有复杂背景。另外,在视频一中存在摄像机的轻微抖动,视频二中由于光线经背景物体的漫反射而存在光线的扰动,视频三和视频四包含了大量不断摆动的树叶等杂乱背景(代表了动态变化场景的一种极限情形)。图 2~5 为本文算法与几个经典的背景建模算法对四个测试视频上的分割结果,所有的分割结果均未做手工处理。

可以看出,虽然利用 Kalman 滤波器的时域递归低通滤波特性对背景进行预测更新,可以减少一部分背景噪声,但是分割出的运动目标容易产生拖影现象,特别是当运动目标速度较慢时,拖影现象十分严重,造成目标严重变形;混合高斯背景建模虽然能够比较完整的分割出运动目标,但背景更新速率慢,在目标边缘出现大量的背景噪声;传统码书背景建模算

法虽然对背景具有一定的自适应性,但由于背景运动的不规则,依然在目标周末产生了大量的虚警,尤其是不能对于背景中像树叶、光线、摄像机抖动这样不规则的运动大面积的消除,使得对运动目标的分割不够准确,并且码书模型中的码字排序无章,减弱了之后的码字匹配速度,大大降低了算法的效率,其处理平均速度为 64.2 ms/f。

与上述方法相比,改进算法利用 Kalman 滤波器的时域递

归低通滤波特点,对码书背景建模进行了改进,能更好的自适应背景更新,极大克服了复杂背景中树叶、光线等干扰影响,更为准确的分割出运动目标的轮廓,可以有效地排除树叶、水波、摄像机抖动等这类假目标,虚警大大低于传统码书方法。并且在构建出背景模型后,每隔 5 帧重新调整一次码字序列,提高匹配速度,其处理平均速度为 57.9 ms/f 左右,提高了运行效率。

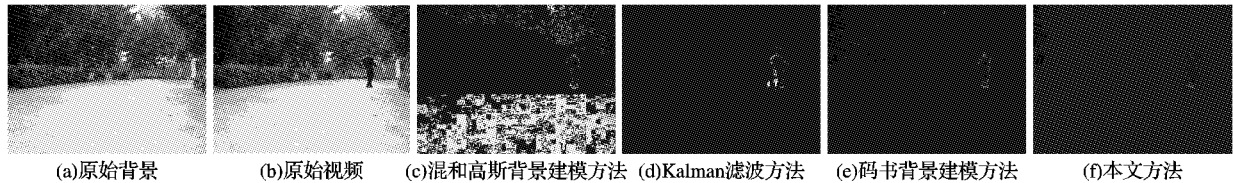


图 2 不同方法对视频一的分割

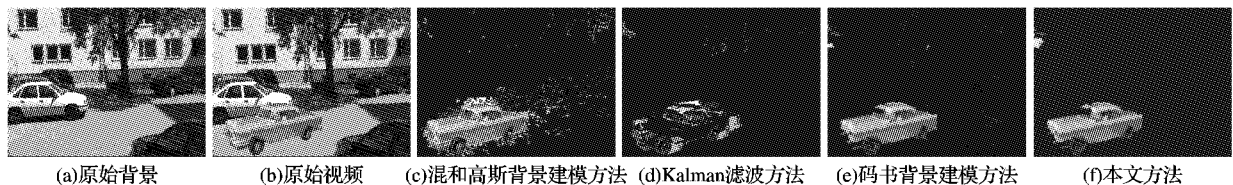


图 3 不同方法对视频二的分割

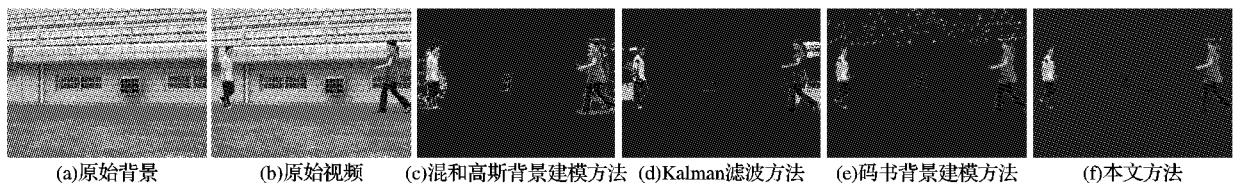


图 4 不同方法对视频三的分割

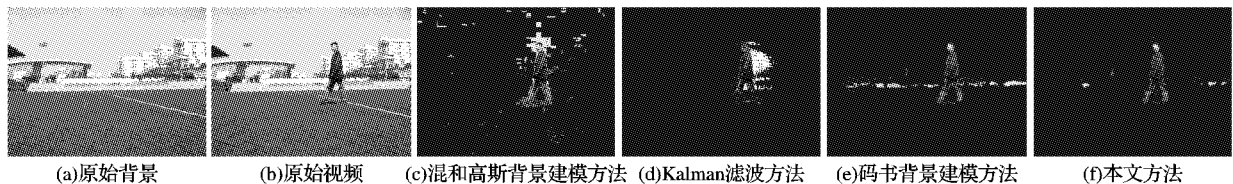


图 5 不同方法对视频四的分割

4 结语

本文利用了混过高斯模型中权值思想和先验知识,对码书的背景更新采用不同的更新率,提高了模型的学习准确程度,特别适合光线缓变或者树叶晃动等复杂场所的背景学习,通过初始统计学习,得出在每个像素点处各颜色值的一组码字,计算每一像素点处是否满足背景方程,最终判断该点是属于背景还是前景。经过试验测试表明,该方法快速、准确,有着广泛的实用性。

本算法中的 α 、 ε_1 、 ε_2 取决于各个复杂背景的变化,对错误率的影响较大,今后的工作将考虑到如何自动选择这 3 个参数。

参考文献:

- [1] 代科学,李国辉,涂丹. 监控视频运动目标检测减背景技术的研究现状和展望[J]. 中国图象图形学报,2006,11(7):22-30.
- [2] 万纛,韩毅,卢汉清. 运动目标检测算法的探讨[J]. 计算机仿真,2006,10(10):56-62.
- [3] CHALIDABHONGSE T H. A real-time system for detecting and tracking people in video [M]. New York: ACM Press, 2001.
- [4] STAUFFER C, GRIMON W. Adaptive background mixture models

- for real-time tracking [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE, 1999:246-252.
- [5] KAEWTRAKULPONG P, BOWDEN R. An improved adaptive background mixture model for Real-time tracking with shadow detection [C]// The 2nd European Workshop on Advanced Video-based Surveillance Systems. Kingston: Kluwer Academic Publishers, 2001:1-5.
- [6] KIM K, CHALIDABHONGSE T H, HARWOOD D, *et al.* Background modeling and subtraction by codebook construction [J]. ICIP '04: 2004 International Conference on Image Processing. Washington, DC: IEEE, 2004:3061-3064.
- [7] KIM K, CHALIDABHONGSE T H, HARWOOD D, *et al.* Real-time foreground-background segmentation using codebook model [J]. Real-Time Imaging, 2005,11(3):172-185.
- [8] LI YONG-BIN, CHEN FENG, XU WEN-LI, *et al.* Gaussian-based codebook model for video background subtraction [M]. Berlin: Springer-Verlag, 2006: 762-765.
- [9] 沈宇键,何昕,郝志航. 图像序列中检测运动小目标的递归算法[J]. 光电工程,2000,27(2):9-13.
- [10] 李彦鹏,黎湘. 一种快速 Kalman 滤波算法实现及效果评估[J]. 电子与信息学报,2005,27(1):153-154.