

面向视频的人物前景分割研究*

徐 宁¹, 王国胤^{1,2}, 龚 勋¹

(1. 西南交通大学 信息科学与技术学院, 成都 610031; 2. 重庆邮电大学 计算机科学与技术研究所, 重庆 400065)

摘 要: 针对目标分割中存在的背景噪声和阴影干扰问题, 提出一种改进的分割方法: 利用码书背景建模对视频进行分割后, 结合梯度因子、Sobel 算子对分割结果进一步约束。实验结果表明, 该方法可以有效地分割出前景目标, 抗干扰能力强。

关键词: 前景分割; 码书; 背景建模; 梯度因子

中图分类号: TP391

文献标志码: A

文章编号: 1001-3695(2012)04-1560-03

doi:10.3969/j.issn.1001-3695.2012.04.100

Foreground segmentation from videos

XU Ning¹, WANG Guo-yin^{1,2}, GONG Xun¹

(1. School of Information Science & Technology, Southwest Jiaotong University, Chengdu 610031, China; 2. Institute of Computer Science & Technology, Chongqing University of Posts & Telecommunications, Chongqing 400065, China)

Abstract: For the problems of background noise and shadow interference in the object segmentation, this paper proposed an improved target segmentation method. After using the codebook background model segmented the video, it constrained the segmentation results for further by combining the gradient factor and the Sobel operator. Experimental results show that the method can effectively segment the foreground targets and has strong anti-interference ability.

Key words: foreground segmentation; codebook; background modeling; gradient factor

随着计算机、通信和图像处理技术的不断发展, 视频监控研究越来越受到重视。运动目标分割作为整个视频监视系统的底层, 是后续各种高级应用的基础, 已经成为这一领域的研究热点。运动目标分割就是在图像序列中将变化区域从背景图像中分割出来。目前主要的前景算法有帧差法(temporal difference)、背景做差法(background subtraction)、光流法(optical flow)^[1,2]三类。

以上各方法往往是针对某一比较理想的场景进行的, 常事先假定在一个固定镜头下场景背景中物体是绝对静态的。然而, 在复杂场景目标分割研究中, 场景往往是非理想化的, 存在光线变化、摄像机抖动、背景扰动、运动阴影及遮挡等因素, 这给运动目标的精确分割带来困难。因此现有目标分割算法在精度和处理速度上有待进一步提高。针对上述几种分割方法, 目前存在一些典型的改进方法: 周西汉等人^[3]利用对称差分法来提取目标的运动轮廓, 降低了帧差法中时隙选择的重要性; 为了增强背景做差法对复杂场景的鲁棒性, 赵晓雄等人^[4]将混合高斯模型与背景图像的相关性运算相结合来提高背景的适应性; 同时, 也出现了将帧差法和背景做差法融合的方法来进行运动目标检测和自适应背景更新^[5,6], 并取得了很好的效果, 但该方法还是存在着目标轮廓分割的不完整性等问题; Elgammal 等人则提出了一种非参数背景模型^[7]实时运动目标检测的背景做差法, 它利用统计学习的模式进行背景更新, 很

好地提取了前景目标, 但基于非参数的运动目标分割算法仅仅依靠序列图像之间的运动信息, 分割的结果对噪声比较敏感, 而且精度受限于遮挡和孔径问题, 得不到准确的目标边界。针对以背景建模为基础的目标分割研究, 最近比较有代表性且效果较好的有: Heikkilä 等人利用基于局部二元模式(LBP)的纹理信息进行背景块建模^[8], 在解决孔径问题的同时很好地检测出目标, 但在边缘处存在较多背景像素; 而基于块的背景建模中, 由于它是一种基于区域块的方法, 使得最后检测得到的目标在边缘处包含很多属于背景的像素; Kim 等人^[9,10]提出的码书模型, 根据颜色相似度及其亮度范围对每个像素点的连续采样值进行码书背景建模, 对待检测帧的每个像素与模型中的每个码书字进行匹配, 算法复杂度小, 占内存少, 取得了较好的检测效果, 但当背景发生剧烈变化时, 背景模型更新速度较慢。

为提高分割的精度并有效抑制背景噪声和运动阴影干扰, 本文结合梯度因子的性质, 对基于码书背景建模的分割结果进行进一步约束, 并有效地分割出了运动阴影, 较好地降低了误检率。

1 相关技术基础介绍

1.1 码书背景建模

码书背景建模算法^[9,10]采用一种量化/聚类的标准, 对一

收稿日期: 2011-07-23; **修回日期:** 2011-09-10 **基金项目:** 国家自然科学基金资助项目(61073146); 重庆市杰出青年科学基金资助项目(2008BA2041)

作者简介: 徐宁(1987-), 男, 硕士研究生, 主要研究方向为模式识别、智能视频系统分析、智能信息处理等(xn0803@163.com); 王国胤(1970-), 男, 教授, 博导, 主要研究方向为智能信息处理、粗糙集、粒计算、网络安全; 龚勋(1980-), 男, 讲师, 硕导, 博士, 主要研究方向为图像处理、模式识别与人工智能。

系列图像序列帧的每个相应像素点进行连续采样,并根据颜色相似度及其亮度范围为采样值生成背景模型。对于建好的模型,每个像素点码书由一个或多个码字组成。

设 $\chi = \{x_1, x_2, \dots, x_N\}$ 是一个像素的序列采样值,其中 $x_t (t=1, \dots, N)$ 是 RGB 向量,并设 $C = \{c_1, c_2, \dots, c_N\}$ 是该像素的码书,每个码书 $c_i (i=1, 2, \dots, L)$ 定义为二元组结构: $v_i = (\overline{R_i}, \overline{G_i}, \overline{B_i})$, $u_i = \langle I_i^{\min}, I_i^{\max}, f_i, \lambda_i, p_i, q_i \rangle$, 其颜色模型和相似度判断如图1所示。

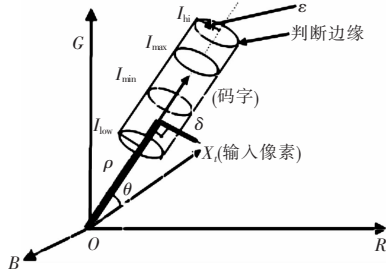


图1 码书颜色模型

主要步骤如下:

a) 置空每个像素的码书,即 $L=0$ 。

b) 对于训练视频的每个像素的序列值

$$\chi = \{x_1, x_2, \dots, x_N\}, x_t = (R_t, G_t, B_t), t=1, \dots, N$$

如码书为空, $L=0$, 则创建一个码字:

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

否则根据亮度和颜色两个约束条件寻找与 x_t 匹配的码字 c_m : 若找不到, 则利用式(1)为该像素创建一个码字; 否则对该码字进行更新。

c) 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, 即对于 $c_i (i=1, \dots, L)$:

$$\lambda_i = \max \{ \lambda_i, (N - q_i + p_i - 1) \} \quad (2)$$

d) 然后利用 λ 消除冗余码字, 得到最能代表真实背景的初始码书 M (M 是码字的索引):

$$M = \{c_k | c_k \in C, \lambda_k \leq T_M\} \quad (3)$$

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。

码书背景模型在背景建模与运动目标分割中取得了较好的效果。但该方法也存在着局限性, 由于码书背景模型是以聚类和量化为标准, 目标中那些与背景颜色或亮度相似的像素点可能会被检测为背景, 造成孔径问题; 同时在复杂多变的场景中, 如有微小物体的运动干扰时不能得到很好的效果, 且该方法在消除运动阴影干扰方面效果还不理想。所以, 如何进一步提高目标分割的精度, 并有效消除背景噪声和运动阴影的影响, 及在保证目标分割精度的前提下提高算法的效率等几个问题就显得尤为重要。

1.2 梯度算子简介

研究证明对于阴影内部每个像素点来说, 其灰度值是背景图在非阴影区域对应点灰度值的 K 倍, 且在整个阴影区内, K 值具有较小的波动, 相当于低频区; 而目标区域的灰度值相对于背景灰度的比值起伏较大, 相当于高频区^[11]。因此可以通过梯度算子区分阴影和目标。

索贝尔算子 (Sobel operator)^[12] 是一种常用的且效果较好的梯度算子, 用来计算图像亮度函数的梯度近似值。在图像的

任何一点使用此算子, 将会产生对应的梯度矢量或法矢量。

该算子包含两组 3×3 的矩阵, 分别为横向及纵向, 将其与图像作平面卷积, 即可分别得出横向及纵向的亮度差分近似值。如以 A 代表原始图像, G_x 及 G_y 分别代表经横向及纵向边缘检测的图像, 其公式如下:

$$G_x = \begin{bmatrix} -1 & 0 & +1 \\ -2 & 0 & +2 \\ -1 & 0 & +1 \end{bmatrix} * A, G_y = \begin{bmatrix} +1 & +2 & +1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A$$

图像每一个像素的横向及纵向梯度近似值可用式(4)结合, 来计算梯度的大小:

$$G = \sqrt{G_x^2 + G_y^2} \quad (4)$$

然后可用式(5)计算梯度方向:

$$\Theta = \arctan(G_x / G_y) \quad (5)$$

如果以上的角度 Θ 等于零, 即代表图像在该处拥有纵向边缘, 左方较右方暗。

2 改进的目标分割算法

由于码书背景建模是根据像素点的颜色相似度与亮度范围生成的背景模型, 而在实际应用中, 目标像素点往往存在与背景像素点灰度相近的部分。与此同时, 由于运动阴影的自身特性^[11], 使得运动阴影与运动目标经常出现误判的情况, 这都对基于码书的分割效果产生着影响。为此, 要进一步提高前景目标的分割精度, 就要更有效地区别背景和阴影, 以减少对目标像素点与阴影像素点的误判, 并提高对背景噪声的抑制作用。本文采用如下改进分割方法:

先用码书建模思想对视频帧进行初步运动目标分割, 得到背景帧差图 S , 图中一般会包括运动目标部分、运动阴影部分及一些背景噪声; 用背景图像 B 除以帧差图 S , 得到输出图像 D 。其中: 背景与阴影点相除所得的值在一个很小的范围变化, 且比较平坦, 是一个低频区域, 而背景与运动目标像素相除的结果则不一定, 是一个高频区域^[13]; 然后用 Sobel 算子计算出图 D 对应的梯度图 G , 并检测出阴影和背景之间因光学增益不同而产生的边缘, 这就可以将运动目标和阴影进一步区分开来; 最后根据一定的阈值 T_n 进行分割, 就可以将阴影区域分割出来从而去除, 进而得到更加精确的前景目标, 并减少了对灰度与背景接近的目标像素的漏识。

具体算法描述如下:

输入: 从视频中获取输入像素 $x_t = (R_t, G_t, B_t)$ 。

输出: 分割出的运动目标像素点。

a) 利用码书建模思想得到每个像素 $x_t = (R_t, G_t, B_t)$ 所对应的码书 c_i 。每个码书 $c_i (i=1, \dots, L)$ 定义为一个二元组结构:

$$\langle v_i, u_i \rangle。其中, v_i = (\overline{R_i}, \overline{G_i}, \overline{B_i}) \\ u_i = \langle I_i^{\min}, I_i^{\max}, f_i, \lambda_i, p_i, q_i \rangle \quad (6)$$

然后, 根据亮度范围及颜色相似度这两个条件

$$\text{亮度约束: } \text{brightness}(I, \langle I_i^{\min}, I_i^{\max} \rangle) = \text{true}$$

$$\text{颜色约束: } \text{colordist}(x_t, v_m) \leq \varepsilon_1 \quad (7)$$

其中: $\text{colordist}(x_t, v_t) = \cos \theta = \frac{\langle x_t, v_t \rangle}{\sqrt{\|x_t\|^2} \cdot \sqrt{\|v_t\|^2}}$

进而利用背景码字 c_m 构造出背景图 B 。

b) 通过背景做差法初步分割出运动目标:

$$d = |V(i, j, n) - B(i, j)|$$

$$S(i, j, n) = \begin{cases} d & d \geq T \\ 0 & d < T \end{cases} \quad (8)$$

其中: V 是当前视频帧; B 是背景亮度分量; n 表示帧数; T 为阈值; S 是背景帧差图。此步分割结果一般包括运动目标部分、运动阴影部分以及一些背景噪声。

c) 背景图像 B 除以图像 S , 得到输出图像 D :

$$D(i, j) = \frac{B(i, j)}{S(i, j)} \quad (9)$$

并利用 Sobel 算子计算图 D 的梯度值:

$$\begin{aligned} x \text{ 方向: } G_x &= |G_x[D(i, j)]| \\ y \text{ 方向: } G_y &= |G_y[D(i, j)]| \\ \text{Grad} &= \sqrt{G_x^2 + G_y^2} \end{aligned} \quad (10)$$

进而得到梯度图 G 。

d) 根据梯度图 G 检测出阴影和背景之间因光学增益不同而产生的边缘, 再根据一定的阈值进行分割, 从而将运动目标和阴影区别开来。

根据 D 的取值将梯度图 G 分成若干部分, 记为 $P_n (n=1, 2, \dots, n)$, 对这 n 个部分分别求出局部阈值 T_n [13]:

$$T_n = \frac{\sum_{i, j \in P_n} G(i, j)}{\sum_{i, j \in P_n} H(i, j)} \quad (11)$$

其中: $\sum_{i, j \in P_n} G(i, j)$ 为该区域总梯度值, $\sum_{i, j \in P_n} H(i, j)$ 为该区域非零梯度像素的个数。然后, 以阈值 T_n 分割梯度图 G , 得到二值目标前景图像 M 。

e) 将得到的二值化前景图像还原成原始前景目标彩色图像。

3 实验及分析

为了验证本文算法的有效性, 采用文献[14]中所使用的视频进行实验。该视频分别针对室内室外的三个不同场景进行采集, 背景复杂, 存在光线扰动与漫反射及运动阴影干扰, 能很好地对本文方法的分割效果进行检测。其中, 视频的分辨率为 320×240 像素, 真彩格式, 采集频率为 27 fps。实验平台为 Intel Core Duo T2300 (1.66 GHz) 处理器, 内存 768 MB。

实验中使用该视频对本文算法进行测试, 并将检测结果与混合高斯背景建模和码书背景建模方法进行对比。

实验 1 室内环境

该场景是在室内环境录制的一段视频数据, 实验结果如图 2 所示。

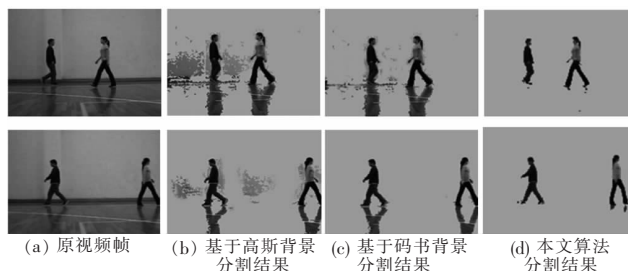


图 2 室内场景分割结果对比

图 2(a) 是原始视频帧, 场景中包含了运动阴影及场景物体漫反射所导致的光线扰动; (b) 是用混合高斯背景建模方法实现的分割效果, 由于光线的扰动, 产生了大块的背景噪声, 而且阴影区域分割效果差; (c) 是用码书背景建模的方法实现的分割效果, 由于码书背景模型是以聚类 and 量化为标准, 目标中那些与背景颜色或亮度相似的像素点会被检测为背景, 而且大

块的运动阴影被误检了出来; (d) 是本文改进方法的检测结果, 其有效地分割了运动阴影部分, 而且减少了对背景像素的误识, 得到了较好的效果。

实验 2 室外环境 1

该场景是在室外环境录制的一段视频数据, 实验结果如图 3 所示。

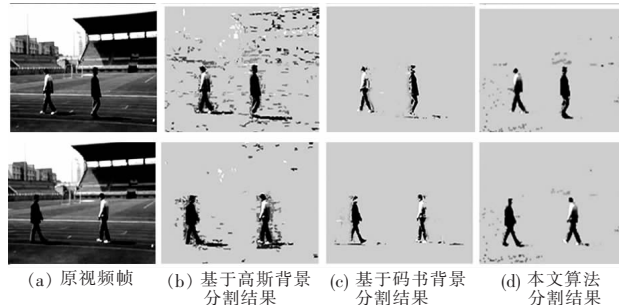


图 3 室外场景 1 分割结果对比

图 3(a) 是原始视频帧, 场景中存在光线变化、背景扰动及摄像机晃动等因素, 而且背景比较复杂; (b) 是用混合高斯背景建模方法实现的分割效果, 由于机器的抖动和背景物体的漫反射, 在背景和运动目标周围产生了较多的噪声点, 而且对阴影的分割效果差; (c) 是用码书背景建模方法实现的分割效果, 其不能有效消除因背景扰动和漫反射而产生的大范围背景噪声, 而且没有很好地分割出运动阴影区域; (d) 是本文改进方法的检测结果, 虽然没有彻底地消除背景扰动的影响, 但从检测结果上看已经有了明显的改进, 而且它有效地分割了运动阴影部分。

实验 3 室外环境 2

该场景是在室外环境录制的一段视频数据, 实验结果如图 4 所示。

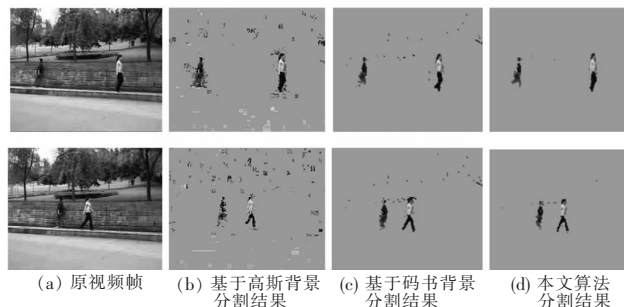


图 4 室外场景 2 分割结果对比

图 4(a) 是原始视频帧, 场景中光线不太稳定, 存在树叶晃动和摄像机抖动等问题, 背景比较复杂; (b) 是用混合高斯背景建模方法实现的分割效果, 由于背景中树叶晃动和机器抖动, 产生了很多误检点, 分割效果差; (c) 是用码书背景建模方法实现的分割效果, 它没有很好地消除背景噪声, 且前景像素丢失严重; (d) 是本文改进方法的检测结果, 虽然没有完全抑制前景像素的丢失, 但在消除背景噪声方面较其他两种方法已经有了明显的改进。

表 1 为本文算法与其他两种经典算法在处理速度上的对比, 其中, 原视频帧为 27 fps。本文方法虽然在检测速度上比其他两种分割方法稍慢, 但在很大程度上克服了复杂背景光线扰动和背景噪声的干扰, 而且较好地实现了运动阴影的分割。此外, 处理速度上的差距可以通过提高硬件条件来进一步缩小。总体而言, 该方法以较小的时间代价换来运动目标检测精度的提高。

(下转第 1587 页)

果(以 ISNR 和 MSE 等客观标准衡量)。在图像复原的速度方面,合成模式下本文算法的速度最快;但在分析模式下,由于该算法用 PCG 方法计算式(15),其速度受到了影响。因此,提高分析模式下该算法的速度是未来进一步研究的方向。

参考文献:

- [1] ELAD M, MILANFAR P, RUBINSTEIN R. Analysis versus synthesis in signal priors [J]. *Inverse Problems*, 2007, 23(3): 947-968.
- [2] DONOHO D L. Compressed sensing [J]. *IEEE Trans on Information Theory*, 2006, 52(4): 1289-1306.
- [3] FIGUEIREDO M A T, NOWAK R D. A bound optimization approach to wavelet-based image deconvolution [C]//Proc of the IEEE International Conference on Image Processing. 2005:782-785.
- [4] CHOUZENOUX E, IDIER J, MOUSSAOUI S. A majorize-minimize strategy for subspace optimization applied to image restoration [J]. *IEEE Trans on Image Processing*, 2011, 20(6): 1517-1528.
- [5] RUGGIERO V, SERAFINI T, ZANELLA R, et al. Iterative regularization algorithms for constrained image deblurring on graphics processors [J]. *Journal of Global Optimization*, 2010, 48(1): 145-157.
- [6] DUPE F X, FADILI J M, STARCK J L. A proximal iteration for deconvolving Poisson noisy images using sparse representations [J]. *IEEE Trans on Image Processing*, 2009, 18(2): 310-321.
- [7] COMBETTES P L, WAJS V R. Signal recovery by proximal forward-backward splitting [J]. *SIAM Journal on Multiscale Modeling and Simulation*, 2005, 4(4): 1168-1200.
- [8] NESTEROV Y. Smooth minimization of non-smooth functions [J]. *Mathematical Programming*, 2005, 103(1): 127-152.
- [9] BECK A, TEOULLE M. A fast iterative shrinkage- thresholding algorithm for linear inverse problems [J]. *SIAM Journal on Imaging Sciences*, 2009, 2(1): 183-202.
- [10] BECKER S, BOBIN J, CANDÈS E J. NESTA: a fast and accurate

- first-order method for sparse recovery [J]. *SIAM Journal on Imaging Sciences*, 2011, 4(1): 1-39.
- [11] CHAMBOLLE A. An algorithm for total variation minimization and applications [J]. *Journal of Mathematical Imaging and Vision*, 2004, 20(1-2): 89-97.
- [12] Van Den BERG E, FRIEDLANDER M P. Probing the Pareto frontier for basis pursuit solutions [J]. *SIAM Journal on Scientific Computing*, 2008, 31(2): 890-912.
- [13] GOLDSTEIN T, OSHER S. The split Bregman method for L1-regularized problems [J]. *SIAM Journal on Imaging Sciences*, 2009, 2(2): 323-343.
- [14] SETZER S, STEIDL G, TEUBER T. Deblurring Poissonian images by split Bregman techniques [J]. *Journal of Visual Communication and Image Representation*, 2010, 21(3): 193-199.
- [15] KONTOGIORGIS S, LEONE R, MEYER R R. Alternating direction splitting for block angular parallel optimization [J]. *Journal of Optimization Theory and Applications*, 1996, 90(1): 1-29.
- [16] STEIDL G, TEUBER T. Removing multiplicative noise by Douglas-Rachford splitting methods [J]. *Journal of Mathematical Imaging and Vision*, 2010, 36(2): 168-184.
- [17] NOCEDAL J, WRIGHT S J. Numerical optimization [M]. 2nd ed. New York: Springer-Verlag, 2006.
- [18] NIESEN U, SHAH D, WORNELL G W. Adaptive alternating minimization algorithms [J]. *IEEE Trans on Information Theory*, 2009, 55(3): 1423-1429.
- [19] WOLKE R, SCHWETLICK H. Iteratively reweighted least squares: algorithms, convergence analysis, and numerical comparisons [J]. *SIAM Journal on Scientific Computing*, 1988, 9(5): 907-921.
- [20] ECKSTEIN J, BERTSEKAS D P. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators [J]. *Mathematical Programming*, 1992, 55(3): 293-318.

(上接第 1562 页)

表1 处理速度对比

场景	方法			fps
	基于高斯背景的目标分割	基于码书背景的目标分割	本文方法	
室内	23.7	21.6	19.4	
室外1	24.6	21.1	19.3	
室外2	23.2	22.3	18.9	

4 结束语

本文利用了码书背景建模的分割方法,并结合梯度因子的思想,得到一种改进的运动目标分割算法。经过实验测试表明,该方法以较小的时间代价,换来运动目标分割精度的提高,对背景噪声和光线扰动有较好的抑制作用,而且能够较好地分割运动阴影。

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] 周西汉,刘勃,周荷琴. 一种基于对称差分背景消除的运动检测方法[J]. *计算机仿真*, 2005, 22(4): 117-119.
- [4] 赵晓雄,张学东. 基于自适应学习速率混合高斯模型背景消除[J]. *鞍山科技大学学报*, 2005, 28(3): 239-242.
- [5] HENRY S, TAKEO K. Object detection using the statistics of parts [J]. *International Journal of Computer Vision*, 2004, 56(3): 151-

177.

- [6] KIM I B, KIM H J. Efficient region-based motion segmentation for a video monitoring system [J]. *Pattern Recognition Letters*, 2003, 24(3): 113-128.
- [7] 王成儒,孟凤. 快速背景重建的在线运动目标检测[J]. *光电工程*, 2007, 34(6): 112-115.
- [8] KIM K, HARWOOD D, DAVIS L, et al. Real-time foreground-background segmentation using codebook model [J]. *Real-Time Imaging*, 2005, 11(3): 172-185.
- [9] KIM K, CHALIDABHOUGSE T H, HARWOOD D, et al. Background modeling and subtraction by codebook construction [C]//Proc of Conference on Image Processing. 2004: 3061-3064.
- [10] LI Yong-bin, CHEN Feng, XU Wei-li, et al. Gaussian-based codebook model for video background subtraction [C]//Proc of the 2nd International Conference on Advances in Natural Computation. Berlin: Springer-Verlag, 2006: 762-765.
- [11] ROSIN P L, TIM E. Image difference threshold strategies and shadow detection [C]//Proc of British Conference on Machine Vision. Surrey: BMVA Press, 1995: 347-356.
- [12] KANOPOULOS N, VASANTHAVADA N, BAKER R L, et al. Design of an image edge detection filter using the Sobel operator [J]. *IEEE Journal of Solid-State Circuits*, 1988, 23(2): 358-367.
- [13] 李忠武,高广珠,余理富,等. 图像序列目标检测中阴影的消除[J]. *计算机应用研究*, 2004, 21(5): 205-206.
- [14] 刘文静,李红波,吴渝. 基于颜色特征的多目标瞬间重合识别和跟踪[J]. *计算机科学*, 2010, 37(10A): 164-167.