

中图分类号: TP391 文献标志码: A 文章编号: 1006-8961(2011)09-1716-06

论文索引信息: 蔡念, 陈世文, 郭文婷, 潘晴. 融合高斯混合模型和小波变换的运动目标检测 [J]. 中国图象图形学报, 2011, 16(9): 1716-1721

融合高斯混合模型和小波变换的运动目标检测

蔡念¹⁾, 陈世文^{1),2)}, 郭文婷¹⁾, 潘晴¹⁾

¹⁾ (广东工业大学信息工程学院, 广州 510006) ²⁾ (嘉应学院, 梅州 514015)

摘要: 当前景目标与背景在颜色上接近时, 仅采用高斯混合模型进行目标检测容易导致误判。为了提高模型分割算法的鲁棒性, 提出一种融合高斯混合模型和小波变换的运动目标检测算法。通过小波变换提取图像的纹理特征信息, 利用高斯混合模型拟合背景信息。将两者融合起来, 把纹理信息作为颜色信息的补偿, 保证了模型在线更新背景信息时模型的稳定性和收敛性, 同时弥补了目标分割中前景与背景颜色信息接近时容易导致误判的不足。实验结果表明, 本文方法比经典高斯混合模型方法具有较高的分割精度。

关键词: 高斯混合模型; 小波变换; 目标检测

Moving object detection using Gaussian mixture model and wavelet transform

Cai Nian¹⁾, Chen Shiwen^{1),2)}, Guo Wenting¹⁾, Pan Qing¹⁾

¹⁾ (School of Information Engineering, Guangdong University of Technology, Guangzhou 510006 China)

²⁾ (Jiaying University, Meizhou 514015 China)

Abstract: Commission errors often arise when Gaussian mixture models (GMMs) are applied to detect moving objects in situations where foreground and background have similar colors. To improve the robustness of the segmentation method, a moving object detection method is proposed by combining a Gaussian mixture model with the wavelet transform. The wavelet transform is employed to extract texture information and a GMM is employed to update the background. Color information and texture information are integrated for segmentation. The method improves the abilities of convergence and stability, and also decreases commission errors occurring in methods which only use color information. Experimental results indicate that the proposed method is superior to the traditional GMM.

Keywords: Gaussian mixture model; wavelet transform; object detection

0 引言

高斯混合模型(GMM)是一种半参数的估计方法, 通过新的学习分布替代旧的高斯分布, 能很好地表征背景变化, 是研究运动目标检测的最有效方法之一。Suo 和 Wang 在高斯混合模型基础上提出一种更新参数方法, 有效地解决了慢运动目标的检测,

但背景是复杂环境时效果却不理想^[1]。Power 和 Schoonees 通过研究高斯混合模型公式及参数更新公式, 提出新的学习率公式, 该学习率使均值收敛快而准确, 但用于方差学习却不够理想^[2]。陈世文等人提出一种加权思想, 在模型更新中赋予均值和方差不同的权值, 但仅利用颜色信息进行前景分割^[3]。Harville 等人结合图像的空间深度信息和色彩亮度信息, 提出一种改进的方法, 该方法较好地检

收稿日期: 2011-01-05; 修回日期: 2011-04-19

基金项目: 国家自然科学基金项目(61001179); 广东省自然科学基金项目(07301038, 9451009001002667)。

第一作者简介: 蔡念(1976—), 男, 副研究员, 硕士生导师。2004年于中国科学院生物物理所获图像处理与模式识别专业博士学位, 主要研究方向为计算机视觉、图像处理、模式识别等。E-mail: cainian@gdut.edu.cn。

测出多目标的运动,但是当不同的目标处在同一像素深度时,尽管有不同的色彩信息,也难检测出运动目标^[4]。原春锋等人提出了基于高斯混合模型和梯度信息的视频分割方法,该方法中结合梯度互相关函数对分割结果进行校正,能适应剧烈光照的变化;但模型收敛速度慢,实时性较差^[5]。刘鑫等人为每个像素构建高斯混合模型,通过融入帧间差分把每帧的图像区分为背景区域,背景显露区域和物体运动区域,能迅速响应实际场景的变化;但不能解决动态背景变化的问题^[6]。

人们通常通过颜色或者亮度信息来分割前景和背景,而颜色和亮度容易受光照及噪声的影响,容易导致误判。一幅图像包含了丰富的信息,除了颜色和亮度外,还包含有轮廓、边缘、纹理等信息,这些特征信息反映了空间的区域结构信息。当目标像素与背景像素在颜色和亮度信息上非常接近时,目标容易被误判为背景,利用纹理信息可以解决这一问题^[7]。文献[8]采用颜色信息与边缘信息对目标进行分割,取得了一定的效果,但当目标边缘不明显时也容易误判。文献[9]用高斯分布去模拟像素在时间轴上的颜色信息,用马尔可夫随机域去模拟纹理信息,然后用贝叶斯推断将两者信息结合起来,但是马尔可夫随机域的参数估计需要耗费较多的时间。本文通过小波变换提取图像的纹理特征信息,利用高斯混合模型拟合背景信息,将两者融合起来,把纹理信息作为颜色信息的补偿;保证了模型在线更新背景信息时模型的稳定性和收敛性,同时弥补了目标分割中前景与背景颜色信息接近时容易导致误判的不足。

1 高斯混合模型

对于随机变量 x , 其概率密度为

$$p(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (1)$$

式中, μ 为高斯分布的期望, σ 为高斯分布的标准差。如果一组数据符合高斯分布, 那么这种数据中的大多数会集中在以 μ 为中心的 -2σ 到 2σ 范围内。

在实际中, 尤其是对于室外的场景, 由于各种噪声的干扰, 其背景变化是比较大的, 会出现树枝摇动、光线变化等不同情况。因此用一个高斯分布就不能完全描述实际的背景情况。这种情况下就需要

采用多个高斯模型来描述动态的背景, 为不同的状态建立不同的高斯模型。

将序列图像中的某一点 (i, j) , 在时间 t 的观察值写为 X_t , 对于给定点 (i, j) 的一系列观察值 $\{X_1, X_2, \dots, X_t\}$, 可以看做是一个与其他点独立的统计随机过程, 用 K 个高斯分布的混合模型去模拟, 在当前时刻 t 点 (i, j) 的概率分布估计为

$$P(X_t) = \sum_{k=1}^K \omega_{k,t} \cdot N(X_t, \mu_{k,t}, \Sigma_{k,t}) \quad (2)$$

式中, $\omega_{k,t}$ 表示 t 时刻第 k 个混合高斯分布的权值, N 表示高斯概率密度函数, $\mu_{k,t}$ 表示第 k 个分布的均值, $\Sigma_{k,t}$ 表示第 k 个分布的方差, 则

$$N(X_t, \mu_t, \Sigma_t) = \frac{1}{(2\pi)^{\frac{n}{2}} |\Sigma_t|^{\frac{1}{2}}} e^{-\frac{1}{2}(X_t - \mu_t)^T \Sigma_t^{-1} (X_t - \mu_t)} \quad (3)$$

k 的选取一般由计算能力和内存来决定, 通常是 3 至 5, 因为虽然更大的 k 值能体现更复杂的场景, 但同时消耗更多的计算时间。

假设像素点的颜色是用 RGB 表示的, 用协方差表示的时候各个分量的方差应该是有差别、相互独立的, 但是为了计算的方便假设协方差矩阵表示为

$$\Sigma_{k,t} = \sigma_k^2 \mathbf{I} \quad (4)$$

式中, \mathbf{I} 为 3×3 的单位矩阵。

实际当中式(4)这种假设是不成立的, 虽然采用这种假设牺牲了一定的准确性, 但是却避免了耗时的矩阵求逆运算。这样, 当前场景中每个像素的观测值就可以用一个高斯混合模型表示。一般地, 一个新像素的值将用混合模型中的主要部分表示, 然后更新模型。

2 融合混合高斯模型和小波变换的目标检测

高斯混合模型能很好地拟合背景颜色信息的变化, 小波变换能较好地提取图像的纹理特征, 而纹理特征反映了图像的区域局部信息, 通常区域空间结构信息可以消除噪声及亮度变化影响。因此, 应用 Mallat 小波提取图像纹理信息, 采用 Li 等人^[10]提出的方法, 即将亮度信息与纹理信息线性组合起来, 进行运动目标检测, 即

$$d_{ii}(p) = \omega_i(p) \cdot d_i(p) + \omega_i(p) \cdot d_i(p) \quad (5)$$

$$d_i(p) = 1 - \frac{1}{\sqrt{2\pi}T} e^{-\frac{\|X_i - \bar{x}\|^2}{\sigma^2 T^2}} \quad (6)$$

式中, \bar{x} 为高斯混合模型中最佳匹配者的均值; T 为阈值, 通过亮度间的差值进行估计; $d_{ii}(p)$ 是背景图像与当前帧图像在相同像素位置 p 处的总体差距; $d_i(p)$ 为颜色信息差异; $d_t(p)$ 是纹理信息差异; $\omega_i(p)$ 为颜色信息贡献权值; $\omega_t(p)$ 为纹理贡献权值。 $\omega_i(p)$ 可根据 $\omega_t(p)$ 计算, 即 $\omega_i(p) = 1 - \omega_t(p)$ 。

如果直接通过小波变换来计算纹理权值 $\omega_t(p)$, 比较困难, 因为不同的纹理具有不同的维数和物理意义。先通过高斯混合模型估计一个 $\omega_i(p)$, 然后计算 $\omega_t(p)$ 。式(5)(6)的各参数通过如下方法来计算: 高斯混合模型更新参数, 计算当前观测值 X_i 与高斯分布的匹配情况, 将最佳匹配者的均值 \bar{x} 定义为更新后的背景颜色, 将对应的标准差 $\sigma_{i,1}$ 定义为阈值 T ; 同时将对应的权值定义为 $\omega_i(p)$ 。这样好处是: 实时更新了背景; 不同的像素采用了不同的阈值; 通过高斯模型间接计算 $\omega_t(p)$, 减少了计算量。

经小波分解后纹理特征信号得到增强, 区域结构信息被提取; 设当前像素 p 处的纹理特征为 $F_i(p)$, 从背景图像获得的纹理特征为 $F_b(p)$, N_p 为以 p 为中心的邻域, p 的纹理差异为

$$d_i(p) = 1 - \frac{2 \sum_{u \in N_p} C_{ib}(u)}{\sum_{u \in N_p} (C_u(u) + C_{ib}(u))} \quad (7)$$

$$C_{ib}(u) = F_i(p) \cdot F_b(p) \quad (8)$$

$$C_u(u) = F_i(p) \cdot F_i(p) \quad (9)$$

$$C_{bb}(u) = F_b(p) \cdot F_b(p) \quad (10)$$

这里 $N_p = 3 \times 3$ 。上式的相似性是通过角相似度来计算的, 纹理特征向量模的大小与角相似度的特点无关, 分析得知, $d_i(p)$ 不影响亮度缩放。

以上计算得到 $d_{ii}(p)$ 后, 根据下式进行目标分割:

$$S(p) = \begin{cases} 1 & d_{ii}(p) > \beta \\ 0 & d_{ii}(p) \leq \beta \end{cases} \quad (11)$$

式中 β 为分割阈值, 一般取 0.5, $\beta \in [0, 1]$ 。如果 $S(p)$ 为 0, 则对应像素被判别为背景; $S(p)$ 为 1 时, 对应像素被判别为前景目标。

目标分割步骤如下:

1) 将前几帧的颜色平均值作为高斯模型的均值, 方差初始化为一个较大的值。用小波变换计算初始背景图像的纹理特征向量。

2) 读入下一帧, 采用加权思想更新高斯参数^[3], 计算 $d_i(p)$, $\omega_i(p)$ 和 $\omega_t(p)$ 。

3) 用小波变换提取纹理特征向量, 用式(7)计算每个像素点的纹理差异信息。

4) 根据式(5)计算 $d_{ii}(p)$, 根据式(11)进行背景分割, 这里 β 取 0.48。

5) 经过以上步骤进行目标分割后, 应用中值滤波及物体空间连通性进行噪声抑制。

3 结果与讨论

实验采用的计算机配置为酷睿 22.0 GHz CPU, 2 Gbit 内存, 采用 matlab7.1 编程实现相关算法。分别采用 3 段视频对本文方法进行验证。视频 1 是光照有点微变的视频, 每帧大小为 240×352 ; 视频 2 是背景为树叶摇摆动态变化的交通视频, 每帧大小为 128×160 ; 视频 3 是大小为 480×1920 , 背景有风吹树叶摆动和太阳光照射, 运动目标是比较复杂的自行车运动, 同时还有阴影; 视频 4 是每帧大小为 384×288 , 视频中运动目标与背景颜色比较接近。分别采用经典高斯混合模型和本文方法对 4 段视频进行运动目标分割。

由图 1—4 可以看出, 从主观视觉效果上来看, 相比于本文方法, 经典高斯混合模型方法具有较多的误分割现象, 尤其随着背景环境复杂度的增加, 误分割现象越严重。在视频 3 实验中, 经典高斯混合模型方法的误分割现象已经严重到淹没了目标对象, 而本文方法仍然较好地分割出目标对象; 在视频 4 实验中, 目标上衣与背景较接近, 本方法检测效果较好, 而经典高斯方法 28 帧没有检测到目标上半部分, 53 帧只检测到目标的脚, 62 帧只检测到目标的头, 丢失目标较严重。因此, 本文方法在运动目标检测上优于经典高斯混合模型。

采用错判率和误判率进行客观评价分析。

1) 错判率指不是目标像素被判为目标像素与目标像素的比值; 2) 漏判率指目标像素未被判为目标像素与目标像素的比值; 这里目标像素通过手动分割获得。由图 5—8 可以看出, 本文方法的漏判率和误判率都低于经典高斯混合模型方法, 尤其当前景和背景颜色信息较为接近时, 本文方法更优于经典高斯混合模型方法, 与主观视觉效果相一致。

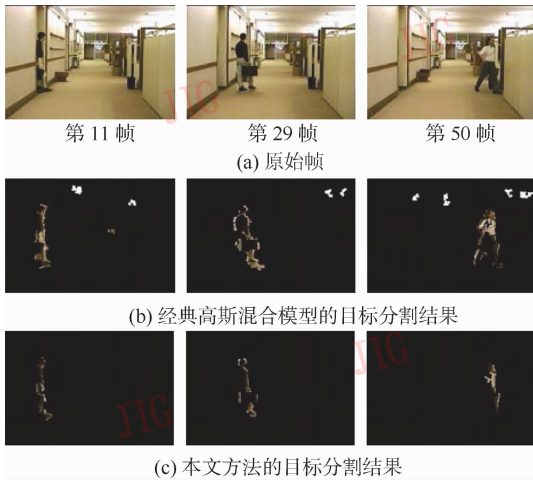


图1 视频1的仿真实验结果

Fig. 1 Experimental results of the 11st,29th,50th frame in video 1

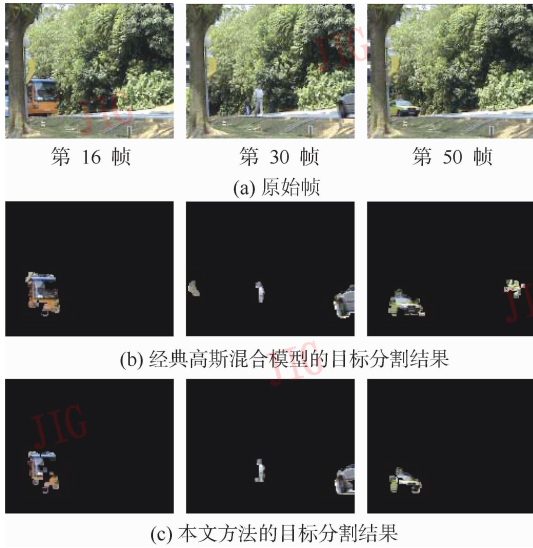


图2 视频2的仿真实验结果

Fig. 2 Experimental results of the 16th,30th,50th frame in video 2

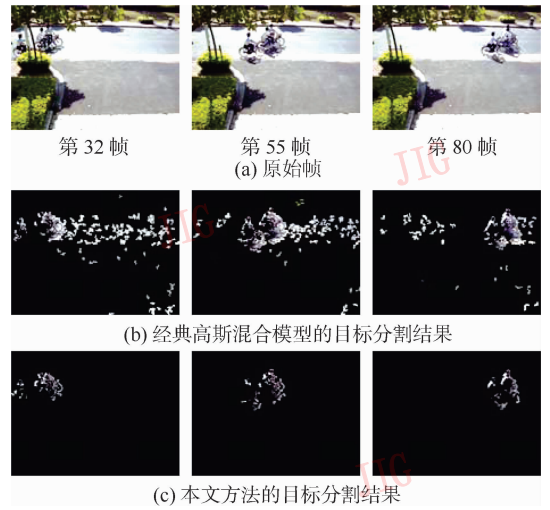


图3 视频3的仿真实验结果

Fig. 3 Experimental results of the 32nd,55th,80th frame in video 3

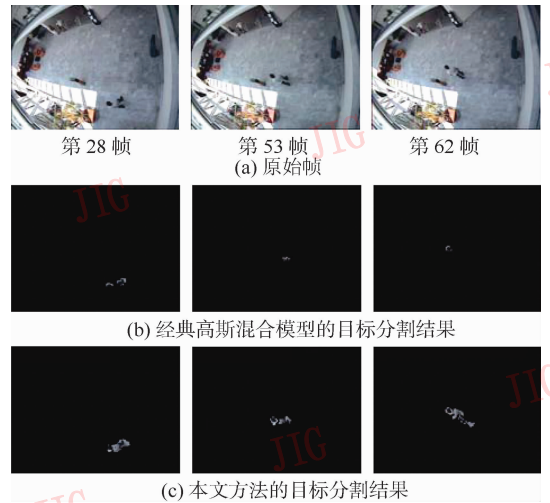


图4 视频4的仿真实验结果

Fig. 4 Experimental results of the 28th,53rd,62nd frame in video 4

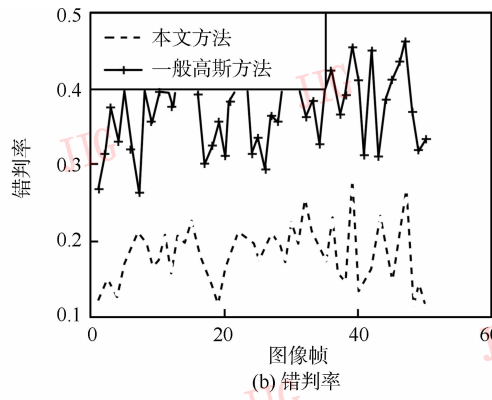
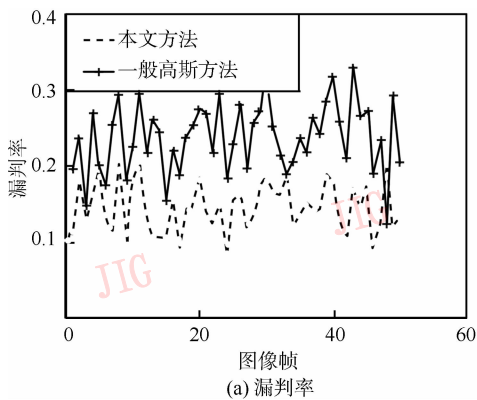


图5 视频1的错判率和漏判率

Fig. 5 Commission and omission error rate for video 1

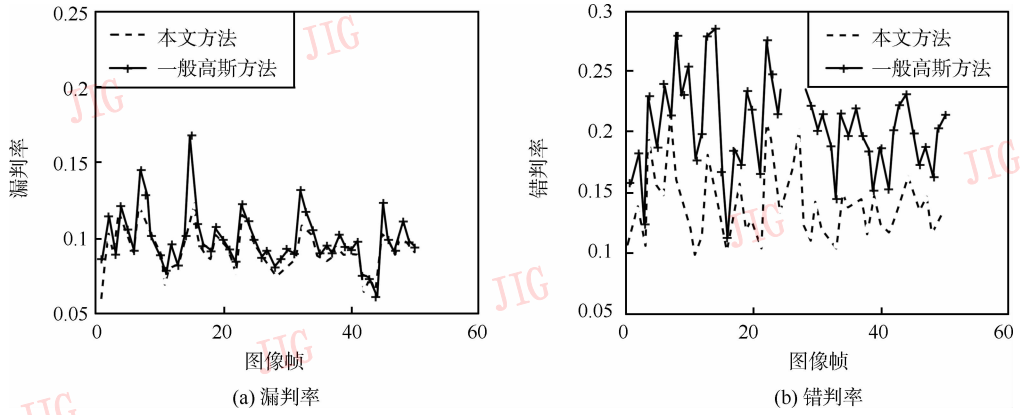


图6 视频2的错判率和漏判率

Fig. 6 Commission and omission error rate for video 2

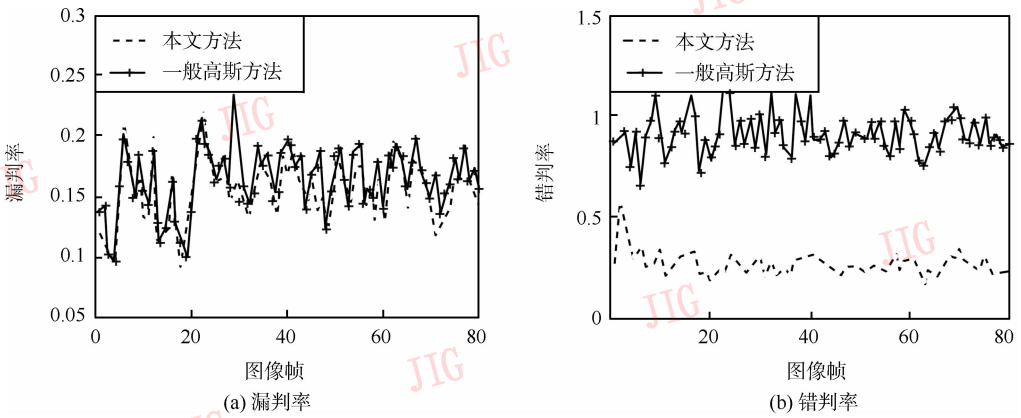


图7 视频3的错判率和漏判率

Fig. 7 Commission and omission error rate for video 3

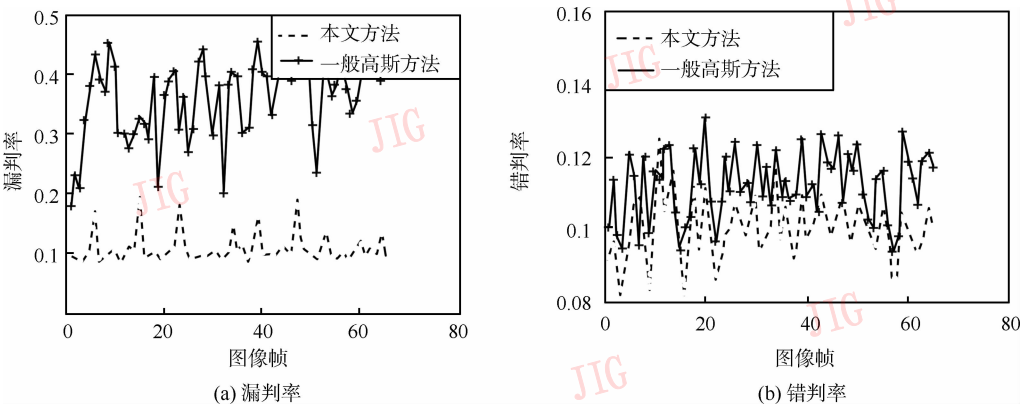


图8 视频4的错判率和漏判率

Fig. 8 Commission and omission error rate for video 4

4 结 论

提出一种基于高斯混合模型和小波变换的目标检测方法。采用高斯混合模型拟合变化的场景;采用小波变换提取运动图像的纹理特征,获得目标的局部结构信息,作为颜色信息的补偿;然后将两者融合起来进行运动目标检测。实验结果表明,该方法能够有效地进行运动目标检测,优于经典高斯混合模型方法。

参考文献(References)

- [1] Peng Suo, Wang Yanjiang. An improved adaptive background modeling algorithm based on Gaussian mixture model [C] // Proceedings of ICSP2008. Beijing: IEEE Press, 2008 :1426-1439.
- [2] Power P W, Schoonees J A. Understanding background mixture models for foregrounds segmentation [C] // Proceedings of Image and Vision Computing. New Zealand: Auckland, 2002; 267-271.
- [3] Chen Shiwen, Cai Nian, Tang Xiaoyan Improved moving object detection algorithm based on Gaussian mixture model [J]. Modern Electronics Technique, 2010, 33 (2): 125-127, 130. [陈世文, 蔡念, 唐孝艳. 一种基于高斯混合模型的运动目标检测的改进算法 [J]. 现代电子技术, 2010, 33 (2): 125-127, 130.]
- [4] Harville M, Gordon G, Woodfill J. Foreground segmentation using adaptive mixture models in color and depth [C] // Proceedings of IEEE Workshop on Detection and Recognition of Events in Video. Vancouver, BC, Canada; USA: IEEE Press, 2001; 3-11.
- [5] Yuan Chunfeng, Wang Chuanxu, Zhang Xiangguang, et al. Video segmentation of luminance abrupt variation based on MOGs and gradient information [J]. Journal of Image and Graphics, 2007, 11(12): 2068-2073. [原春锋, 王传旭, 张祥光, 等. 光照突变环境下基于高斯混合模型和梯度信息的视频分割 [J]. 中国图象图形学报, 2007, 11(12): 2068-2073.]
- [6] Liu Xin, Liu Hui, Qiang Zhenping, et al. Adaptive background modeling based on mixture Gaussian model and frame subtraction [J]. Journal of Image and Graphics, 2008, 4 (13): 729-735. [刘鑫, 刘辉, 强振平, 等. 混合高斯模型和帧间差分相融合的自适应背景模型 [J]. 中国图象图形学报, 2008, 4(13): 729-735.]
- [7] Zhong J, Sclaroff S. Segmenting foreground objects from a dynamic textured background via a robust Kalman filter [C] // Proceedings of International Conference on Computer Vision. Nice, France: IEEE Press, 2003: 44-50.
- [8] Jabri S, Duric Z, Wechsler H. Detection and location people in video images using adaptive fusion of color and edge information [C] // Proceedings of International Conference on Pattern Recognition. Barcelona, Spain: IEEE Press, 2000, 627-630.
- [9] Ercan Ozyildiz, Nils Krahnstover, Rajeev Sharma. Adaptive texture and color segmentation for tracking moving objects [J]. Pattern Recognition, 2002, 35 (10): 2013-2029.
- [10] Li Liyuan, Leung K H Maylor. Integrating intensity and texture differences for robust change detection [J]. IEEE Trans. Image Processing, 2002, 11(2): 105-112.