

文章编号:1001-9081(2007)10-2518-04

基于改进的 tMHI 的实时视频运动目标检测及跟踪

孟繁锋¹, 曾庆双¹, 姜永林¹, 高晓颖²

(1. 哈尔滨工业大学 航天学院, 哈尔滨 150001;

2. 中国运载火箭技术研究院 宇航智能控制技术国防科技重点实验室, 北京 100854)

(mengfanfeng@sohu.com)

摘要:提出了一种对视频图像进行实时目标分割及跟踪的新方法。该方法利用基于时间片的运动历史图像(tMHI)的灰度阶梯轮廓,对存在的子运动区域进行包围划分并予以标记,实现视频图像中运动目标的实时分割,进而将每帧 tMHI 图像中各个运动区域同场景中运动目标连续关联起来,实现对多运动目标的轨迹跟踪。为了提高分割质量,对 tMHI 进行了改进处理,去除了大部分噪声干扰,取得了明显的改善效果。实验表明,该方法可以有效地分割并跟踪视频中的多个运动目标,鲁棒性好,检出率较高,并且处理速度较快,达到了实时性的要求,还解决了局部粘连的问题。

关键词:基于时间片的运动历史图像;目标跟踪;阶梯轮廓;背景提取

中图分类号: TP391.41 **文献标志码:** A

Improved tMHI based real-time video moving object detecting and tracing

MENG Fan-feng¹, ZENG Qing-shuang¹, JIANG Yong-lin¹, GAO Xiao-ying²

(1. School of Astronautics, Harbin Institute of Technology, Harbin Heilongjiang 150001, China;

2. State Laboratory of National Defense for Aerospace Intelligent Control,
China Academy of Launch Vehicle Technology, Beijing 100854, China)

Abstract: A new method for segmenting and tracing real-time moving object from video image was purposed, this method using the step-down grey value silhouette of timed Motion History Image (tMHI). By enveloping and marking the motion silhouette regions, it can segment the moving objects from video image real-timely. By associating the motion silhouette regions in every tMHI with the moving objects in real scene, it achieved the trace of moving objects. To get high quality segmentation, the primary tMHI was improved by removing most of the noise. The method was demonstrated with a traffic video, and some moving objects were segmented effectively. It appeared that the method has good locating robust; high detecting rate and fast processing speed that satisfied real time usage. Part conglutination was basically avoided also.

Key words: timed Motion History Image (tMHI); object tracing; step-down silhouette; background extraction

0 引言

在视频监控领域,积累了多种多样的目标检测提取方法,根据方法对图像源的要求,可以归纳为三大类:1)基于单帧或双帧视频图像的方法^[1-5];2)基于多帧序列视频图像的方法^[6-14];3)基于帧对序列视频图像的方法^[15-18]。

对于实时视频检测,一方面要尽量降低对外界各种干扰的灵敏度,另一方面还要保证较高的处理速度、较好的鲁棒性以及精确的检测结果。前述各种视频检测方法为视频检测应用提供了足够的备选方案。在实际应用中,有些检测方法虽然简单、处理速度快,但抗干扰性差;有些方法效果虽好,但计算耗费高,无法用于实时处理。这就需要针对具体情况进行速度与效果的折中,来挑选合适的方案。

文献[19]提出了一种视频分割识别的方法,称为基于运动历史图像(Motion History Image, MHI)的方法,该方法对多帧视频进行分析,充分利用了信息资源,且相对简便、快速高效,可以实时地得到分析结果。文献[20]MHI 做出改进,提

出了基于时间的运动历史图像(timed Motion History Image, tMHI),并成功地将 tMHI 应用在人体姿态识别相关系统中。

基于 tMHI 的方法虽然较好地实现了对视频中运动目标的分割及姿态识别,但是却没有解决目标跟踪问题,且该方法对户外场景的噪声干扰敏感。本文改善了 tMHI 的输入图像,去除了噪声干扰,并提出了基于 tMHI 的视频运动目标跟踪算法,最后给出了实际应用效果。

1 tMHI

1.1 tMHI 的定义及更新

tMHI 的更新公式为:

$$tMHI_{\delta}(x, y) = \begin{cases} \tau, & (x, y) \in \text{当前轮廓} \\ 0, & tMHI_{\delta}(x, y) < \tau - \delta \end{cases} \quad (1)$$

其中, τ 为当前系统时间, δ 为当前 tMHI 所能表示的时间片段范围。这种定义方式使得 tMHI 对运动目标的表达独立于处理系统的工作性能。于是,基于不同处理性能及图像采集帧率的系统在同一参数、同一时间段下所得到的 tMHI 的历史

收稿日期:2007-04-04;修回日期:2007-06-25。

作者简介:孟繁锋(1976-),男,黑龙江哈尔滨人,博士研究生,主要研究方向:信号与信息处理、图像处理、视觉控制;曾庆双(1964-),男,吉林扶余人,教授,博士生导师,主要研究方向:惯性技术、运动控制、智能控制与智能系统;姜永林(1965-),男,黑龙江哈尔滨人,博士研究生,主要研究方向:工业自动化控制、智能交通;高晓颖(1969-),男,河北人,研究员,主要研究方向:智能控制、飞行器导航。

时间范围是相同的,图 1 为基本 tMHI 的实验结果。

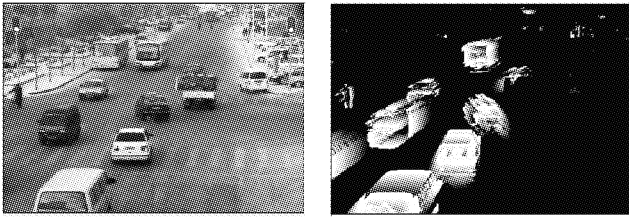


图 1 tMHI 结果

1.2 tMHI 的不足

上述实验的结果表明,tMHI 存在不足之处:

1) tMHI 中,运动目标的位置及大小的定位不稳定,偶尔会出现偏差,导致运动目标的错误划分,这影响到对运动目标的跟踪检测,因此,有待于提高对运动目标的定位精确度。

2) tMHI 虽然使用多帧序列视频图像信息构造,通过分割得到各运动目标信息,但其结果是基于单幅 tMHI 的,是针对图像全局处理而非针对具体目标的处理,使得 t 时刻从 tMHI 辨识获得的若干运动目标与 $t+1$ 时刻辨识获得的若干运动物体之间缺少关联信息,无法对同一目标进行连续跟踪定位并确定其运动轨迹。因此,需要引入新的目标跟踪定位算法策略,实现对具体运动目标的轨迹跟踪。

3) tMHI 的构造过程中,受外界光线变化和噪声干扰的原始视频帧图像被直接差分处理,差分后的图像清晰度较低,存在大量噪声,影响了 tMHI 的质量和目标定位精度。因此,需要对原始视频帧图像做预处理,提高 tMHI 的质量和结果精度。

2 tMHI 输入轮廓图像的改进

为了提高 tMHI 的质量,需要对输入视频序列进行处理优化,去掉冗余的信息。一种方法是将视频序列图像进行前景(运动目标)与背景(场景)的分离,然后从视频序列图像中去除背景,将剩余图像前景作为 tMHI 的输入,可以得到更精确的 tMHI。

2.1 实时背景提取及更新

最简单的背景提取方法就是在场景中没有任何前景目标时,手工获取背景。由于现场的光线以及其他因素的影响,需要适时地进行手工获取活动,增加了人力、物力需求,限制了应用范围,并且不适应快速变化的现场情况。要使视频系统快速适应亮度改变,遮挡、重现等环境的变化,就必须自动完成从视频图像序列中获取和更新背景的工作。

目前,主要的背景模型方法可以归纳为三大类,一类是基于统计的模型方法,主要包括高斯分布模型(Single Gaussian model)、混合高斯分布模型(Mixture of Gaussian, MOG)以及非参数化模型(Non-parametric Model)等;另一类是基于预测的模型方法,主要包括维纳滤波器方法(Wiener filter)以及卡尔曼滤波器方法(Kalman Filter, KF)等。此外还有近似中值滤波方法(Approximated median filter)及贝叶斯估计法等。

理论上讲,视频场景中的“背景”就是视频图像序列中灰度值变化范围很小的那些像素点的集合。与之对照,从统计学观点来说,“背景”中像素的灰度值是用统计的结果来表示的,是依据对视频图像序列的统计得出的各像素最可能出现的值。可以据此建立背景更新的统计学模型,并在此基础上形成一系列的基于统计的背景更新方法。典型的基于统计的背景更新公式如下:

定义 t 时刻背景图像 $Back_t(x, y)$ 为:

$$Back_t(x, y) = W(p_t(x, y), p_{t-1}(x, y), \dots, p_{t-(n-1)}(x, y), u(t)Back_{t-1}(x, y)) \quad (2)$$

其中, $W(\cdot)$ 为对背景进行更新的函数, $p_x(x, y)$ 为 i 时刻的视频帧图像, $t-(n-1) \leq i \leq t, n$ 为参与更新的图像序列的帧数, $u(t)$ 为权值函数。选取不同的更新函数 $W(\cdot)$, 可以得到不同的背景更新算法。

文献[21]认为一个系统状态的最优信息是通过估计获得的,并引入了 Kalman 滤波器。该方法对噪声有较强的抑制作用,模型参数对各种天气条件下的背景提取效果影响不大,只适合背景变化非常缓慢的场合。

基于以上背景模型的研究,为提高背景更新质量,本文采用一种改进的基于形态学的方法完成背景的实时更新。首先,为了最大限度地减少噪声干扰,特别是摄像设备引起的高斯噪声,采用高斯滤波器对采集得到的视频图像进行了滤波去噪处理。然后,将灰度图像序列进行逐帧差分:

$$e_i(x, y) = p_i(x, y) - p_{i-1}(x, y) \quad (3)$$

其中, $e_i(x, y)$ 为 i 时刻灰度图像与 $i-1$ 时刻灰度图像的差分图像。

为了保持视频帧图像中运动区域的范围稳定,消除区域中的砂眼噪声,对逐帧差分图像 $e_i(x, y)$ 进行灰度形态学闭运算:

$$E_i(x, y) = e_i(x, y) \cdot H(j, k) \quad (4)$$

其中, $H(j, k)$ 是一个 $l \times l$ 的结构单元。 $E_i(x, y)$ 为包含完整运动区域的最终差分图像。对视频背景图像的更新,采取有条件的统计更新算法,差分图像 $E_i(x, y)$ 作为背景更新算法的条件函数。定义 i 时刻背景图像 $Back_i(x, y)$ 为:

$$Back_i(x, y) = W(p_i(x, y), \dots, p_{i-(n-1)}(x, y), E_i(x, y), u(i)Back_{i-1}(x, y), c_i(x, y)) \quad (5)$$

其中, $u(i)$ 为前一帧背景的加权因子, $c_i(x, y)$ 为阈值函数。差分图像中小于阈值 $c_i(x, y)$ 的像素点 $E_i(x, y)$ 视为未发生变化的当前背景,参考背景图像中对应像素点被算法加权更新;差分图像中大于阈值的那些像素点视为运动区域,保持参考背景图像中对应点值不变,这样就生成了新的参考背景。

将原始当前视频帧图像与背景图像作差分,就得到了当前场景中的运动目标前景:

$$M_i(x, y) = p_i(x, y) - Back_i(x, y) \quad (6)$$

得到的图像 $M_i(x, y)$ 只包含当前的运动目标,作为 tMHI 方法的输入图像。

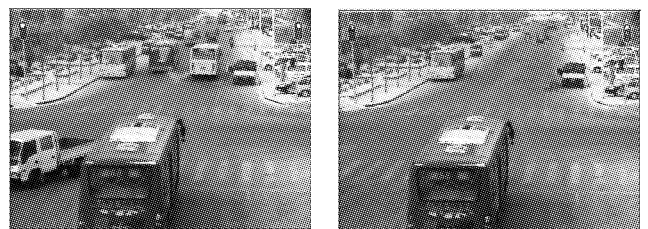


图 2 实时背景提取结果

2.2 改进的 tMHI 更新方法

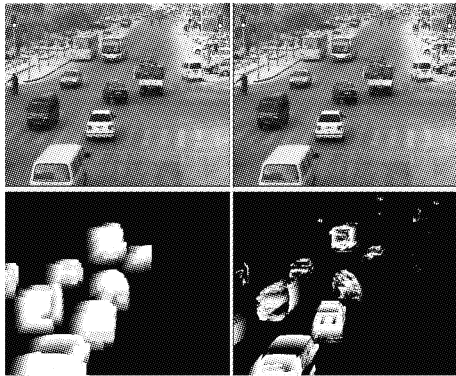
对应视频场景的变化,tMHI 图像需要实时地更新。与前述基本 tMHI 方法不同的是,本文采用了先填充再累加的方法,减少了原方法的运算复杂度,提高了执行效率。首先,对图像 $M_i(x, y)$ 进行填充,将像素值大于一定阈值的那些区域点的像素值重新填充为指定值:

$$Fill_i(x, y) = \begin{cases} C_{cur}, & M_i(x, y) \geq l \\ 0, & M_i(x, y) < l \end{cases} \quad (7)$$

其中, C_{cur} 为设定的最大像素值, $0 < C_{cur} \leq 255$, l 为判断阈值: $0 < l < C_{cur} \leq 255$ 。将当前 $Fill_i(x, y)$ 合并到 MHI 图像中:

$$MHI_{i+1}(x, y) = \begin{cases} [MHI_i(x, y) - C_{step}] + Fill_i(x, y), & MHI_i(x, y) - C_{step} > 0 \\ Fill_i(x, y), & MHI_i(x, y) - C_{step} \leq 0 \end{cases} \quad (8)$$

其中, C_{step} 为 tMHI 中的像素阶梯深度。可见, 运动焦点区域的像素值最大, 这就突出了运动目标。这样得到的不断更新的 tMHI 图像, 就可以应用各种方法对其进行细致分析, 以从中获得需要的信息。图 3 为改进后的 tMHI 效果与原效果的比较。



(a)改进后效果 (b)原效果

图3 效果比较

3 目标的分割及跟踪

为实现运动目标的检测跟踪, 需要将单元目标从实时更新的 tMHI 序列中分割出来, 然后分别加以标记, 才能达到实际应用的要求。

图像的分割算法很多, 有基于二值图像的、基于灰度的、和基于彩色的, 这些算法的结果都是被分割物体的具体轮廓区, 绝大多数分割算法尽管结果很精确, 但是算法复杂度较高, 算法执行时间较长, 达不到实时性的要求。由于 tMHI 是介于纯粹灰度图像与二值图像之间的简单均匀阶梯灰度值图像, 可以设计一种全新的针对 tMHI 的分割算法, 以满足实时性要求。

对于 tMHI 中运动目标的遮挡情况, 利用 tMHI 的阶梯均匀灰度值可以解决。采用包围分割子区域的最小矩形来表示分割出的各个运动目标的运动历史区, 并分别标记出来。 $M_k(\min V, \max V, step V)$ 为被分割的 tMHI 图像, 其中 $\min V$ 为最小阶梯灰度值, $\max V$ 为最大阶梯灰度值, 而 $step V$ 为灰度阶梯, 且 $\max V - \min V = \delta$ 。此分割算法描述如下:

$$Area_i(k) \rightarrow rect = \begin{cases} Rect_j(k), & Diff(Area_i(k-1) \rightarrow rect, Rect_j(k)) < Limit \\ Area_i(k-1) \rightarrow rect, & other \end{cases} \quad (10)$$

$$Area_i(k) \rightarrow old = \begin{cases} zero, & Diff(Area_i(k-1) \rightarrow rect, Rect_j(k)) < Limit \\ add1, & other \end{cases} \quad \text{if } (Area_i(k) \rightarrow old > OldLim), clear(Area_i) \quad (11)$$

其中, $Limit$ 为区域偏差上限, $OldLim$ 为衰老上限;

3) 更新运动目标的运动轨迹。利用人形心法, 取运动目标出现区域的形心为运动物体当前的轨迹点, 将此点加入到

1) 算法从图像原点开始逐行搜索图像, 直至遇到一个像素为最大阶梯值的点, 该点属于运动区域之一;

2) 从该点开始, 对其所属运动区域进行向下的边界搜索, 以得到区域的轮廓;

3) 搜索的有效目标运动区域的下限像素值为 $\min V$, 如果遇到像素值跨越几个阶梯的增大突变, 则表明前一运动目标的历史残影与下一个运动目标重叠, 此处为两个运动区域的分界线;

4) 当搜索到的轮廓边缘形成闭合时, 则该运动区域搜索完毕;

5) 将得到的运动区域标记为已读, 于是后续的操作将此区域视为 0;

6) 将搜索得到的运动区域边界轮廓的坐标进行比较, 得到横、纵坐标的最大、最小值 $\min x, \min y, \max x, \max y$, 它们所包围的矩形区即是所分割出的单元目标的运动历史区;

7) 回到 1), 继续进行搜索, 直至将整个图像搜索完毕;

8) 根据所关注目标的尺度范围, 滤除过大或者过小的分割区;

9) 若最终分割得到 m 个运动区域, 则将所得参数记入记录 $Rect_i(k)$, 即第 k 帧图像中第 i 个运动区域, 其中 $i = 1, 2, \dots, m$, 算法结束。

这样, 就得到了对 tMHI 图像的分割结果。

为了取得运动目标的各项参数, 以便对运动目标进行分析, 就需要对分割出的若干运动目标区域进行跟踪, 进一步获得运动目标的运动轨迹。在上述图像分割算法中, 已经获得了每帧 tMHI 中各运动目标区域, 现在需要把每帧图像中的各个运动区域同场景中的运动目标连续关联起来, 这可以通过算法将运动区域对照相应运动目标分别标注记号来实现。

令 $Area\{rect, object, trace, old\}$ 为存储运动目标参数的结构, $Area_i(k)$ 表示第 i 个运动物体在 k 时刻图像中的出现区域, 则有如下算法:

1) 初始化, 将首次分割出来的 m 个运动区域分别记入 m 个运动目标结构记录中, 即:

$$Area_i(0) \rightarrow object = i;$$

$$Area_i(0) \rightarrow rect = Rect_i(0);$$

$$i = 0, 1, \dots, m \quad (9)$$

2) 对于 t 时刻新获得的帧图像中的 n 个运动区域, 由于图像序列的时空连贯性, 将 $t-1$ 时刻结构记录中的各个运动目标区域与这 n 个区域进行逐个比较, 如果找到满足条件的相近区域, 则该区域为该运动目标的当前出现区域, 记录新的区域值, 并标记该区域编号为此运动目标号。如果找不到相近区域, 为了防止当前帧分割出现的偶然误差导致运动目标区域的辨认缺陷, 未立即删除该运动目标记录, 而是将此运动目标结构记录中的衰老计数值 $Area_i(k) \rightarrow old$ 增 1, 若此值超过一定的阈值, 则视为当前场景中已经不存在该运动目标, 停止相应记录更新。

轨迹点集, 对更新后的轨迹点集, 通过最小二乘法进行曲线拟合, 得到更新的运动目标轨迹方程;

4) 对于没有对应运动目标的运动区域, 创建新的运动物

体记录,并将运动区域参数赋值给该记录;

5) 根据运动目标记录,对运动目标进行特定的分析;

6) 返回 1),对下一帧图像进行处理。

4 实验结果

在 Visual C++ 编程环境下基于以上算法编制了验证程序,对某交通道口监控视频进行验证分析。结果如图 4 所示。

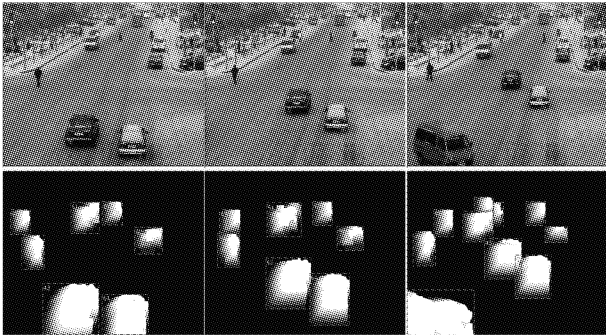


图 4 视频跟踪处理结果

图中,上半部分的三幅图像显示的是某三个连续相近时间点的三帧现场视频图像,按时间顺序从左到右。下半部分的三幅图像显示的是分别在其上对应的视频图像的时间点上算法的当前处理结果。小于预定阈值的过小运动目标已经被滤除掉。结果中,每一个被分割出的运动目标都被用矩形框根据其轮廓大小表示出来,且按出现时间早晚分别被标记上识别序号。从中可以看出,算法可以区分局部粘连的运动目标。

视频图像帧率为标准 PAL 制,实验结果表明处理算法满足实时性及鲁棒性要求。

5 结语

本文提出了一种对视频图像进行实时目标分割及跟踪的新方法,该方法利用 tMHI 取得了良好效果。实验结果表明,该方法实现了对视频场景中多运动目标的快速实时分割及跟踪,解决了局部粘连的问题,鲁棒性强。下一步工作将集中在进一步提高效能,以便能够增强在更加复杂的现场中识别运动目标的鲁棒性(这些目标可能具有更加多样的运动方向及速度)。

参考文献:

- [1] PARK Y. Shape-resolving local thresholding for object detection [J]. Pattern Recognition Letters, 2001, 22(5): 883 - 890.
- [2] WAN C L, DICKINSON K W, BINNIE T D. A cost-effective image sensor system for transport applications utilizing a miniature CMOS single chip camera [C]// IFAC Transportation systems, Tianjin, Proceedings 1994.
- [3] SHIMIZU K, SHKIGEHARA N. Image processing system used cameras for vehicle surveillance [C]// IEEE Second International Conference on Road Traffic Monitoring, Conference Publication Number 1989, 299: 61 - 65.
- [4] KLAUSMANN P, KROSCHEL K, WILLERSINN D. Performance prediction of vehicle detection algorithms [J]. Pattern Recognition, 1999, 32: 2063 - 2065.
- [5] FATHY M, SIYAL M Y. An image detection technique based on morphological edge detection and background differencing for real-time traffic analysis [J]. Pattern Recognition Letters, 1995, 16: 1321 - 1330.
- [6] PARAGIOS N, DERICHE R. Geodesic active contours and level sets or the detection and tracking of moving objects [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(3): 266 - 280.
- [7] KIM J B, PARK H S, PARK M H, *et al.* A real-time region-based motion segmentation using adaptive thresholding and K-means clustering [C]// Proceedings of the 14th Australian Joint Conference on Artificial Intelligence: Advances in Artificial Intelligence LNCS 2256. London: Springer-Verlag, 2001: 213 - 224.
- [8] BADENAS J, SANCHIZ J M, PLA F. Motion-based segmentation and region tracking in image sequences [J]. Pattern Recognition, 2001, 34: 661 - 670.
- [9] MICHALOPOULOS P G, PANDA D P. Derivation of advanced traffic parameters through video imaging [C]// IFAC Transportation System. Chania, Greece: Elsevier, 1997: 975 - 980.
- [10] PARAGIOS N, DERICHE R. Geodesic active contours and level sets for the detection and tracking of moving objects [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(3): 266 - 280.
- [11] NAGEL H H, ENKELMANN W. An investigation of smoothness constrains for the estimation of displacement vector fields from image sequences [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986: 565 - 593.
- [12] HORN B K P, SCHUNCK B G. Determining optical flow [J]. Artificial Intelligence, 1981, 17: 185 - 203.
- [13] KORIESS R, NAGEL H H, ZIMMERMANN G. Motion detection in image sequences: an evaluation of feature detectors [C]// Proceedings of International Joint Conference on Pattern Recognition. Montreal: [s. n.], 1984: 778 - 780.
- [14] CARLSSON S, EKLUNDH J O. Object detection using model-based prediction and motion parallax [C]// Proceedings of Europe Conference on Computer Vision. New York: Springer-Verlag, 1990: 297 - 306.
- [15] BERTOZZI M, BROGGI A. Vision-based vehicle guidance [J]. Computer Vision, 1997, 30 (7).
- [16] POMERLEAU D, JOCCHEM T. Rapidly adaptive machine vision for automated vehicle steering [J]. IEEE Machine Vision, 1996, (4): 19 - 27.
- [17] BERTOZZI M, BROGGI A. GOLD: a parallel real-time stereo vision system for generic obstacle and lane detection [J]. IEEE Transaction Image Processing, 1998, 7(1).
- [18] KOLLER D, DANILIDIS K, NAGEL H. Model-based object tracking in monocular image sequences of road traffic scenes [J]. International Journal Computer Vision, 1993, 10: 257 - 281.
- [19] DAVIS J, BOBICK A. The representation and recognition of human movement using temporal templates [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001, 23(3): 257 - 267.
- [20] BRADSKIL G R, DAVIS J W. Motion segmentation and pose recognition with motion history gradients [J]. Machine Vision and Applications, 2002, 13(7): 174 - 184.
- [21] RIDDER C, MUNKELT O, KIRCHNER H. Adaptive background estimation and foreground detection using Kalman - filter [C]// Proceedings of the Int'l Conference on Recent Advances in Mechatronics, ICRAM'95, Istanbul, Turkey, 1995: 193 - 199.