

# 一种基于区域的视频车辆跟踪系统

石时需<sup>1,2</sup>, 郑启伦<sup>1</sup>, 曹 波<sup>1</sup>

(1. 华南理工大学计算机科学与工程学院, 广州 510640; 2. 广东新粤交通投资有限公司, 广州 510410)

**摘 要:** 基于区域跟踪法提出一种视频车辆跟踪系统, 改进了系统中涉及的多个方法。以车辆的提取检测结果为基础, 提出基于混合差分的车辆检测方法, 使检测更快速, 检测结果质量更高。通过设计堆栈搜索技术和车辆识别规则, 提高了基于差分法的识别准确率, 改进了区域跟踪效果。实验结果证明了系统的可行性与实用性。

**关键词:** 基于区域的跟踪; 视频车辆跟踪; 混合差分; 视频事件检测

## Vehicle Video Tracking System Based on Region

SHI Shi-xu<sup>1,2</sup>, ZHENG Qi-lun<sup>1</sup>, CAO Bo<sup>1</sup>

(1. College of Computer Science and Engineering, South China University of Technology, Guangzhou 510640;

2. Guangdong Xinyue Communication Investment Corporation Limited, Guangzhou 510410)

**【Abstract】** This paper proposes a video tracking system for moving vehicle based on region and several improved techniques of the system. The tracking quality based on region lies on the detecting results. A hybrid difference strategy is introduced to attain higher detecting quality than other methods. A stack searching and a discriminating rule are designed to optimize the vehicle recognition. Experimental results indicate the video tracking system is feasible and useful in application of intelligent transportation system.

**【Key words】** tracking based on region; vehicle video tracking; hybrid difference; video incident detection

### 1 概述

近年来随着计算机和图像处理技术的不断发展, 利用视频技术进行车辆检测成为一种很有潜力的方法<sup>[1]</sup>。视频运动目标跟踪是计算机视觉技术的一个重要研究课题, 在智能交通系统和视频监控等方面均有广泛的应用<sup>[2]</sup>。视频车辆检测与跟踪系统通常包括感兴趣区域(Region Of Interesting, ROI)提取、车辆检测、车辆对象识别、车辆跟踪等4个模块。ROI提取常用的方法有背景差分法<sup>[3]</sup>和混合高斯模型(Gaussian Mixture Model, GMM)法<sup>[4-5]</sup>: 背景差分法速度快, 但容易造成后期对象识别不准确; GMM方法检测效果好, 但计算量太大以致运行速度慢。

在ROI结果的基础上, 可以用阈值法<sup>[6]</sup>进行运动车辆的检测, 然而这类方法容易出现噪音干扰。在进行车辆跟踪之前, 需要从检测结果中识别出车辆对象, 区域跟踪法<sup>[3]</sup>就是在车辆检测结果中识别出一个个像素连通区域, 这些区域表示检测出的车辆。这种方法在车辆稀少时效果很好, 且得到的区域可以提供丰富的信息, 如大小、形状、密度<sup>[7]</sup>。用于车辆跟踪的主要技术是滤波跟踪, 有卡尔曼滤波<sup>[8-9]</sup>和粒子滤波<sup>[10]</sup>等。

为了改善跟踪效果, 本文设计了一种基于区域的视频车辆跟踪系统, 并以道路车辆的视频监控视频测试了 GMM 算法、背景差分法和改进方法的车辆识别效果、车辆跟踪准确率和处理时间。实验结果表明: 本文的跟踪系统在保证准确率的前提下, 比 GMM 算法更快速, 抗噪声能力更强; 比背景差分法识别的对象轮廓更加清晰, 提取信息更准确, 且鲁棒性更好。

### 2 视频车辆跟踪系统流程

视频车辆跟踪系统实现的一般流程如图1所示。

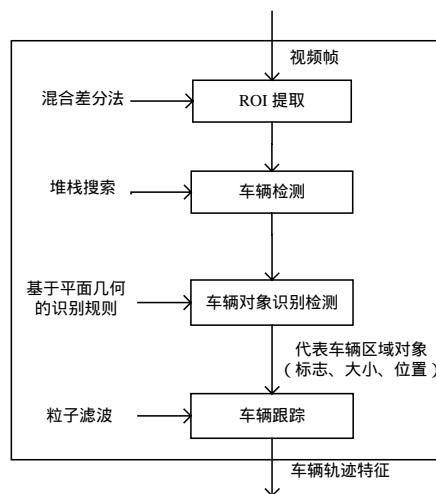


图1 基于区域的视频车辆跟踪系统流程

如图1所示, 本文采用混合差分法进行ROI提取, 快速得到高质量的前景二值图。在二值图基础上, 设计了基于堆栈的搜索技术, 对车辆在视频帧的区域进行了检测和标识。对标识后的区域提出了基于平面几何的识别规则, 从而将视频中的车辆识别出来作为跟踪分析的对象。最后将卡尔曼滤波技术<sup>[8]</sup>应用于车辆跟踪, 得到跟踪轨迹。

**基金项目:** 广东省自然科学基金资助项目(07006474); 广东省科技攻关基金资助项目(2007B010200044)

**作者简介:** 石时需(1974-), 男, 博士研究生, 主研方向: 智能视频技术, 智能交通; 郑启伦, 教授、博士生导师; 曹 波, 硕士研究生

**收稿日期:** 2008-01-02

**E-mail:** shishixu@126.com

### 3 混合差分法

首先用混合差分法实现跟踪系统的 ROI 提取。

差分方法是基于图像矩阵模型的，即把视频任意帧的图像( $n \times m$ 个像素点)用一个 $n \times m$ 的矩阵 $G$ 进行刻画。 $G^t(i, j)$ 表示第 $t$ 帧( $i, j$ )位置像素点的像素值，其中， $i=1, 2, \dots, n, j=1, 2, \dots, m, t=1, 2, \dots, n_0$ 。

最简单的差分法是用当前帧与前一帧相减，计算同一位置像素值之差，如下式所示：

$$D^t(i, j) = |G^t(i, j) - G^{t-1}(i, j)| \quad (1)$$

其中， $D^t(i, j)$ 为差分后所得绝对值，表示第 $t$ 和第 $t-1$ 帧( $i, j$ )位置像素值之差， $t=1, 2, \dots, n_0$ 。而两两差分后的图像由连续 2 个两帧相减所得(该图像为二值图像)：

$$B^t(i, j) = \begin{cases} 1 & \text{if } D^t(i, j) > T \\ 0 & \text{if } D^t(i, j) < T \end{cases} \quad (2)$$

其中， $T$ 为阈值，一般根据经验设定。这种方法虽然简单且易于实现，但检测出的视频对象并不完整清晰。为了避免相邻 2 帧某些像素点上相差不明显，可以采用间隔差分法作为改进，即取当前帧和前面某一帧相减。常用的做法是隔  $k$  帧相减， $k$  一般取 3，如下式所示：

$$D^t(i, j) = |G^t(i, j) - G^{t-k}(i, j)| \quad (3)$$

由于实现简单且响应速度快，式(1)和式(3)所示的差分法常用于辅助视频处理<sup>[11-12]</sup>。但当视频目标对象内部的灰度比较均匀时，间隔差分法可能在对象的重叠部分形成空洞，难以检测出物体的整个轮廓。如公路上的汽车，一般内部的灰度都比较均匀，且移动速度较快，要对这样的对象进行识别提取，需要改进差分法的性能。

为了提高对视频前景识别的敏感度，本文设计了一种新的混合差分策略。新策略保存前 3 帧图像，用当前帧分别与前 3 帧图像两两差分，如式(4)所示：

$$D_k^t(i, j) = |G^t(i, j) - G^{t-k}(i, j)| \quad (4)$$

其中， $D_k^t(i, j)$ 表示第 $t$ 和第 $t-k$ 帧( $i, j$ )位置像素值之差，式(2)改为

$$B^t(i, j) = \begin{cases} 1 & \text{if } \bigvee_{s=1}^k D_s^t(i, j) > T \\ 0 & \text{if } \bigvee_{s=1}^k D_s^t(i, j) < T \end{cases} \quad (5)$$

只要在相邻  $k$  帧中有像素值的变化，混合差分策略就可以识别出。 $k$  取值太小时，得到的差分较小，接近式(3)的效果，提取的对象易出现不清晰； $k$  取值太大时，会使检测过于敏感，很可能把静态的背景误识别为前景，出现较大噪声。经过多组实验，确定  $k$  的最优取值为 3。

混合差分法得到的结果是一个二值图，像素值为 1 的是前景点，像素值为 0 的是背景点。

### 4 堆栈搜索标识算法

在车辆检测阶段，在混合差分法得到的二值图中，将前景点分类标识成区域，一个区域代表一个运动对象(车辆)。这里设计了一种基于堆栈的搜索标识算法，实现对车辆区域的标识。

像素点用一个五元组  $B(x, y, f, s)$  记录，像素点的相关数据描述如下：

(1)  $b$  为像素的灰度值， $b=1$  为前景， $b=0$  为背景。

(2)  $f$  为当前旋转访问的角度，共有 8 个，取值为 1~8，顺时针访问，如图 2 和定义所示。

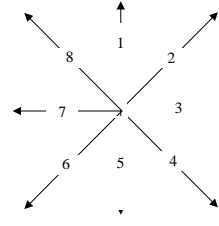


图 2 访问函数  $H(x, y, f)$  8 个角度的旋转

定义  $B(x, y, f, s)$  的访问函数  $H(x, y, f)$  为

$H(x, y, 1)=(x, y+1)$ ,  $H(x, y, 2)=(x+1, y+1)$

$H(x, y, 3)=(x+1, y)$ ,  $H(x, y, 4)=(x+1, y-1)$

$H(x, y, 5)=(x, y-1)$ ,  $H(x, y, 6)=(x-1, y-1)$

$H(x, y, 7)=(x-1, y)$ ,  $H(x, y, 8)=(x-1, y+1)$

其中， $f=1, 2, \dots, 8$ ，表示 8 个角度的旋转访问。

(3)  $s$  表示访问标志， $s=0$  表示未访问的点， $s=1, 2, \dots, N$  表示已经访问的点， $N$  是区域数目，同一区域的像素点  $s$  取值相同。

基于堆栈的区域标识算法描述如下：

初始化  $x=0, y=0, N=0$ ，每个像素点的值  $f$  取 1。

(1)  $x=x+1$ ，如果  $x>n$  且  $y>m$ ，算法退出；如果  $x>n$  且  $y \leq m$ ，则  $y=y+1, x=0$ ；读取像素点( $x, y$ )的信息至  $B$ 。

(2) 如果  $B.s=0$  (未访问过的点)，则用  $C$  记录  $B$  的信息， $N=N+1$ ，转(3)；否则， $B.s=1$  (即访问过的点)，转(1)。

(3) 判断  $C.b=0$  是否成立：

1) 成立：如果栈非空，则出栈至  $C$ ，转(4)；否则，栈空，转(1)；

2) 不成立：将  $C$  入栈，即  $C.b=1, C.s=N$ ，转(4)。

(4) 如果  $C.f \neq 8$ ，用  $T$  记录  $H(C.x, C.y, C.f)$  的信息， $C.f=C.f+1$ ，将  $C$  入栈，用  $C$  取  $T$  的信息，转(3)；否则， $C.f=8$ ，即 8 个角度都已访问，判断栈是否为空：如果非空，则出栈至  $C$ ，转(4)；如果栈空，转(1)。

以上算法可以将不同区域用不同数字作标识，相同区域的标识值相同，如图 3 所示。

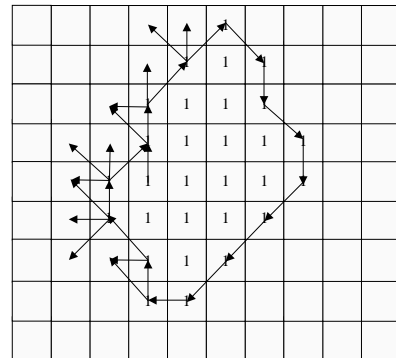


图 3 第 1 个区域的标识示例

最后用区域扫描算法记录标识了的区域信息，算法描述如下：

初始化  $k=0$ 。

(1)  $k=k+1$ 。

(2) 扫描二值图中  $s$  值为  $k$  的像素点，记录像素点数(面积)、最大  $y$  值  $y_{\max}$ 、最小  $y$  值  $y_{\min}$ 、最大  $x$  值  $x_{\max}$ 、最小  $x$  值  $x_{\min}$ 、中心点坐标  $(x_c, y_c)$ 。

(3) 如果  $k=N$  ( $N$  为标识得到的区域总数)，则算法退出；否则，转(1)。

所记录的区域信息将作为之后车辆对象识别的依据。

## 5 基于平面几何的车辆识别规则

本文在车辆检测区域图的基础上，研究设计了一种基于平面几何的识别规则，用于识别具体的车辆对象，将车辆对象从视频图像中分割出来。

假设在车辆检测中得到的区域数目为 $N$ ，对于任意区域以一个矩形作为标记，矩形中心为 $(x_c, y_c)$ ，宽高为 $W=x_{\max}-x_{\min}$ ， $H=y_{\max}-y_{\min}$ 。车辆识别本质上是一个区域合并的过程，如果所有区域都不需要合并，则原来的 $N$ 个区域直接标识为 $N$ 个车辆对象。

区域两两合并的过程如下：

- (1) 求出矩形 $A, B$ 的中心点 $(x_c^{(a)}, y_c^{(a)})$ ， $(x_c^{(b)}, y_c^{(b)})$ 。
- (2) 若2个矩形的宽分别为 $W_a, W_b$ ，高分别为 $H_a, H_b$ 。

令变量

$$w = (W_a + W_b) / 2$$

$$h = (H_a + H_b) / 2$$

(3) 令

$$u = |(x_c^{(a)} - x_c^{(b)}) - w|$$

$$v = |(y_c^{(a)} - y_c^{(b)}) - h|$$

$$u' = |x_c^{(a)} - x_c^{(b)}|$$

$$v' = |y_c^{(a)} - y_c^{(b)}|$$

(4) 取

$$\alpha = \max(W_a, W_b) / 5$$

$$\beta = \max(H_a, H_b) / 5$$

当以下3条规则成立时，矩形 $A, B$ 代表的2个区域合并：

- 1)  $u < \alpha$  &  $v < \beta$ ；
- 2)  $u' < \alpha$  &  $v < \beta$ ；
- 3)  $u < \alpha$  &  $v' < \beta$ 。

当没有区域需要合并时，每个区域作为一个车辆对象从图像中分割出来，并用一定的形状来区分汽车块。为了避免标识符号重叠，采用了按一定比例缩小的椭圆取代原来的矩形符号。具体实现方法如下：

对图像 $P$ 提取空间矩：

$$M_{x\_order, y\_order}(P) = \sum_{x,y} [I(x, y) \cdot x^{x\_order} y^{y\_order}]$$

其中， $I(x, y)$   $P$  是像素 $(x, y)$ 的亮度值。

先求矩形 $R$ 的中心矩：

$$M_{0,0}(R) = \sum_{x,y} [I(x,y) \cdot x^0 y^0]$$

椭圆的中心为

$$x_M = M_{1,0}(R) / M_{0,0}(R)$$

$$y_M = M_{0,1}(R) / M_{0,0}(R)$$

椭圆的长轴为

$$xx = 4 \cdot \sqrt{M_{2,0}(R) / M_{0,0}(R) - x^2}$$

椭圆的长轴为

$$yy = 4 \cdot \sqrt{M_{0,2}(R) / M_{0,0}(R) - y^2}$$

以代表车辆的椭圆作为跟踪对象，再用卡尔曼滤波<sup>[8]</sup>实现车辆的跟踪，由此实现的基于区域的视频车辆跟踪系统如图1所示。

## 6 实验结果与分析

本文进行了2组对比实验：第1组是背景差分法<sup>[13]</sup>和本

文方法在跟踪图像上的直观比较；第2组是GMM方法<sup>[4]</sup>、背景差分法<sup>[13]</sup>和本文方法在跟踪准确率和处理时间上的对比。在测试平台(CPU为P4 2.4 GHz、内存为512 MB)上实现了多组高速公路交通视频检测跟踪，这里随机抽取了一个例子作为演示。

图4是用背景差分法<sup>[13]</sup>实现的车辆识别效果图，左上角的卡车由于过大被分割为2个对象，出现了车辆的误识别。图5是用本文方法实现的车辆识别效果图，其中左上角的卡车被准确地识别为一个对象。用背景差分法跟踪时，图6中顶部带反光玻璃的车辆被识别为2个对象，也出现了误检测。同样的视频用本文方法跟踪时，车辆都被准确无误地识别出来，如图7所示，这说明本文方法的车辆识别准确率高于背景差分法。



图4 背景差分法的车辆识别图示1



图5 本文方法的车辆识别图示1



图6 背景差分法的车辆识别图示2



图 7 本文方法的车辆识别图示 2

表 1、表 2 给出了 3 种方法对一组视频在跟踪准确率和处理时间方面的对比。

表 1 3 种跟踪方法的跟踪准确率对比

算法	实际车辆	检测出	正确检测	少检测	多检测	误测原因
背景差分	84	94	74	10	10	检测不完整漏识别；把对象一分为二
GMM	84	87	83	1	4	将噪声误检成汽车
本文方法	84	83	83	1	0	将相邻的车辆合并

表 2 3 种跟踪方法的处理时间对比 s

对比算法	60 s 测试视频	120 s 测试视频	180 s 测试视频
背景差分	65	110	198
GMM	330	669	1 029
本文方法	45	89	136

从表 1 可以看出，本文方法的跟踪准确率最高，几乎所有的车辆对象都能被准确跟踪到，并且没有被噪声干扰。基于背景差分法的跟踪<sup>[13]</sup>则因为不完整的前景识别而漏检了多个车辆对象，且由于将一个车辆对象分割为 2 个对象而出现了多检测。GMM 算法因为噪声过大造成了几次多检测。从总体效果上看，GMM 识别的精度最高、范围最广，但其抗噪能力弱，容易将噪声当作车辆处理。背景差分效果最差，检测失误率很高，尤其是在检测颜色与路面相近的车辆时，效果非常差。本文方法的抗噪能力最强，这带来的轻微副作用就是：较远处相邻车辆区域间隔不明显时会被误合并为一个对象，因此，会少检测视频中远处模糊不清的车辆。

表 2 显示本文算法的响应速度最快，实时性最高，基本没有延时。GMM 的响应时间很长，实时性最差。背景差分法介于混合差分与 GMM，随着时间的推移，它的响应时间会加快。

综上所述，相比以往基于差分的方法<sup>[13]</sup>，本文基于区域的跟踪系统不会出现漏检、多检的情况，大大提高了视频跟踪中车辆识别的准确率。相比 GMM 方法<sup>[4]</sup>，本文的跟踪系统具有更强的抗噪能力，而且在响应时间上有着非常明显的优势。

## 7 结束语

本文通过改进基于背景差分的检测和车辆识别方法，实

现了一种基于区域的视频车辆跟踪系统。在 ROI 提取部分，设计了混合差分法，获得了比传统差分法质量更高的检测二值图，且比 GMM 方法更快速、抗噪能力更强。在车辆检测和识别部分，通过堆栈搜索和车辆识别规则设计，克服了以往方法因分割对象而造成误识别的缺点，大大提高了车辆识别的准确率。测试实验的直观图示和总体统计数据验证了本系统的有效性和实用性。

进一步的研究是将系统运用于视频车辆的行为跟踪以及视频交通事件的检测等。

## 参考文献

- [1] Tomizuka M. Automated Highway Systems: An Intelligent Transportation System for the Next Century[C]//Proceedings of the IEEE International Symposium on Industrial Electronics. [S. l.]: IEEE Press, 1997.
- [2] 左军毅, 潘 泉, 梁 彦, 等. 基于模型切换的自适应背景建模方法[J]. 自动化学报, 2007, 33(5): 467-473.
- [3] Zhang Hongjiang, Gong Yihong. Moving Object Detection, Tracking and Recognition[C]//Proc. of the 3rd International Conference on Automation, Robotics and Computer Vision. Singapore: IEEE Press, 1994: 18-26.
- [4] Stauffer C, Grimson W. Adaptive Background Mixture Models for Real-time Tracking[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Fort Collins, Colorado, USA: IEEE Press, 1999: 245-252.
- [5] Zivkovic Z. Improved Adaptive Gaussian Mixture Model for Background Subtraction[C]//Proceedings of the 17th International Conference on Patter Recognition. Cambridge, United Kingdom: IEEE Press, 2004: 28-31.
- [6] Park Y. Shape-resolving Local Threshold for Object Detection[J]. Pattern Recognition Letters, 2001, 22(5): 883-890.
- [7] Badenas J, Sanchiz J M, Pla F. Motion-based Segmentation and Region Tracking in Image Sequences[J]. Pattern Recognition, 2001, 34(8): 661-670.
- [8] Ridder C, Munkelt O, Kirchner H. Adaptive Background Estimation and Foreground Detection Using Kalman-filtering[C]//Proceedings of International Conference on Recent Advances in Mechatronics. Istanbul, Turkey: [s. n.], 1995.
- [9] Zhong Jing, Sclaroff S. Segmenting Foreground Objects from a Dynamic Textured Background via a Robust Kalman Filter[C]//Proceedings of IEEE International Conference on Computer Vision. Nice, France: IEEE Press, 2003: 44-50.
- [10] Nummiaro K, Koller-Meier E, Van Gool L. An Adaptive Color-based Particle Filter[J]. Image and Vision Computing, 2003, 21(1): 99-110.
- [11] 施 华, 李翠华, 韦凤梅, 等. 基于像素可信度和空间位置的运动目标跟踪[J]. 计算机研究与发展, 2005, 42(10): 1726-1732.
- [12] 付 萍, 方 帅, 徐心和, 等. 视频监控系统中运动目标检测的阴影去除方法[J]. 计算机工程, 2007, 33(10): 22-24
- [13] Hariharakrishnan K, Schonfeld D, Raffy P, et al. Video Tracking Using Block Matching[C]//Proceedings of IEEE International Conference on Image Processing. [S. l.]: IEEE Press, 2003.