

# 基于视频的交通目标跟踪方法研究

雷云,王夏黎,孙华

(长安大学信息工程学院,陕西西安710064)

**摘要:**交通目标的检测是智能交通系统(ITS)中的一项关键技术,基于视频跟踪方法的检测技术是目前研究的热点。介绍了近年来提出的一些主要的基于视频的运动目标跟踪方法,对各种方法进行了归类,并分析比较了这些方法的优缺点。在此基础上,着重介绍了一种快速运动目标跟踪方法——MeanShift算法。该算法主要利用图像的颜色统计直方图作为特征,利用 Bhattacharyya 距离作为目标匹配相似性测度,采用梯度优化方法完成对运动目标的快速跟踪。该方法非常适合对交通目标的跟踪。

**关键词:**智能交通系统;视频图像处理;目标跟踪;MeanShift 算法

中图分类号:U491.1<sup>+</sup>16

文献标识码:A

文章编号:1673-629X(2010)07-0044-04

## The Research about Transport Target Tracking Based on Video

LEI Yun, WANG Xia-li, SUN Hua

(School of Information Engineering, Chang'an University, Xi'an 710064, China)

**Abstract:** The detection of traffic target is a key technology in the intelligent transportation systems (ITS), and at present, the detection technology based on video tracking is a research topic of much interest. Some of the major video-based tracking methods about moving targets in recent years are introduced in this paper, which are classified and analyzed the advantages and disadvantages. On this basis, focus on a fast moving target tracking method - MeanShift algorithm. The algorithm mainly uses the color curve and surface of the image as the main feature, and use Bhattacharyya distance as a similarity measure standard to match the target, and at last use gradient optimization method to complete the fast tracking of moving targets. This method is suitable for tracking targets on traffic.

**Key words:** ITS; video image processing; target tracking; MeanShift algorithm

### 0 引言

随着经济的发展和社会的进步,对交通场景的监控已经成为一个十分重要的课题。现如今通过摄像机来监控交通场景,早已被广泛应用于社会生活当中,但是在很多场合下,由人来分析视频序列是不合适的。而基于计算机视觉和图像处理的视觉监控技术提供了一种更加先进和可行的方法,其广泛的应用前景和潜在的商业价值已经引起了国内外学术界和企业界的极大关注。

文中将对视觉交通监控中的交通目标跟踪方法进行研究。

### 1 常用的目标跟踪算法

目标跟踪,就是通过分析然后在视频序列的每幅图像中找到所感兴趣的运动目标的位置,建立起运动目标在各幅图像中的联系。简单的说就是在视频序列的每一幅图像中对运动目标进行定位。基于视频或图像序列的交通目标跟踪长期以来都是计算机视觉、数字视频与图像处理和模式识别领域中一个非常重要和活跃的研究课题。

目标跟踪的作用非常重要,因为它是衔接运动目标检测和目标行为分析与理解的一个重要环节。在实际应用中,交通目标跟踪不仅可以提供目标的运动轨迹、运动参数和准确的位置,也为进行场景中运动目标的运动分析和场景分析提供了可靠的数据来源,同时运动目标的跟踪信息也反过来为运动目标的正确检测以及运动目标的识别提供了帮助,从而更有利于交通运动目标的跟踪。

当前流行的跟踪算法有的是针对刚体目标(例如

收稿日期:2009-10-27;修回日期:2010-01-29

基金项目:西安市科技攻关项目(GG06014)

作者简介:雷云(1983-),男,山西中阳人,硕士研究生,研究方向为数字图像处理、智能交通;王夏黎,副教授,硕士生导师,研究方向为图形图像处理与智能交通系统。

汽车)的跟踪,有的是针对非刚体(行人)运动目标的跟踪,归纳起来,主要有基于光流、变形模型、区域和特征的四种跟踪方法。

### 1.1 基于光流的跟踪

光流的概念是 Gibson 于 1950 年首先提出的,所谓光流场是指序列图像中由运动产生的瞬时位移场。连续光流法是通过提取运动目标的光流场,在经过一系列光流场区域处理,对相近速度区域进行聚类,较完整地提取运动目标区域。从含运动目标的图像序列中抽取光流场,筛选出光流较大的运动目标区域并计算目标的速度矢量,从而实现运动目标的跟踪。主要包括图像预处理、初始光流场计算、光流场改进计算、光流区域聚类几个部分。虽然基于光流的方法不需要进行连续图像间特征的匹配,但它也存在着一些缺点。首先,光流的计算需要微分运算,而图像的微分运算对噪声很敏感;其次光流的计算常用松弛迭代算法,运算量大算法费时,难以满足实时控制的要求,对有较大速度的交通目标跟踪不太理想。

### 1.2 基于变形模型的跟踪

在跟踪过程中最典型的变形模型就是 1987 年 Kass<sup>[1]</sup>等人提出 Snake 的主动轮廓模型。Snake 模型是一种有效的分割和跟踪工具,它可以用来检测目标边缘和跟踪运动目标。

Snake 的分割和跟踪是通过能量最小化的原则进行的。Snake 模型依赖于图像中细微的变化,因为它的解决方案是建立在变化基础之上。2001 年 Won<sup>[2]</sup>等人提出了 Snake 的跳跃模型,当序列图像中连续两帧图像不存在目标重叠现象时,Snake 的跳跃模型就可以用来跟踪物体,这种 Snake 的跳跃模型理论是假设在图像流的处理中能够获得运动方向的基础上的,从先前每一帧图像中获得的 Snake 的节点跳跃到目标的边界,在确定目标分割半径的基础上,位移到另一个区域并且重新初始化。通过图像流信息的反复分割和跳跃来达到能量最小化。

Snake 的跳跃模型能有效解决快速运动物体的跟踪问题。Snake 的跳跃模型使用了摄像机拍摄到的多帧连续图像。Snake 轮廓由  $N$  个节点组成,再由  $N$  个节点计算出一个目标中心点。图像分割过程通过分割汽车的边界完成。在图像序列中使用了 Snake 的跳跃模型,图像流信息被用于 Snake 轮廓跳跃,操作是 Snake 能量的最小化的过程。能量的最小化是通过搜索过程中每一步能量的比较,得出能量变化小于阈值时结束。能量最小化的方法能够相当有效地跟踪出车辆。Snake 的跳跃模型用来跟踪运动目标的前提是序列图像中连续两帧图像不存在目标重叠现象。但由于

Snake 模型存在初始化轮廓问题和对图像噪声以及车辆的遮挡现象非常敏感,所以该方法需要与别的方法结合着使用。

### 1.3 基于区域的跟踪

基于区域的跟踪算法的基本思想是:首先得到包含目标的模板,该模板通过图像分割获得或是预先人为确定,模板通常为略大于目标的矩形,也可为不规则形状;然后在序列图像中,运用相关算法跟踪目标,对灰度图像可以采用基于纹理和特征的相关,对彩色图像还可利用基于颜色的相关。最常用的相关准则是平方和准则 SSD(the sum of squared differences),该算法还可和多种预测算法结合使用,如线性预测、二次曲线预测、卡尔曼预测等,以估计每幅图像中目标的位置。这种算法的优点在于当目标未被遮挡时,跟踪精度非常高,跟踪非常稳定。但它也存在着很大的局限性,其缺点首先是费时,当搜索区域较大时情况尤其严重;其次,算法要求目标变形不大,且不能有太大遮挡,否则相关精度下降会造成目标的丢失。由于实际应用中车辆自身的影子以及车辆之间的遮挡问题普遍存在,所以区域跟踪法还需要做进一步的改进。

### 1.4 基于特征的跟踪

用于目标的跟踪的个体特征有许多,不管是刚体运动目标,还是非刚体运动目标,在序列图像中相邻的两帧图像,由于图像序列间的采样时间间隔很小,可以认为这些个体特征在运动形式上具有平滑性,因此可以用直线<sup>[3]</sup>、曲线<sup>[4]</sup>、参照点<sup>[5]</sup>等个体特征来跟踪运动目标。Liu<sup>[3]</sup>等人介绍了灰度图像中一种边缘直线匹配的算法。在边缘直线的提取中,首先,用图像边缘聚焦技术处理图像数据,消除不必要的图像噪声,形成了一个边缘,然后从边缘中分割出直线,并从中提取直线。用一种以直线的几何关系和灰度图像的信息为基础的匹配函数描述了两幅图像边缘直线的相似性,在连续帧图像中采用直线匹配的方法进行了运动参数的估计。

●基于特征的跟踪方法有其显著的优点:

- 1)由于使用的符号模型运动方式简单,运动具有平滑性,因此跟踪目标的算法就简单了;
- 2)这种方法已经假设特征符号运动是相互独立的运动,因此在运动分析时可以不区分运动物体是刚体还是非刚体,也不用管它的几何形状;
- 3)跟踪过程中符号特征容易捕捉,能够匹配到每一个特征符号。

●但是,基于特征的跟踪方法也有其致命的缺点<sup>[6]</sup>:

- 1)伴随着复杂运动的简单运动,刚体运动目标的

特征提取就会产生困难,如圆柱旋转式运动时,运动目标不可能是匀加速运动,更不可能是匀速运动;

2)运动初始化时的难点。刚体的一些特征会因为遮挡而无法识别,因此,基于特征的跟踪算法必须解决目标跟踪过程中的运动初始化的难点,但这些问题的解决又会使跟踪算法变得非常复杂;

3)在改变符号参数和 3D 目标运动参数时,这些参数是非线性的,因此特征跟踪中恢复的 3D 运动参数对噪声相当敏感。

MeanShift 就是一种利用目标本身的特征进行跟踪的算法,因为其计算速度快,而且可以较好地解决目标遮挡的问题,所以近几年变的较为流行,下面将重点介绍下 MeanShift 算法。

### 2 MeanShift 快速跟踪算法

在使用颜色特征进行运动目标跟踪算法分为两大类:参数估计的统计方法和无参估计的统计方法,这两种方法都是对颜色区域建立颜色分布的模型。在参数估计的统计方法中,文献[7]通过对颜色空间(YUV 颜色空间)使用单高斯分布来建立对目标区域的颜色分布模型,然而对目标区域使用单高斯模型限制了目标只能是单色彩的,这样降低了这种方法的适用性。使用统计直方图的无参估计方法可以很好地克服参数估计方法的缺陷,目前广泛地应用在基于颜色直方图统计特征的运动目标跟踪算法当中。MeanShift<sup>[8,9]</sup>是一种较为典型的无参估计目标跟踪算法。

MeanShift 算法利用梯度优化方法来减少特征搜索匹配的时间,实现快速的目标定位,同时利用 Bhattacharyya 距离作为对目标模板和候选目标的相似性测度,完成特征的匹配,并利用图像的颜色统计直方图作为整个搜索匹配的特征。

#### 2.1 理论概述

MeanShift 算法是一种半自动跟踪方法,在起始跟踪帧,通过手动或检测方法确定搜索窗口来选择运动目标。计算核函数加权下的搜索窗口的直方图分布,用同样的方法计算当前帧对应窗口的直方图分布,以两个分布的相似性最大为原则,使搜索窗口沿密度增加最大的方向移动到目标的真实位置。

MeanShift 算法流程图如图 1 所示。

下面介绍目标模型和候选模型的建立及其 MeanShift 跟踪过程。

##### 2.1.1 目标模型

首先是确定跟踪目标区域,一般可以手动选定。假定目标区域的中心为  $x_0$ ,那么,对目标模型可以描述为在目标区域上的所有特征值的概率值。目标特定

值设为  $u = 1, 2, \dots, m$ ,则该物体可以表示为:

$$\hat{q}_u = C \sum_{i=1}^n k(\| \frac{x_i - x_0}{h} \|^2) \delta[b(x_i) - u] \quad (1)$$

式中  $x_0$  是搜索窗口( $n$  个像素)的中心像素坐标,  $x_i$  是第  $i$  个像素的坐标;  $k(\|x\|^2)$  是核函数,  $h$  表示核函数的带宽,一般等于窗口宽度的一半,函数  $b$  和  $\delta$  的作用是判断  $x_i$  处的颜色值是否属于特征值  $u$ ,  $C$  是一个标准化的常量系数,使得所有特征值的概率和为 1。

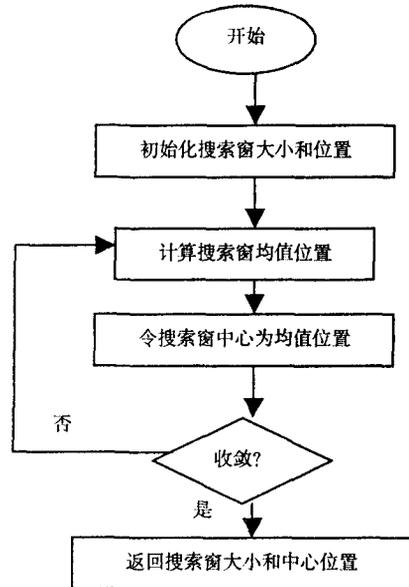


图 1 MeanShift 算法流程图

##### 2.1.2 候选模型

运动目标在当前帧中包含目标的区域称之为候选区域,设其中心坐标为  $y$ ,也是核函数的中心坐标。候选区域特征值  $u = 1, 2, \dots, m$  的概率密度为:

$$\hat{p}_u(y) = C_h \sum_{i=1}^n k(\| \frac{x_i - y}{h} \|^2) \delta[b(x_i) - u] \quad (2)$$

##### 2.1.3 相似性度量

相似性函数描述目标模型和候选模型之间的相似程度,在理想情况下两个模型的概率分布是完全一样的。在 MeanShift 算法中,采用 Bhattacharyya 系数来作为相似性函数度量分布,即:

$$\hat{\rho}(y) = \rho[\hat{p}(y), \hat{q}] = \sum_{u=1}^m \sqrt{\hat{p}_u(y) \hat{q}_u} \quad (3)$$

$\hat{\rho}(y)$  的值在 0 ~ 1 之间,  $\hat{\rho}(y)$  值越大表示两个模型越相似。

##### 2.1.4 目标定位

寻找最优匹配目标可划为使  $\hat{\rho}(y)$  最大,在当前帧的目标中心先定义为前一帧中目标中心的位置(用  $y_0$  表示),从这一点开始寻找最优匹配的目标,式(3)

在  $\hat{p}_u(y)$  点泰勒展开可得:

$$\rho[\hat{p}(y), \hat{q}] \approx \frac{1}{2} \sum_{u=1}^m \sqrt{p_u(y_0) q_u} + \frac{C_h}{2} \sum_{i=1}^n w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right) \quad (4)$$

其中:

$$w_i = \sum_{u=1}^m \delta[b(x_i) - u] \sqrt{\frac{\hat{q}_u}{\hat{p}_u(y_0)}} \quad (5)$$

式(5)右边的第二项就是有权值的核密度估计,可以利用 MeanShift 法进行最优化,在每次 MeanShift 迭代中,目标区域的中心位置都会由  $y_0$  移动到新的位置  $y_1$ :

$$y_1 = \frac{\sum_{i=1}^n x_i w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)}{\sum_{i=1}^n w_i k\left(\left\|\frac{y - x_i}{h}\right\|^2\right)} \quad (6)$$

MeanShift 算法以  $y_0$  为起点,以两个模型相似比最大的方向移动,具体迭代过程如下:

- (1) 初始化目标模板  $q_u$ ;
- (2) 在当前帧中  $y_0$  处,计算  $\{p_u(y_0)\}_{u=1, \dots, m}$ , 由公式(4) 计算目标模型和候选模型的相似比系数;  $y_0 \leftarrow y_1$
- (3) 计算权值  $w_i$ ;
- (4) 由公式(6) 寻找下一个位置  $y_1$ , 计算  $y_1$  处的候选目标模型,重新计算  $\rho(y_1)$ ;
- (5) 若  $\rho(y_1) < \rho(y_0)$ ,  $y_0 \leftarrow 1/2(y_1 + y_0)$ , 若  $\|y_1 - y_0\| < \epsilon$ , 停止计算; 否则  $y_0 \leftarrow y_1$  转到第二步计算。

MeanShift 搜索当前目标位置的过程概括为:首先在跟踪区域内,根据先前帧位置计算目标模型,当前帧建立初始候选模型,然后利用加权的 MeanShift 迭代不断改变候选模型位置,最后找到目标中心位置,从而确定当前帧中目标的位置。

## 2.2 算法分析

MeanShift 算法用于目标跟踪有几个优点:首先,该算法本身计算量不大,在目标区域已知的情况下完全可以做到实时跟踪;其次,MeanShift 算法作为一个无参数密度估计算法,很容易作为一个模块和别的模块的算法集成;另外采用核函数直方图建模,对边缘遮挡、目标旋转、变形和背景运动不敏感。这些优势使得 MeanShift 算法广泛地应用于目标跟踪领域。

但是 MeanShift 算法本身也存在着一些缺点:一是跟踪过程中不能自动调整跟踪窗口的大小,当目标有尺寸变化时,可能会造成跟踪失败;二是直方图是一种比较弱的对目标特征的描述,当背景和颜色分布相

似时,算法效果欠佳;三是当场景中目标运动速度很快时,目标区域在相邻帧间会出现没有重叠的区域的情况,目标这时一般收敛于背景中与目标的颜色分布相似的物体,而不是场景中的目标,通常采用卡尔曼滤波对目标位置进行预测后再运用该算法可以取得较好的效果。

## 3 结束语

文中介绍了基于视频的目标检测技术中常用的一些跟踪方法,分析比较了各方法的特点。在此基础上,文章还着重介绍了 MeanShift 算法的原理以及算法本身的优缺点。

在视觉跟踪领域,MeanShift 算法是一个非常优秀的算法,是国外最近几年才发展起来的,对于具有交通特性的运动目标的跟踪,该方法是适合的。

## 参考文献:

- [1] Kass M, Witkin A, Terzopoulos D. Snakes: Active contour models[J]. International Journal on Computer Vision, 1998(4):321-331.
- [2] Kim W, Lee C Y, Lee J J. Tracking moving object using Snake's Jump based on image flow[J]. Mechanics, 2001(11): 119-216.
- [3] Liu Y, Huang T S. Determining straight line correspondences from intensity images[J]. Pattern Recognition, 1991, 24(6): 489-504.
- [4] Hang Z, Faugeras O D. Three dimensional motion computation and object segmentation in a long sequence of stereo frames [J]. International Journal on Computer Vision, 1992(3): 211-241.
- [5] Ferruz J, Ollero A. Integrated real time vision system for vehicle control in non-structured environments[J]. Engineering Applications of Artificial Intelligence, 2000(13): 215-236.
- [6] Jung S K, Wohn K Y. A model based 3D tracking of rigid objects from a sequence of multiple perspective views[J]. Pattern Recognition on Letters, 1998(19): 499-512.
- [7] Huang J, Kumar S, Mitra M, et al. Spatial Color Indexing and Applications[J]. Computer Vision, 1999, 35(3): 245-268.
- [8] Comaniciu D, Ramesh V, Meer P. Real-Time Tracking of Non-Rigid Objects Using Mean Shift [C] // IEEE Conf. Computer Vision and Pattern Recognition. [s. l.]: [s. n.], 2000: 142-149.
- [9] Cheng Yizong. Mean Shift, Mode Seeking, and Clustering[J]. IEEE Transactions on Patterns and Machine Intelligence, 1995, 8(17): 790-799.