

智能交通系统中基于视频图像处理的 车辆检测与跟踪方法综述*

王圣男¹, 郁梅^{1,2}, 蒋刚毅^{1,2}

(1. 宁波大学 信息与工程学院, 浙江 宁波 315211; 2. 北京大学 视觉与听觉信息处理国家重点实验室, 北京 100871)

摘要: 与传统的车辆检测器相比, 基于视频图像处理与视觉技术的车辆检测器具有处理速度快、安装维护便捷且费用较低、可监视范围广、可获取更多种类的交通参数等诸多优点, 因而近年来在智能交通系统(ITS)中得到了越来越广泛的应用。针对摄像头拍摄得到的交通序列图像, 人们提出了很多视频图像处理和分技术, 其中最基本的研究领域就是交通场景中车辆对象的检测与跟踪。介绍了近年来提出的一些主要的车辆检测与跟踪技术, 并根据核心处理方法(基于特征、区域或模型等)及处理域(空域、时域)的不同对这些技术进行了分类, 同时分析比较了各种方法的优缺点。最后, 说明了这一领域仍然存在的问题和对可能的研究方向进行了一定的预测。

关键词: 智能交通系统; 交通监视系统; 车辆检测; 车辆跟踪

中图法分类号: TN919.81 文献标识码: A 文章编号: 1001-3695(2005)09-0009-06

Review on Vehicle Detection and Tracking Techniques Based on Video Processing in Intelligent Transportation Systems

WANG Sheng-nan¹, YU Mei^{1,2}, JIANG Gang-yi^{1,2}

(1. College of Information Science & Engineering, Ningbo University, Ningbo Zhejiang 315211, China; 2. State Key Laboratory of Machine Perception, Peking University, Beijing 100871, China)

Abstract: Compared with traditional traffic detectors, the video sensor has lots of advantages such as fast response, easy installation and maintenance, the ability to monitor wide areas and obtain more kinds of traffic parameters, and as a result, it has been widely used in Intelligent Traffic System(ITS) in recent years. Up to now, a number of video processing and analyzing methods have been proposed for vehicle detection and tracking. In this paper, vehicle detection and tracking techniques are introduced and categorized based on their key processing techniques (feature, area or model) and processing domain (spatial or temporal), the advantages and disadvantage of each method are also analyzed. Additionally, the problems that still exist and the trend of the research in this field in the future are indicated.

Key word: Intelligent Transportation Systems(ITS); Traffic Monitoring Systems; Vehicle Detection; Vehicle Tracking

1 引言

智能交通系统(ITS)是将先进的信息技术、数据通信传输技术、电子传感技术、电子控制技术、计算机处理技术等有效地集成运用于整个交通管理而建立的一种在大范围内、全方位发挥作用的、实时、准确、高效的综合交通运输管理系统^[1]。交通监测系统是智能交通系统的重要环节,负责采集有关道路交通流量的各种参数。交通环境的车辆检测研究可以追溯到20世纪70年代。传统车辆检测器如磁感应线圈有着诸多缺点和局限,鉴于这种情况,人们不断提出新的替代方案,如采用雷达、超声波、红外线、微波、声频及视频图像等技术的悬挂式传感器。近年来随着计算机和图像处理技术的不断发展,利用

机器视觉检测器来进行车辆检测成为一种特别有潜力的替代方法,有望取代传统检测器成为现代智能交通系统的一个重要组成部分。

1978年,美国JPT(加州帕萨迪纳市的喷气推进实验室)首先提出了运用机器视觉来进行车辆的检测的方法,指出其是传统检测方法的一种可行的替代方案。几年后,美国明尼苏达大学的研究人员研制了第一个可以投入实际使用的基于视频的车辆检测系统。该系统使用了当时最先进的微处理器,在不同场景和环境下的测试结果良好^[2],说明了利用视频传感器实时检测车辆的可行性。同期基于视频的车辆检测的研究也在欧洲和日本广泛展开^[3-5]。

此后十年间基于视频的车辆检测技术取得了长足的进步。1991年,美国加州理工大学对在高速公路上运用视频方法的检测技术进行了评估,在评估报告中对当时采用的不同的视频车辆检测技术详尽地进行了分类。三年后,美国休斯飞机公司评测了当时存在的几种检测技术,包括视频检测技术,测评结果指出基于视频图像处理的车辆检测系统已经具备了投入实

收稿日期: 2004-09-12; 修返日期: 2004-11-12

基金项目: 浙江省自然科学基金青年人才基金(RC01057); 浙江省自然科学基金(601017); 宁波市重点博士基金(2003A61001, 2004A610001); 人事部留学回国人员科研启动基金项目

际使用的潜力。1994 年 Mn/DOT(明尼苏达运输部)为 FHWA(美国联邦公路局)进行了更详尽严格的测评,结果表明视频检测器的检测准确性和可靠性可以达到令人满意的程度。同时随着视频车辆检测技术的发展,人们已不满足于仅仅检测出车辆, FHWA 进一步利用此技术来提取交通参数,如交通流量,十字路口的车辆转向信息等^[9]。事实上,与其他几种车辆检测方法相比,基于视频图像技术的方法具有直观、可监视范围广、可获取更多种类的交通参数以及费用较低等优点,因而可广泛应用于交叉道口和公路干线的交通监视系统中^[6]。

在交通监测系统中,视频检测的传感器(即摄像头)被安放在道路的上方以获得道路和过往车辆信息,安装的高度一般在 5m~6m,以保证对整个交通场景有很好的视点,且得到的视频图像序列可以为车辆检测和跟踪提供足够的信息。

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、车辆跟踪等三个模块,流程如图 1 所示。首先由摄像头拍摄得到实时交通场景的视频序列图像,然后对序列图像进行 ROI 提取,并将提取到的 ROI 送到车辆检测模块以根据一定的图像处理方法和准则判断某 ROI 区域是否为车辆。检测出车辆后,可在跟踪模块对车辆进行跟踪。由检测和跟踪的结果可以分析提取出交通流量参数,如车速、车流密度、转向信息等。这类实时道路交通信息及各种服务信息汇总到交通管理中心并经集中处理后,将传送至公路交通系统的各个用户,使公众能够高效地使用公路交通设施,从而到达提高道路负载能力和行驶效率以及节约能源等目的。

交通场景中车辆对象的实时检测与跟踪是基于视频的交通监测系统中最重要也是最基本的步骤,是视频检测法的核心,检测与跟踪的正确与否直接关系到智能交通系统决策的正确性。针对视频图像传感器所得到的交通序列图像,人们提出了许多视频图像处理和分析方法。

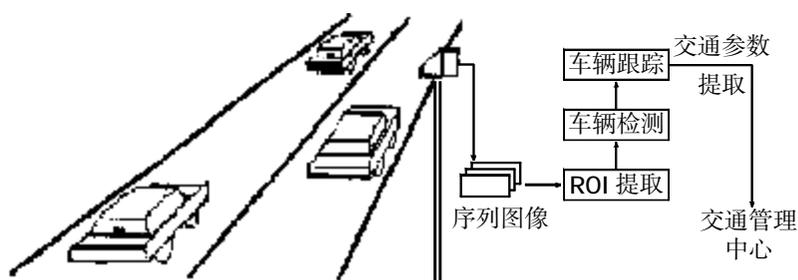


图 1 车辆检测与跟踪系统流程图

2 ROI 提取模块

ROI 提取的目的是对交通场景视频图像中的车辆进行粗分割,将有可能是车辆的区域从复杂的交通场景中分割出来以便进行后续的检测与跟踪操作。其方法包括对视频图像序列的时域、空域及时-空域分析,但其中利用时域分析重于利用空域分析;换言之,对一个像素,检测主要是利用了视频图像序列帧中处于同一位置的像素信息,而不是仅仅利用一帧中该像素的周围像素信息^[7],这主要是由视频图像的特点决定的。对 ROI 的提取主要有帧差法、背景差法和非监督分割法等几类方法。

2.1 帧差法

帧差法将两帧(或多帧)连续图像逐像素相减,以去除静止或移动缓慢的物体及背景^[8],它是消除两帧连续图像中的静止物体以及提供运动物体(车辆)运动轨迹最直接的方法。

除了最简单的逐像素相减,帧差图像还可以由两组属于相邻图像帧的像素(如相邻的四个元素)的均值相减得到。

帧差法的优点是计算简单且不易受环境光线变化的影响,但它不能检测静止车辆,且处理效果与图像采样频率以及被检测车辆的车速有关。如果视频检测器采样频率过小,而车速较快,可能会造成误分割;反之如果采样频率过大且车速较慢,又会造成过度覆盖,极端情况下运动物体可能完全重叠,类似于静止车辆,从而导致无法分割出运动物体。为了克服这些问题, Kameda 采用概率统计的方法通过三帧连续图像间的运动信息找出 ROI^[9]。而在文献[10]中,除了利用连续三帧图像外,还使用单帧图像静止分割的结果来修正帧差法所得到的运动信息。文献[11]结合了背景差法来提高处理的准确性。文献[12]则利用基于空间 Markov 随机场的连续帧图像间的差值统计量来描述帧间差异。

2.2 背景差法

背景差法的基本思想是先形成交通场景的背景图像,然后将待检测图像与背景图像逐像素相减(理想情况下的差值图像中非零像素点就表示了运动物体),进而就可运用阈值方法将运动物体从背景中分离出来^[13],即计算 $D_{i,j} = |C_{i,j} - B_{i,j}|$,其中 $\{C_{i,j}\}$ 和 $\{B_{i,j}\}$ 分别为当前帧和背景图像, $\{D_{i,j}\}$ 为背景差图像,则车辆检测可表示为 $V_{i,j} = \begin{cases} 1 & \text{若 } D_{i,j} > T \\ 0 & \text{其他} \end{cases}$,这里 $V_{i,j} = 1$ 表示对应像素为车辆区域,否则属背景区域。

精确可靠的背景图像是背景差法能否成功提取 ROI 的关键。背景图像可由人工拍摄一幅没有车辆的图像来得到,也可以通过序列图像的平均来得到。显然,建筑物阴影、浮云或光照的变化都会造成背景环境明显的变化;由于这些环境变化因素,作为参照物的背景需要定时更新。目前有多种背景更新方法,最常用的方法是多帧平均(Frame Averaging)法^[14]和选择更新(Selective Updating)法^[15]。

最简单的多帧平均法是直接将前面若干输入图像的平均值作为当前背景。文献[14]采用如下方式更新背景:

$$B_{\mu} = K B_{\mu-1} + (1 - K) C_{\mu-1}, 0 < K < 1 \text{ 且 } K = N / (N + 1) \quad (1)$$

其中, B_{μ} 为当前更新的背景, $B_{\mu-1}$ 是前一帧背景, $C_{\mu-1}$ 为上一幅场景图像, N 为一整数, K 是指更新率。 K 越趋向 1, 新背景就越接近于上一幅背景。多帧平均法的缺点在于对场景改变过于敏感性,而这种改变并不总是由车辆运动引起的,同时 K 的正确取值也是个难题。

选择更新法的基本思想是只将没有检测到运动的区域即真正的背景进行更新。文献[22]在背景更新前,先逐像素判断相邻两幅图像的差值,如果小于一个阈值,则说明在这个像素位置没有检测到车辆,可进行背景更新操作。这种方法依赖于阈值的选取,如果阈值选择不合理,背景图像将很快变得不可用。文献[17]在选择更新背景前进一步加入了背景调整步骤,即首先计算当前帧和上帧图像间的平均灰度之差,若差值大于设定阈值,则认为当前帧的平均灰度发生突变,就要在背景图像上逐像素加上此差值。

基于背景差的车辆检测的准确性很大程度上依赖于背景图像的可靠性。总的来说,背景差法对环境光线的变化非常敏感,背景图像需不断地被更新以迎合环境光线、阴影和天气的变化等,因而背景更新中的误差累计是影响背景差法精度的重要因素。

2.3 非监督视频分割法

非监督视频分割法就是在不需要人为操作的情况下自动将视频序列图像分割为代表不同物体的连通区域的技术,因而也可以运用到基于视频的车辆检测中。

2.3.1 k -均值聚类法

k -均值聚类法是按照一定的聚类准则将整个数据集分成 k 个聚集的算法。在视频交通检测应用中,利用 k -均值法就是将拍摄到的交通场景中的像素聚类为 k 个表示不同物体的区域块。Kottle 在文献[16]中提出了一种 k -均值法的分割算法,其聚类准则是利用了像素的三个特征:纵横坐标和灰度值,其中灰度值可以将属于不同物体的像素分离,而坐标则决定了像素的归属区域。然而这种方法的缺点是分割区域的数目要作为初始参数,显然在复杂的交通场景中,事先预测可以正确分割图像的区域数目是比较困难的,因为这取决于场景中的车辆数目。为了解决这个问题,文献[18]提出了一种多段(Multi-stage)分割方法,事先固定初始分割区域数目(6~10个),然后根据初始分割区域的密度和大小来决定是否对此区域继续划分。

k -均值法产生的最后分割区域数目不是固定的,产生的每一个区域块属于单一物体,但一个物体却可能包含多个区域。这就需要后续的融合操作将多个区域合并为单个的 ROI。由于不同物体间存在相对运动,而属于同一物体的区域则不存在这种相对运动,文献[19]先利用了这一特征来初步合并区域,再从时域分析区域块的运动矢量进一步合并得到 ROI。

k -均值法是一种有效的分割方法,但是它需要事先知道 k 的值,同时在聚类过程中涉及到迭代运算,这使其在实时性要求极高的交通检测应用中受到了一定的限制。

2.3.2 SPCPE(同步分割与类别参数估计)法

SPCPE(Simultaneous Partition and Class Parameter Estimation)法也是一种无监督视频分割方法^[20],其基本思想是假设已知待分割图像中物体的类别(这在交通检测中是可行的,一般交通场景中包含车辆和道路两种类别),用给定模型描述类别,在给定光流数据的条件下,迭代搜索分割标记的最大后验概率,使得当前分割与期望分割符合的程度最大,得到最优的分割结果。

在 SPCPE 法中,一个给定的类别描述决定了一种分割情况。类似的,一个给定的分割也产生一个类别描述,所以分割和类别参数必须同步估计。SPCPE 法中第一帧图像的初始划分是随机的,然后迭代寻找最大后验概率来进行分割。由于视频图像序列的连续性,可以利用前一帧分割的结果作为当前帧图像的初始分割来减少计算量。为了进一步减少计算的开销,文献[21]利用背景差法所得到的结果作为第一帧的初始划分。

3 车辆检测模块

3.1 阈值法

车辆是致密的运动物体和周围静止背景的灰度值存在差异。这样通过设定阈值就可以将车辆从背景中提取出来。最简单的车辆检测方法就是对提取的 ROI 区域进行阈值处理,高于阈值的像素属于运动物体(车辆),反之就是背景。然而不恰当的阈值选取会造成车辆的误检测,阈值过高会造成漏检,阈值过低又会把背景检测为车辆或将相邻车辆检测为同一辆车,所以在判断 ROI 是否为车辆时阈值的选取至关重要。

阈值选取分为全局阈值和局部阈值选取。

文献[22]提出了一种二维最小平方方法(Two-dimensional Least Square)的全局最优阈值选取方法。但全局阈值选取可能会造成相邻的车辆相连,给车辆检测带来困难。而且如果车辆的灰度值与背景差异不大,则很可能会造成漏检。

文献[23]先找出一个阈值平面(Threshold Surface),每个阈值平面对应一个阈值,这样就可以通过可变的阈值去适应背景和光照环境的改变,但它只利用了局部的直方图分布而没有考虑分割形状,不能保证准确地检测重叠车辆。文献[24]提出了用于背景差图像二值化的考虑形状分解的局部阈值法。这种方法可以分解出重叠的车辆。图2是这种方法的一个实例,图中横坐标是递减的多重阈值,纵坐标为比值,对应致密度和区域变化率。在不同的阈值层次下首先显示出来的是种子区域,即有可能是车辆的区域。令 P_k 表示第 k 个可能是车辆的种子区域, A_{ki} 表示以 P_k 为种子在第 i 个阈值时的区域,一系列的阈值层次的区域变化率定义为 $\rho_{ki} = A_{k(i+1)} / A_{ki}$ 。则在比值最大时的情况下就可以阻止相邻车辆的融合;同时该方法还有一个致密性准则,令 N_{ki} 表示 A_{ki} 的像素数, M_{ki} 表示完全包括 A_{ki} 区域的矩形框的像素数,则致密度 δ_{ki} 定义为 N_{ki} / M_{ki} 。最后对每个种子区域找出满足约束 $\delta_{ki} > \delta_1$ 条件下的最大值就可以得到每个区域的局部最优阈值,这里 δ_1 为一常数,决定了检测的灵敏度。在对交通场景图像进行车辆检测时,先求出背景差图像,然后对背景差图像及其反转图像分别采用上述形状分解局部阈值法进行二值化以分别寻找亮的和暗的车辆区域,再将两个结果合并从而得到最终结果。这种方法的缺点在于不能避免对阴影的误检测,且容易丢失与环境灰度值相近的部分车体。

另外,作为阈值选取法的辅助手段,文献[25]提出了利用二值数学形态学将相对分散的像素聚集为致密的物体对象。

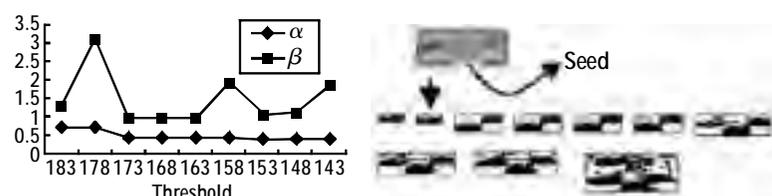


图2 形状分解局部阈值实例

3.2 检测线法

检测线法就是在待检测图像上的合适位置设置检测线,检测线的作用类似于埋于地下的感应线圈传感器,它的方向与车辆行驶方向垂直,如图3所示。保存位于检测线位置上的参考图像(背景图像或其边缘检测图像),当车辆经过检测线时,检测线位置上的图像就会由于车辆的覆盖而发生改变,当运动物体覆盖检测线的宽度大于某个阈值时,就认为检测到一辆车辆^[27]。文献[27]中采用的是固定的人工设置的宽为三个像素的检测线。

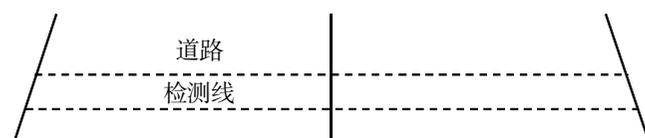


图3 检测线法示意图

文献[28]利用道路上的白色字符作为路面标记以检测车辆并排除阴影干扰,其原理和检测线法是相同的。图4(a)、图4(b)中白色字符标记或未被任何物体所覆盖,或位于运动或静态阴影中,或被车辆所覆盖。在获得输入图像路面标记区域

如图 4(c)、(d) 所示的二值边缘图像后, 将其与参考边缘图像进行匹配, 以检测标记是否可见。在有车辆经过路面标记区域的情况下, 匹配值会大于一个阈值, 从而检测出车辆。

检测线法是一种简单有效地检测车辆的方法, 在高速公路这样的交通场景可以取得较好的检测结果, 但不太适用于行驶行为复杂的场合如十字路口这样的交通场景, 能获取的交通参数种类较少。



图 4 基于路面标记的车辆检测

3.3 边缘检测法

车辆目标最基本的特征之一是边缘。基于车辆边缘的边缘检测法通常效率较高, 甚至可以检测出静止车辆。当图像亮度发生变化时表现也较为稳健, 因为边缘信息即使是在各种昏暗的光照环境下仍较为明显。

用于检测车辆边缘的方法有多种, 如梯度算子、Laplacian 算子、Kirsch 算子等。另外由于形态边缘检测法的优良性能, 所以也常常被采用^[29,30], 文献[30]就采用了 SMED 的形态学边缘检测, 表示如下:

$$SMED = D(S(f)) - E(S(f)) \quad (2)$$

其中 $S(f)$ 是对图像 f 进行可分离中值滤波后的结果, $D(\cdot)$ 和 $E(\cdot)$ 分别是膨胀和腐蚀操作。形态边缘检测法检测是否成功很大程度上取决于结构元素的选择, 文献[31]采用了图 5 所示的 3×7 结构元素。

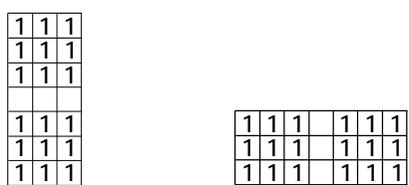


图 5 3×7 结构元素

对图像进行边缘检测所得到的结果通常是一组边缘, 比较稠密的部分一般代表车辆, 而稀疏的部分一般为背景或阴影。所以对车辆的检测就可以依据边缘的稠密程度来差别, 这可以通过直方图来进行分析^[32]。另外, 这些检测出的边缘可以被聚集起来形成车辆的边界。顺着这个方向, 我们必须先制定一个聚集原则, 通过这个原则聚集起来的每一个特征集都表示一个可能的车辆或障碍物, 接着必须识别相关的特征(通常是线段集)。这样车辆或障碍物的检测就可以看成在图像中寻找包含特征集的矩形框的过程。文献[33]提出了检测矩形框方法如图 6 所示, 它先在整幅图中寻找左上角模式(图 6(a)), 对每一个检测到的左上角寻找与之相匹配的左下角(图 6(b)), 然后依次再找右上角(图 6(c))和右下角(图 6(d))。为了减少计算量, 对最后检测到的矩形框数量要进行限制: 矩形框必须大于一个阈值(矩形框大小由透视公式计算); 如果两个矩形框重叠部分大于一个阈值, 则较小的框并入较大的框中。

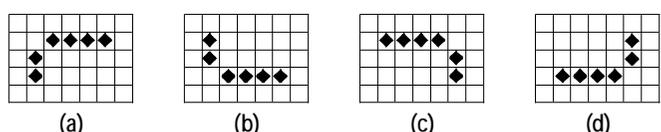


图 6 检测矩形框模式图

为了得到较好的边缘检测结果以利于后续操作, 文献

[34] 利用 Hough 变换来提取车辆轮廓线, 并用形态学方法来恢复轻微断裂的轮廓。

3.4 时域运动估计法

这种车辆检测方法的核心思想是通过时域的运动估计跟踪图像序列中的运动物体来提高分割的准确性, 将车辆跟踪过程和图像分割过程结合起来, 进而减少计算复杂度, 提高系统的实时处理能力。它综合考虑了空域和时域信息。在匹配不同图像帧中表示同一车辆的运动块的过程中, 可以得到被跟踪车辆在 n 帧序列图像中的形态演变, 因而也就有可能预测其在 $n+1$ 帧中的形状, 从而能纠正车辆图像的误分割情况, 如运动块突然出现、消失以及形状突变等。这种方法和其他车辆检测跟踪方法主要的不同在于分割过程和跟踪过程是同步进行的, 而不是在跟踪过程前就有了明确的分割结果^[18,19]。换句话说, 分割结果不仅可以为跟踪过程提供信息, 反过来跟踪过程也可提高分割的准确性。

时域运动估计法中的一个重要步骤是对 ROI 的运动矢量进行估计, 通过一定准则就当前帧中每个 ROI 在下一帧中找出与之最相似的对应 ROI, 这样当前帧的每个 ROI 都可以求得一个运动矢量。运动矢量相近的 ROI 就可以认为属于同一运动物体, 这些 ROI 就可以合并, 从而提高了空域分割的准确性。此外得到的运动矢量还可以被用来进行车辆跟踪及车速估计等。

3.5 模型法

基于模型的车辆检测方法与前述的车辆检测方法相比其优点在于能获得对图像内容的理解^[35-37], 而前述方法实质上都是对待检测图像中的一组像素进行检测和跟踪, 属于非模型方法。换句话说, 非模型方法只是将处理得到的待检测图像中的运动块看作一组像素的集合, 其缺点在于有可能把误分割形成的像素集合也检测为一个车辆对象; 而基于模型的方法将这些像素看作是三维世界中的车辆在二维图像平面上的投影, 经过与预先建立的模型在图像块同一位置的投影相匹配, 可以直接得到车辆的长、宽、高及车辆类型等信息, 它是一个二维到三维的匹配过程。

建立适当的模型是这种方法能够获得对图像内容理解的前提条件, 预先建立的模型通常包括摄像机模型和车辆模型。摄像机模型描述了摄像机与交通场景之间的空间几何特征, 包括摄像头与水平面倾斜的角度、光心的空间位置、焦距等信息; 车辆模型则描述了车辆的先验知识(长宽高等信息)。通常车辆模型的建立采取如图 7 所示的 3D CAD 或线框(Wireframe)模型的方法。

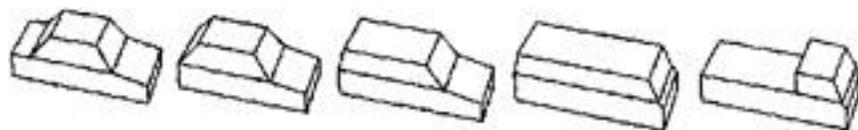


图 7 3D CAD 或线框模型

基于模型的车辆检测方法步骤大致如下:

- (1) 求出 ROI 的中心 (x, y) 和运动方向, 利用摄像机模型将中心点反投影到真实三维世界坐标 (X, Y, Z) ;
- (2) 将 (X, Y, Z) 和 作为车辆模型在真实世界的位置参数, 再利用摄像机模型将车辆模型投影到图像平面上, 产生投影结果, 将投影结果与原 ROI 按一定准则进行匹配;
- (3) 依次将所有的车辆模型投影到图像平面, 找到匹配结果最好的模型, 若匹配度大于一个阈值, 基本就可以认定 ROI

区域对应于匹配度最好的模型,这时就可以利用车辆模型的先验知识直接得到长宽高和类型等信息。

由于模型匹配的固定性,系统必须为每一类车辆都设定一个模板。这就存在一些问题,因为即使是属于同一类的车辆仍可能有不同的几何形状。同时模板方法通常都是假设车辆的亮度特征变化不大,但在实际情况中,环境光照、阴影和车辆重叠以及车身和车窗的光反射都可能造成同一类型车辆的特征差异很大。为了解决这些问题,文献[38]首先用多幅车辆位置已经手工标定的图像对神经网络进行训练,然后用经过训练后的神经网络来识别车辆。此外,利用可变形模型分割和识别车辆也是解决车辆外形差异较大所带来问题的一种可行方法^[39,40]。

虽然模型法能够获得对图像内容的理解,但建立摄像机模型时需要测量详细的摄像机与交通场景之间的空间几何特征,这使这种方法在实际应用中存在着很大的局限性。同时模型法稳健性也不够高,当摄像机由于外力原因产生微小角度变化时就可能造成检测失败,且对遮挡情况下的车辆也会发生误检。

4 车辆跟踪模块

检测出车辆后,车辆跟踪就比较容易了。大多数车辆跟踪方法都遵循一个基本原则,即用空间距离判断两相邻帧中的车辆是否为同一辆车,进而完成时域上车辆的跟踪。空间距离可以是最简单的欧几里德距离,也可以是其他距离标准如 Hausdorff 距离^[41]。跟踪方法大致有以下四类。

4.1 基于模型的方法

这种跟踪方法为模型法检测车辆的后续操作^[35-37]。如前文所述,这种方法的核心是建立的已知车辆对象的精细三维模型与待检测图像之间的匹配操作。其缺点是对车辆模型的过分依赖,而很明显的是不可能为公路上行驶的每种车辆都建立精细的模型。此外这种方法的计算量较大,不利于实时处理。

4.2 基于区域的方法

基于区域的跟踪方法就是在时域上跟踪车辆检测模块检测出的一个个像素连通块,这些块区域表示检测出的车辆。这种方法在车辆稀少时效果很好,且块区域可以提供丰富的信息如大小、形状、密度等^[8,19]。

4.3 基于动态轮廓的方法

动态轮廓模型跟踪方法的主要思想是先初始勾勒出车辆的轮廓并且不断地在后续帧更新轮廓进而达到跟踪的目的^[42,43]。然而基于区域的方法在阴影和道路拥挤的情况下其效果会变得很差,因为阴影和车辆之间的遮挡都会将本来相邻的多个连通块变为一个,造成漏检和误检。这种方法其实是基于区域方法的一个变形,与区域方法相比优点在于计算量低,而缺点是存在初始化困难的问题。它存在和区域法一样的问题,即在阴影和拥塞情况下效果欠佳。

4.4 基于特征的方法

上述方法都是将车辆作为最小单元进行跟踪,而基于特征的方法则是将车辆的特征作为最小跟踪单元。这些特征可以是点^[46]、线^[47]或者曲线^[48]等,这些点、线条可能代表了车辆的保险杠、车窗、车顶棚等。此方法的优点在于即使是在车辆间相互遮挡的情况下,车辆的很大一部分特征还是可见的,可

以为跟踪过程提供依据。但它存在需要特征聚类的问题,即众多的特征中分析哪些是属于同一辆车的。文献[53]采用运动限制作为聚类的依据,即属于同一辆车的特征总是以相同的速度运动,取得较好的效果。

5 存在的问题和展望

视频检测器与传统检测器相比有其明显的优势,近年来在智能交通系统中得到了越来越广泛的应用。虽然视频检测器有着诸多优点,但仍然存在许多需要解决的问题。一个可以投入实际使用的基于视频图像的交通监测系统应该具备处理时间短、计算量低和可靠性高的特点^[6]。而且,这种系统所采用的方法必须对重建3D场景的误差、车辆运动所引起的图像噪声、视频检测器的偏移等干扰有足够的稳健性。实际情况是视频检测器的检测精度是随着光照情况的变化而变化的,当光照良好时如正午时检测精度最好,反之如傍晚、雨雪天气则较差。另外一个问题是阴影问题,阴影是造成视频检测方法误检测的主要原因,阴影通常有三种:车辆自身的运动阴影、道路场景中的静态阴影、缓慢移动的阴影如浮云造成的阴影。阴影问题的解决好坏直接关系到检测结果的正确性。同时车辆在道路场景中的相互遮挡也是必须要考虑的问题。

未来这一领域的发展应该是围绕上述问题的解决而展开。研究近年来人们提出的种种车辆检测方法,我们不难发现由于受限于当时计算机硬件的处理速度,早期的检测方法都较为简单,对实时性的要求要大于对准确率的要求。而随着更快的中央处理器的不断研发,中后期的检测方法则越来越复杂,在一个系统中往往使用了多种车辆检测方法来提高检测的准确率,同时采用了更多新技术如 Kalman 滤波器、模糊判决理论等。所以注重准确率、综合使用多种检测方法是未来车辆检测的一大发展趋势。另外,车辆检测中的多检测器信息融合也是未来研究的重点。摄像头获取的视频图像序列可以被看作3D真实空间向2D图像平面投影的过程,这种投影往往会造成信息的丢失,为了获取足够高的检测准确率,需要更多的空间-时间信息。而通过设置多个检测器就可以给检测系统提供更多的信息。另外,与基于视频图像车辆检测紧密联系的相关研究领域,交通视频图像的压缩和多媒体数据挖掘,也是未来研究的热点。在ITS中,车辆检测是将摄像头实时拍摄到的视频图像传输到交通管理中心进行处理分析,直接传输数据量巨大的视频图像是不经济也不现实的,所以如何有效地压缩交通视频图像有着非常重要的意义。

当前交通检测系统中检测到的车流量数据并没有与数据库相关联,现有的检测系统无法为ITS提供有组织的、非监督的、可方便存取的多媒体数据信息,而利用多媒体数据挖掘技术^[21,48]就可以从车流量数据中发现一些重要且之前未知的知识,这些知识可以帮助ITS更好地进行决策。

参考文献:

- [1] Tomizuka. Automated Highway Systems: An Intelligent Transportation System for the Next Century[J]. IEEE International Symposium on Industrial Electronics, 1997, (1): 1-4.
- [2] Micchalopoulos P G. Vehicle Detection Through Video Image Processing: AUTOSCOPE System[J]. IEEE Transactions on Vehicular Technology, 1991, 40(1): 21-29.
- [3] J Versavel, et al. Camera and Computer Aided Traffic Sensor[C]. IEE the 2nd International Conference on Road Traffic Monitoring, 1989. 21-29.

- [4] M Takatoo, *et al.* Traffic Flow Measuring System Using Image Processing[J]. SPIE, 1989, (1) : 1190-1197.
- [5] J M Blosseville, *et al.* Titan: A Traffic Measurement System Using Image Processing Techniques[C]. IEE the 2nd International Conference on Road Traffic Monitoring, 1989. 56-62.
- [6] 郁梅, 蒋刚毅. 智能交通系统中的计算机视觉技术应用[J]. 计算机工程及应用, 2001, 37(10) : 101-103.
- [7] M Papageorgiou. Video Sensors. Papageorgiou Markos (Ed) [J]. Concise Encyclopedia of Traffic and Transportation systems, 1991, 24(2) : 610-615.
- [8] Hongjiang Zhang, Yihong Gong, *et al.* Moving Object Detection, Tracking and Recognition[C]. The 3rd International Conference on Automation, Robotics and Computer Vision, 1994.
- [9] Kameda Yoshinari. Minoh Michihiko a Human Motion Estimation Method Using 3-successive Video Frames[C]. Proceedings of International Conference on Virtual Systems and Multimedia, 1996. 135-140.
- [10] 王春波, 张卫东. 智能交通系统运动车辆的视觉检测[J]. 红外与毫米波学报, 2001, 20(2) : 81-86.
- [11] J B Kim, H S Park, *et al.* A Real-time Region-based Motion Segmentation Using Adaptive Thresholding and K-means Clustering[J]. M Brooks, D Corbett, M Stumptner. Springer, 2001, 30(6) : 213-224.
- [12] N Paragios, R Deriche. 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.
- [13] P G Michalopoulos. Vehicle Detection Video Through Image Processing: The Autoscope System[J]. IEEE Trans. on Vehicular Technology, 1991, 40(1) : 21-29.
- [14] Takaba S, *et al.* A Traffic Flow Measuring System Using a Solid State Sensor[C]. Proc. IEE Conf. on Road Traffic Data Collection, London, UK, 1984. 45-56.
- [15] Hoose N. IMPACT: An Image Analysis Tool for Motorway Analysis and Surveillance[J]. Traffic Engineering Control, 1992, 23(4) : 140-147.
- [16] Kottle D, Sun Y. Motion Estimation via Cluster Matching[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1994, 16(2) : 1128-1132.
- [17] 蒋刚毅, 郁梅, 叶锡恩. 一种基于视觉的车辆跟踪及交通流量参数估计新方法[J]. 电路与系统学报, 2001, 6(4) : 69-73.
- [18] J Badenas, M Bober, F Pla. Motion and Intensity-based Segmentation and Its Application to Traffic Monitoring[C]. Proceedings of the International Conference on Image Analysis and Proceedings ICIAP '97, Florence, Italy, IAPR, 1997. 502-509.
- [19] J Badenas, J M Sanchiz, F Pla. Motion-based Segmentation and Region Tracking in Image Sequences[J]. Pattern Recognition, 2001, 34(8) : 661-670.
- [20] S Peeta, R L Kashyap. Unsupervised Video Segmentation and Object Tracking[C]. IEEE International Conference on Image Processing, 1999. 223-241.
- [21] Shu-Ching Chen, Mei-Ling Shyu. Multimedia Data Mining for Traffic Video Sequences[J]. Journal of Intelligent Information Systems, 2002, 19(1) : 61-77.
- [22] Abutaleb A S. Automatic Thresholding of Gray-level Pictures Using Two-dimensional Entropy[J]. Computer Vision Graphics and Image Processing, 1989, 47(2) : 22-32.
- [23] Yanowitz S, Bruckstein A. A New Method for Image Segmentation [J]. Computer Vision Graphics and Image Processing, 1989, 46(4) : 82-95.
- [24] Y Park. Shape-resolving Local Thresholding for Object Detection[J]. Pattern Recognition Letters, 2001, 22(5) : 883-890.
- [25] J M Blosseville, C Krafft, F Lenoir, *et al.* TITAN: New Traffic Measurements by Image Processing[C]. Proc. of IFAC Transportation Systems, 1999, 14(1) : 277-296.
- [26] N Hoose, L G Willumsen. Automatically Extracting Traffic Data from Video-tape Using the CLIP4 Parallel Image Processor[J]. Pattern Recognition Letters, 1987, 6(3) : 199-213.
- [27] T Abramczuk. A Microcomputer based TV Detector for Road Traffic[J]. In Symposium on Road Research Program, T, 1984, 3(2) : 145-147.
- [28] 郁梅, 蒋刚毅, 贺赛龙. 基于路面标记的车辆检测和计数[J]. 仪器仪表学报, 2002, 23(4) : 386-391.
- [29] W Kasprzak. An Iconic Classification Scheme for Video-based Traffic Sensor Tasks[C]. W Skarbek. CAIP 2001, Springer, Berlin, 2001. 725-732.
- [30] M Fathy, M Y Siyal. An Image Detection Technique Based on Morphological Edge Detection and Background Differencing for Real-time Traffic Analysis[J]. Pattern Recognition Letters, 1995, 16(2) : 1321-1330.
- [31] B Ross. A Practical Stereo Vision System[C]. Proceedings of International Conference on Computer Vision and Pattern Recognition, 1993, 3(5) : 148-153.
- [32] Joon Woong Lee. A Machine Vision System for Lane-Departure Detection[J]. Computer Vision and Image Understanding, 2002, 86(1) : 52-78.
- [33] M Bertozzi, A Broggi, S Castelluccio. A Real-time Oriented System for Vehicle Detection[J]. Journal of System Architecture, 1997, 43(3) : 317-325.
- [34] X Li, Z-Q Liu, K-M Leung. Detection of Vehicles from Traffic Scenes Using Fuzzy Integrals[J]. Pattern Recognition, 2002, 35(2) : 967-980.
- [35] C Setchell, E L Dagless. Vision-based Road-traffic Monitoring Sensor [J]. IEE Proceedings - Vision, Image and Signal Processing, 2001, 148(1) : 78-84.
- [36] E C Di Mauro, T F Cootes, C J Taylor, *et al.* Active Shape Model Search Using Pairwise Geometric Histograms[C]. Proc. of British Machine Vision Conference, 1996, 51(6) : 353-362.
- [37] G D Sullivan, K D Baker. Model-based Vehicle Detection and Classification Using Orthographic Approximations[J]. Image and Vision Computing, 1997, 15(8) : 649-654.
- [38] S Mantri, D Bullock. Analysis of Feedforward-backpropagation Neural Networks Used in Vehicle Detection[J]. Transportation Research Part C, 1995, 3(3) : 161-174.
- [39] M P Dubuisson, S Lakshmanan, A K Jain. Vehicle Segmentation and Classification Using Deformable Templates[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(3) : 293-308.
- [40] J Ferryman, A Worrall, *et al.* A Generic Deformable Model for Vehicle Recognition[C]. Proceedings of British Machine Vision Conference, 1995. 127-136.
- [41] D P Huttenlocher, GA Klandeman. Comparing Images Using the Hausdorff Distance[J]. IEEE Trans. on Pattern Anal. Machine Intell, 1993, 15(9) : 850-863.
- [42] Jen-Chao Tai, Shung-Tsang Tseng. Real-time Image Tracking for Automatic Traffic Monitoring and Enforcement Applications. Visual tracking[J]. Image and Vision Computing, 2004, 22(6) : 640-649.
- [43] N Ferrier, S Rowe, A Blake. Real-time Traffic Monitoring[C]. Proceedings of the 2nd IEEE Workshop on Applications of Computer Vision, 1994. 81-87.
- [44] D Beymer, P McLauchlan, *et al.* A Real-time Computer Vision System for Measuring Traffic Parameters[J]. Computer Vision Pattern Recognition, 1997, 2(9) : 495-501.
- [45] R Deriche, O Faugeras. Tracking Line Segments[J]. Image Vision Computer, 1990, 8(4) : 261-270.
- [46] Z Zhang, O D Faugeras. Three-dimensional Motion Computation and Object Segmentation in a Long Sequence of Stereo Frames[J]. Int. J. Comput. Vision, 1992, 7(3) : 211-241.
- [47] B Coifman, D Beymer, *et al.* A Real-time Computer Vision System for Vehicle Tracking and Traffic Surveillance[J]. Transportation Research Part C, 1998, 6(2) : 271-288.
- [48] Shu-ching Chen, Mei-ling Shyu, Cheng-cui Zhang. An Intelligent Framework for Spatio-temporal Vehicle Tracking [C]. Intelligent Transportation Systems Conference Proceedings, 2001, 3(4) : 25-29.

作者简介:

王圣男(1980-),男,安徽合肥人,硕士生,研究方向为图像处理与计算机视觉及其应用;郁梅(1968-),女,江苏无锡人,教授,博士,研究方向为多媒体信号处理、编码与通信、计算机视觉及其应用;蒋刚毅(1968-),男,浙江绍兴人,教授,博士生导师,博士,主要研究领域为多媒体信号处理、视频压缩与通信、基于视觉的智能控制。