VEHICLE DETECTION AND CLASSIFICATION BY CONTOUR SHAPE MODEL

Mikhail Taranin, Victor Glazov, Sergey Landyshev, Sergey Taranin
Research Center "Module"
3, Eight March 4th Street, box 166, Moscow, 125190, Russia
Phone: +7 495 152-9802 Fax: +7 495 152-4661 E-mail: taranin@module.ru

ABSTRACT

A video-based system for vehicle detection is considered. We propose a contour shape method to detect and recognize the vehicles in the complicated conditions. The method is good compromise between the detection quality and computational cost. The real-time version of contour shape algorithm has been implemented in "TrafficMonitor-C" visual surveillance system.

INTRODUCTION

The objective of this paper is to present an algorithm for vehicle detection from monocular imagery that is based on shape analysis using templates. Usually the shape models suffer from objects of unexpected shapes. To resolve this problem, deformable templates have been proposed [1]. Nevertheless deformable templates require a heavy calculation for finding deformable parameters. To remove deformations and reject the effects of rotation as well scale and perspective a projective transform method (homography) has been proposed [2]. But this method presumes the restricted requirements to position of video camera against the road. Actually this method is applicable when the camera location is high enough (12 meters or higher), then a "bird's-eye view" image is formed and vehicles can be represented by 2D models. However in many applications this requirement is not practicable. For example, typical camera height in tunnel is about 5-6 meters. In this case we have a scene where vehicles are 3D objects. It is essential to find some kind of object model or template that stable for shape deformations and allows detecting of 3D objects. In this article we propose the contour shape method for vehicle detection that is more robust against camera position and is enough efficient for real-time applications.

IDEA

3D object detection requires surface models: surface and significant curves on surfaces must be present in the representation. A vehicle can be modeled by a set of planes as primitive elements (surfaces) and their relationships. Consider projecting an object in 3D space into 2D plane. The object projection consists of elementary projections of its surfaces or planes. Each plane projection is defined by a contour. We mean the contour is a locus of the connected points on the high contrast object boundary. The apparent contours of three-dimensional surfaces are known to be rich source of information. Consequently, all detected objects (vehicles) can be represented by set of connected contours on 2D image. In order to impose constraint to the contour shape, one can use the following assumptions: all significant contours are closed and convex. These assumptions are valid in case of vehicle detection. To detect the object on the image it is need to detect all contours and then examine their combination for correspondence to a reference object model.

Two basic questions must be considered:

1. What method must be used for closed contours detection?
2. What information must be present in reference object model?

The task of finding closed contours is difficult enough. The object and part boundaries do not completely correspond to image boundaries. The detected boundaries are often fragmented and many boundaries due to surface markings and noise are present. It is a problem to robustly extract curves and apparent contours without a priori model. Vehicle surfaces can be approximately described by contours of quadrangular shape as shown on figure 1.

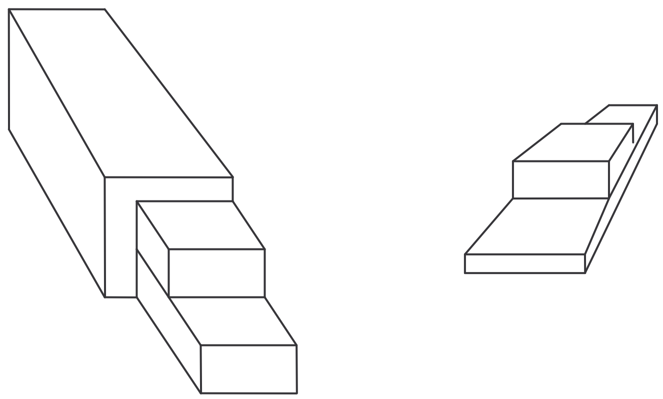

Figure 1. Vehicle model

Therefore, it is reasonable to search only quadrangular contours on 2D image. In the next section, we describe algorithm for quadrangular contour detection that is efficient for real-time applications.

Second question points to the idea of the representation structure. The idea is that of describing things at a variety of levels of models. In general, each level may require completely different conceptual tools for the analysis. We start from coarse level model. This model can be flexible and simple for vehicles. The coarse model is very suitable for such objects finding. We define the model in the following way: vehicle object is a group of adjoining quadrangular closed contours. Note this model describes vehicles of various types (car, truck). To separate vehicles from false objects and to classify vehicles the “fine level models” can be used. Fine level model is represented by projection histograms as shown on figure 2.

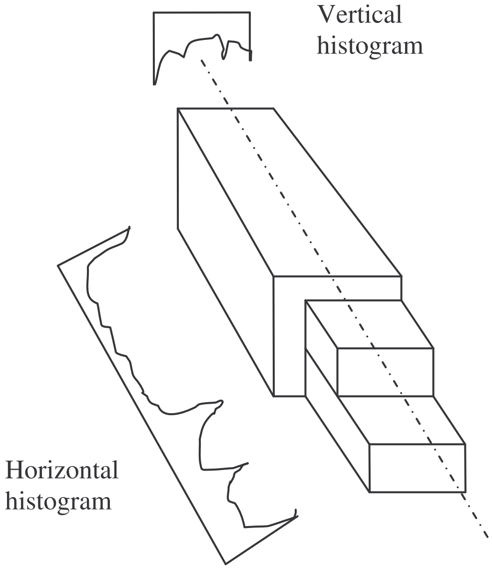

Figure 2. Projection histograms

For each detected object (after coarse model), vertical and horizontal projection histograms are computed by projecting object region on an axis perpendicular to the major axis and along the major axis, respectively. Projection histograms are normalized by rescaling projections onto fixed length. The fine model was established by studying a database of images containing typical vehicles. We use a group of fine models for every vehicle type.

ALGORITHM

The algorithm is shown on figure 3. It starts from edge detection procedure to extract edge pixels from input image.

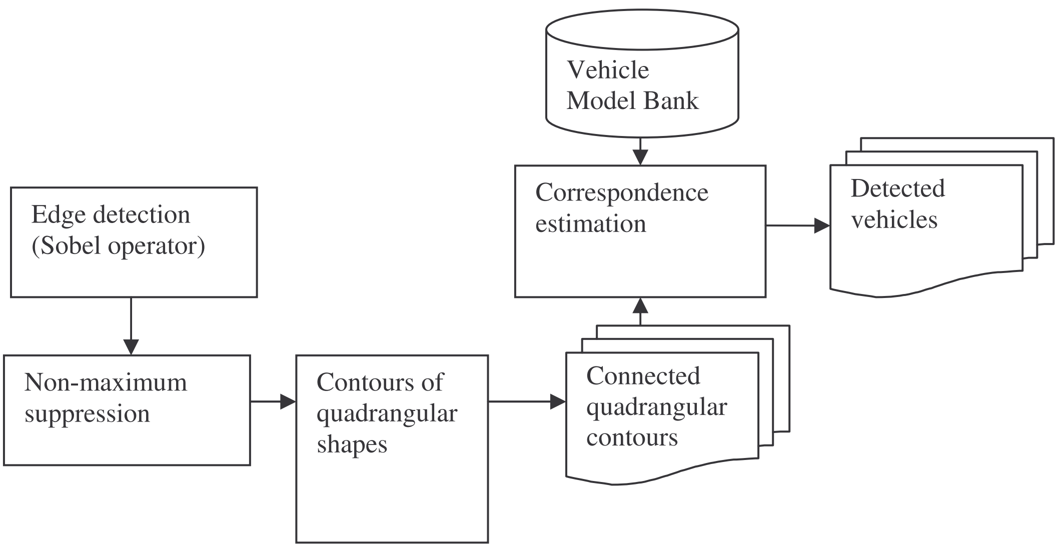

Figure 3. Algorithm scheme

We use horizontal and vertical Sobel operators to extract horizontal and vertical edges. As we are interested in contour detection, edges should be placed at the points of maximum; or rather non-maximum points must be suppressed. It is preferable to suppress non-maximum perpendicular to the edge direction, since the edge strength is expected to continue along an extended contour. We apply this procedure for vertical and horizontal edges respectively. Finally, algorithm uses the so-called "hysteresis" threshold. Points which lie between two thresholds are accepted if they are connected to pixels which exhibit strong response.

Next step is selection of contours that correspond to quadrangular shapes. Every contour is defined by two horizontal edges and two vertical edges. The algorithm looks over all horizontal edges and vertical edges to find closed convex contours. The most probabilistic combination is classified as closed contour shape. Results are shown on figure 4. All unclosed and non-convex contours have been rejected. At this stage every contour is a locus of the connected points on the object boundary. We can describe a contour using four vertices that are points there vertical and horizontal edges crosses. This representation is shown on figure 5. Finally, we seek objects of interest by grouping connected quadrangular contours.

Figure 4. Detected contours

Figure 5. Contours representation

Next step is separation of vehicles from false objects and classification. Vehicles can be in many positions while they are moving along the road. It is supposed that the angle between the view direction and the road direction is in range from 0° to 45°. We collect examples of vehicles over a wide range of views to extract vehicle shapes. Vehicle shape is represented by the normalized horizontal and vertical projection histograms. Average normalized horizontal and vertical projection templates for each vehicle type were computed experimentally. The algorithm compares the observed objects shape with the projection templates of vehicles using sum of absolute difference method to estimate the most similar vehicle type. Let S_i be the similarity between the detected shape and ith vehicle template, H_i and V_i the horizontal and vertical projections of ith template, and H and V the horizontal and vertical projections of the detected shape. S_i is calculated as:

$$S_i = -\log \sum_h \sum_v (H_i^h - H)^2 + (V_i^v - V)^2.$$

We determine the most similar template by using the highest score.

After labeling all detected object we apply a tracking algorithm. The goals of the vehicle tracking stage are to determine when a new vehicle enters or exits the system and estimate the position and velocity of each vehicle. The system employs a motion model for each vehicle to estimate its location in subsequent frames. The prediction from this model is used to estimate a bounding box location for each vehicle. These predicted bounding boxes are then compared to the actual bounding boxes of the detected vehicles. Given that a vehicle is matched to a box, the system has to determine the current position of the vehicle to update its motion model.

IMPLEMENTATION

The proposed algorithm has been implemented in "TrafficMonitor-C" - a high performance video image processing system for real time traffic data collection [3]. TrafficMonitor-C device (hereinafter referred to as TMC) is designed for video surveillance of a road section, for real time measuring the characteristics of a traffic situation, and for transmitting the measured information to a remote computer. TMC is shown on figure 6.

Figure 6. TrafficMonitor-C

The TMC is an all-weather integrated device that intended for mounting under the roads and consists of a color high-resolution video camera, high performance video processor, weatherproof housing and wire and wireless communication interfaces.

The TMC provides traffic analysis for several (up to six) lanes. The following parameters are determined: number of vehicles having passed in the time of the analysis, the average speed of all vehicles on a lane, the average distance between vehicles and occupancy of a lane. Any moving direction is admissible. Also the device provides vehicle classification among the following types: car, light truck or pickup, heavy good trucks (longer than 12 m), bus and motorcycle.

The device can detect some incidents, such as excess of the allowed speed, wrong vehicle direction, illegal stop and traffic jam. Video of an incident prehistory can be saved for further analysis.

When the arrangement of video cameras is optimal, the relative error of determining the characteristics of a traffic situation must not exceed 5% for the number of vehicles and 10% for the mean speed.

The TMC operates in real-time at 25 frames per second (PAL) or 30 frames per second (NTSC).

CONCLUSION

We have described an algorithm for detecting and tracking vehicles in outdoor environment. It operates on monocular gray-scale or color video imagery from general CCTV camera. The algorithm employs a combination of contour shape analysis and tracking to locate vehicles and classify them into five classes. We use universal vehicle templates that are combination of apparent closed contours of quadrangular shape for coarse vehicle detection. To separate vehicles from false objects and classify those fine level models is used. Fine level models are represented by projection histograms. The dynamics/kinematics model of vehicle motion is utilized to help the tracking process. The system has been successfully tested under various lighting (day, night) and weather (rain, snow) conditions.

REFERENCES

1. M.-P. Dubuisson, S. Lakshmanan, and A. K. Jain, “Vehicle Segmentation And Classification Using Deformable Templates”, IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 18, no. 3, march 1996.
2. M. Taranin, E. Zakharov, S. Landyshev, S. Taranin “NeuroMatrix TrafficMonitor System for Real–Time Vehicle Detection”, Proceedings of 11th World Congress on ITS, Nagoya, Aichi 2004, Japan.
3. “TrafficMonitor from Russia”, ITS International, July/August 2006, p. 66.