Vision-based Industrial Automatic Vehicle Classifier

[1]Timur Khanipov, [2]Ivan Koptelov, [3]Anton Grigoryev, [1]Elena Kuznetsova, and [1]Dmitry Nikolaev
[1]IITP RAS
[2]Visillect Ltd
[3]MIPT (SU)
This work is supported by Russian Foundation for Basic Research (grant 13-01-12106)

ABSTRACT

The paper describes the automatic motor vehicle video stream based classification system. The system determines vehicle type at payment collection plazas on toll roads. Classification is performed in accordance with a preconfigured set of rules which determine type by number of wheel axles, vehicle length, height over the first axle and full height. These characteristics are calculated using various computer vision algorithms: contour detectors, correlational analysis, fast Hough transform, Viola-Jones detectors, connected components analysis, elliptic shapes detectors and others. Input data contains video streams and induction loop signals. Output signals are vehicle enter and exit events, vehicle type, motion direction, speed and the above mentioned features.

Keywords: video processing, vehicle classification, industrial systems, toll roads.

1. INTRODUCTION

Worldwide popularity of toll roads has been increasing over time and lately they have become a very wide-spread type of roads. Amount of the toll varies by vehicle type which usually depends on number of axles, full vehicle height, vehicle height over the first axle, vehicle length or weight. Accurate weight measuring of a moving vehicle is a difficult engineering problem and thus is used rarely.

Vehicle type is determined at special lanes of toll plazas. Classification may be performed in one of the following modes: manual mode (by a human operator), semi-automatic mode (operator confirms or corrects automatically calculated vehicle type) and automatic mode (all operations are performed without staff). While using automatic classification greatly reduces economic costs of maintaining a toll road and helps reduce human factor in classification mistakes it is also the only possible operation mode at fast lanes of toll plazas where vehicles move at high speed.

Until recent time most automated vehicle classification systems implemented the following methods of measuring vehicle characteristics. Number of axles is usually calculated using optical sensor pairs, pressure sensors and high-frequency induction loops, while height and length are determined using laser profilers and distance meters. Applying these methods demands costly equipment and special lane construction requirements. Maintenance of such lanes is difficult and expensive. The Automatic Video Classifier (AVC) [1] was developed to deal with these issues. AVC is a complicated software and hardware system which uses digital video camera as a primary input data source for vehicle classification. AVC is currently installed and successfully performs classification at more than 100 lanes at 7 toll plazas in Russia.

2. AUTOMATIC VIDEO CLASSIFIER OVERVIEW

AVC has two separate classification subsystems which are responsible for two consecutive zones of classification which are located at the beginning and the end of the toll lane respectively: preclassifier and postclassifier (Figure 1). The preclassifier is responsible for classifying and adding a newly entered vehicle to the end of the virtual vehicle queue with an appropriate type and the postclassifier is responsible for removing vehicles which finished payment and are leaving the lane from the head of the queue (the postclassifier also performs classification to verify type correspondence). Also the postclassifier sends closing signal to an exit barrier to prevent the next vehicle proceeding freely and the preclassifier removes vehicles which are reversing from the end of the queue. Both pre- and postclassifier use two cameras which are located to the left and to the right of the lane. A special shield with contrast black and white

Seventh International Conference on Machine Vision (ICMV 2014), edited by Antanas Verikas,
Branislav Vuksanovic, Petia Radeva, Jianhong Zhou, Proc. of SPIE Vol. 9445, 944511
 · doi: 10.1117/12.2181557

checkerboard coloring is installed in front of each camera on the opposite side of the lane. Induction loops are used as additional vehicle presence detectors in each classification zone. During the day different cameras are activated to prevent their overexposure by the sun. IR-lights are used to provide additional illumination during night.

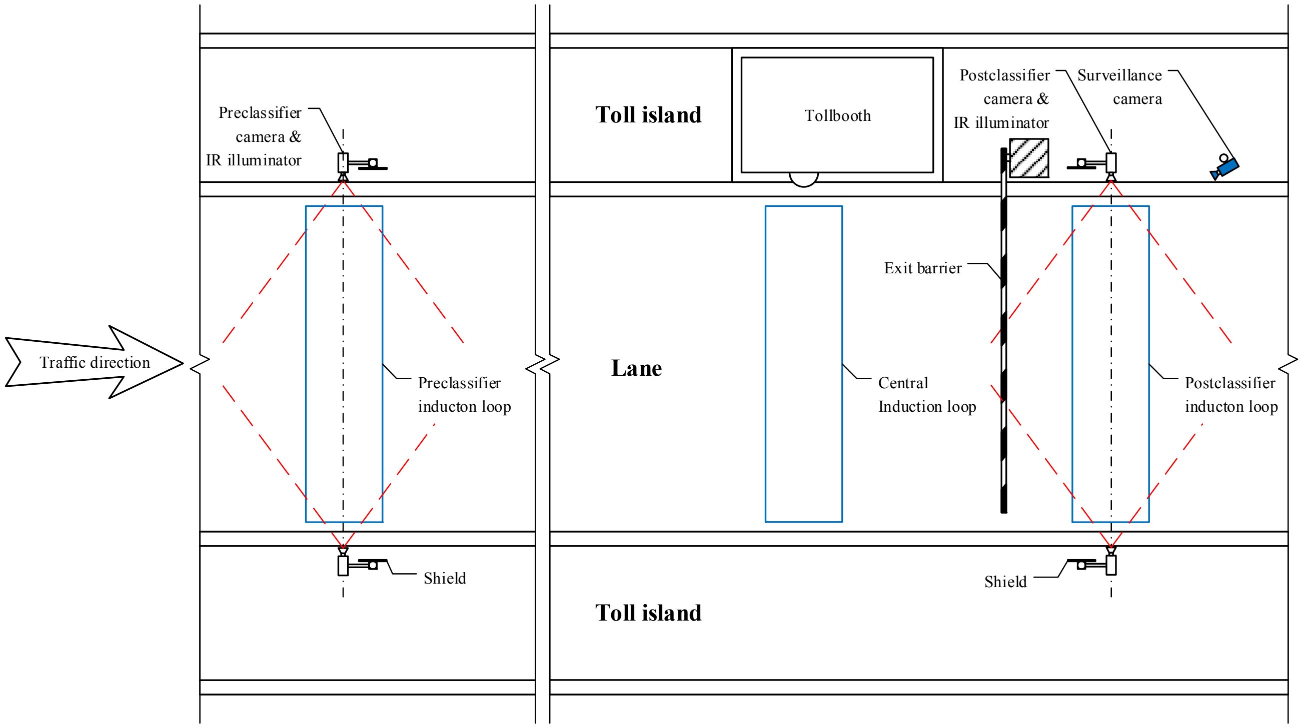

Figure 1. Lane diagram

3. SOFTWARE

AVC performs continuous real-time video stream processing to detect presence of a passing vehicle in classification zone and to calculate vehicle characteristics needed to determine vehicle type. AVC computes number of wheel axles, full height and height over the first axle.

AVC consists of a single dispatcher process and 4 recognition core processes. Dispatcher coordinates recognition cores and maintains correct vehicle queue state at the lane. The cores perform video processing (each with the corresponding camera) and classify vehicles.

3.1 Dispatcher

The dispatcher transmits vehicle type signals (as well as queue control commands) from cores to a lane controller and induction loop and exit barrier states from lane controller to cores. It also controls core states for instance for scheduling cores. The schedule is computed automatically based on geographical coordinates, geodetic azimuth of the lane and the day of year.

The dispatcher controls vehicle queue state, detects and corrects queue errors which occur after vehicle “splits”, “merges” and wrongly determined direction of vehicles due to core errors and after faulty manual editing of the queue by an operator. A queue error, if not corrected, leads to a permanent vehicle shift and a sequence of classification mismatches at pre- and postclassifier which is a serious incident. To avoid it AVC uses a special automatic queue correction mechanism. The queue corrector maintains a record of type mismatches and performs correction if two consecutive mismatches may be explained by a shift. With the help of special heuristics certain cases are also corrected. If the queue is not empty and the lane is inactive for too long (i.e. no signals come from cores or induction loops) then it gets cleared of phantoms. If the queue is empty and the payment induction loop is occupied then a vehicle has been missed by preclassifier and it is added to the queue with a default type.

3.2 Recognition cores

AVC recognition cores receive video stream and signals from vehicle presence sensors (induction loops) and control the cameras.

Each recognition core consists of separate modules which are responsible for image correction, camera control, pass detection, shield occlusion, height measurement and axles counting. The output data of these modules is processed by the classification module which transmits classification data (vehicle type and direction) to the dispatcher.

3.2.1 Image correction and camera control

The image correction and camera control module is used to maintain image characteristics appropriate for further processing. The module performs its function by means of controlling camera gain, exposure and lens iris based on analyzing specific image ROI. It also corrects radial distortion and performs different geometric transforms (rotation, scaling) for different subsystems.

3.2.2 Pass detector

Vehicle pass is detected based on a combination of the following signals: shield occlusion, significant ROI image change (using Pearson's correlation coefficient and exponential background estimation), induction state change. Shield occlusion and induction loop states are filtered first to prevent false vehicle detection due to signal instability.

Since each of the mentioned three signals does not ideally correspond to a real vehicle passage, the pass is detected only if at least two of the three signals indicate vehicle presence. There also is a set of additional decision rules which avoid most cases of splitting vehicles with long coupler trailers and merging closely moving vehicles. Those rules were created by experts and they analyze various delays between events so pass detector past state is important.

3.2.3 Height calculation

Vehicle height at a fixed lengthwise section is defined as the distance between road plane and the highest vehicle or load point. Full height (the maximum of all position heights) and height over the first wheel axle are used for classification. First of all to determine vehicle height its vertical profile must be measured. The first way to measure it is to analyze shield occlusion pattern and detect the highest occlusion point (after appropriate filtering because cells may be falsely occluded). The second algorithm compares the current frame with the accumulated since the beginning of the pass background in the specific ROI, which is a vertical rectangle located in the middle of the image. The first method gives the most accurate result however it fails to detect heights which are higher than the top of the shield.

Vehicle type depends on which metric interval specific height (full or first axle) belongs to (for instance, vehicles of type 1 have height under 2 m while vehicles of type 2 have height over 2 m). Cameras are installed at a height which equals to one of the height border values (i.e. border between height intervals). For this particular case height border can be marked in image coordinates at the shield (which has appropriate vertical alignment) and it is only necessary to check if profile (or a single profile segment if first axle height is needed) intersects the corresponding horizontal border line without calculating height in meters. In this situation geometric scene model inaccuracies do not influence classification result which increases classification quality and simplifies the system configuring process.

In general case it is necessary to calculate height in meters and scene geometry must be used to transform vehicle top and bottom borders in the image to real world coordinates (and measure height). The bottom border is deduced from wheel tracks by the wheel detection module.

3.2.4 Wheel axles calculcation

Wheel axles are calculated in two steps: wheel detection and summarizing detection results.

Several methods are simultaneously used for wheel detection. The first one is machine learning based Viola-Jones (VJ) detector [2], [3] which robustly detects common wheel types but sometimes fails to detect rare wheel types that were not present in the learning set. The second one is elliptic objects detector [4] which can detect wheels of any type but with lower than the VJ detector quality. All detectors process each frame independently and have false positives, false negatives and multiple detections of the same wheel, hence for calculating the actual number of axles special postprocessing is needed: filtering and integration of detections.

Filtering is performed based on geometric limitations: all wheels must be positioned on the same horizontal line and have the same size for each axle.

Counting wheel axles is performed by detecting and counting wheel tracks on a so-called xt-diagram. The xt-diagram is a raster image in which the vertical coordinate corresponds to frame number while the horizontal coordinate matches the horizontal coordinate of the image (Figure 3 and Figure 4). Each wheel produces a continuous line on this diagram. One of the two following algorithms is applied to find and calculate tracks depending on time length and motion type. In case of uniform motion the tracks are straight lines and the fast Hough transform is used to detect them (Figure 2) [5]. In case of non-uniform motion the vehicle speed is relatively small, wheel detections have high density and tracks are detected using connected components analysis on the diagram (Figure 3). Also the connected components search method is used if fast Hough transform indicates presence of a priori incorrect tracks (for instance, if they intersect, Figure 4).

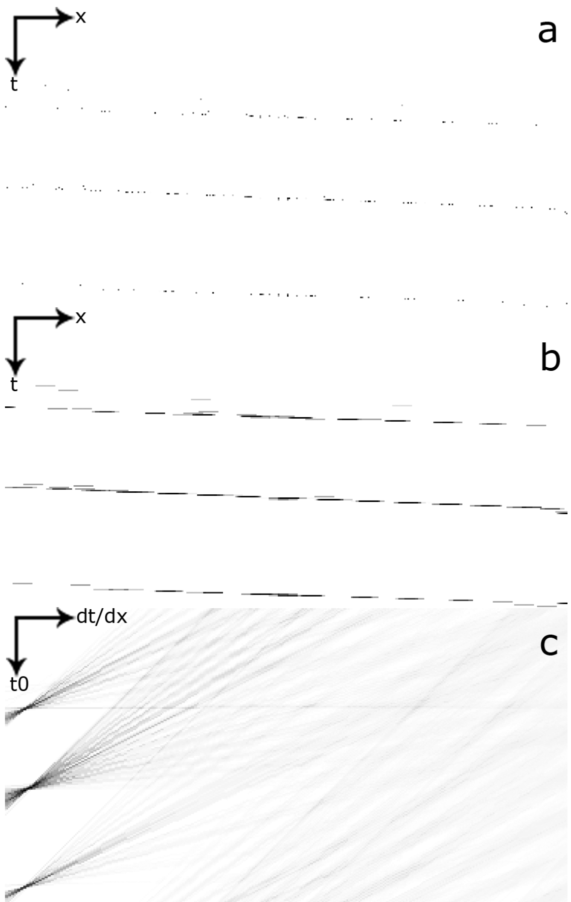

Figure 2. Fast Hough transform

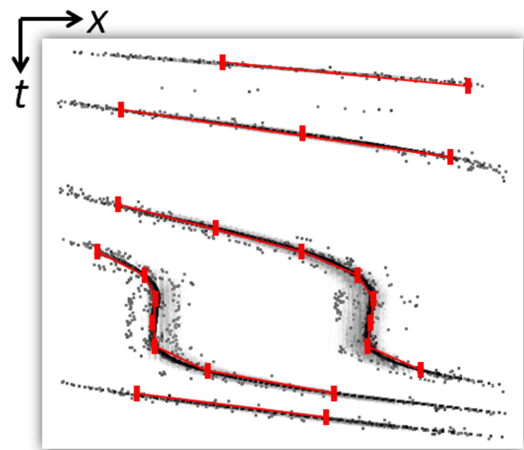

Figure 3. Connected components search

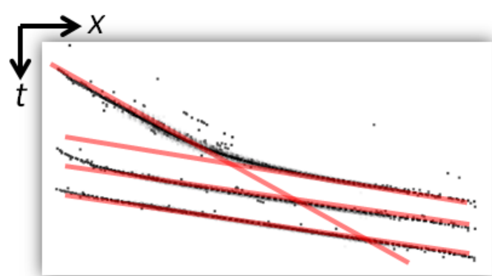

Figure 4. Example of incorrect track detection using FHT

3.2.5 Direction detector

Vehicles may pass classification zone in different directions. In case of forward or backward directions appropriate queue modifications must be performed. Vehicles moving forward may stop at some moments and even have segments of reverse motion. In case of “mixed” direction (for instance the vehicle entered the classification zone in forward direction, stopped and left it by reversing) queue should not be modified. The correct detection of direction is a very important task because in case of an error real and virtual vehicle queue mismatch will occur.

The primary method for detecting direction is similar to the Lucas-Kanade algorithm which analyzes optical flow [6]. The secondary method uses wheel counting module data and deduces direction from wheel tracks [7]. The second method is less accurate and is only used if the first method fails (i.e. rejects input data).

4. HARDWARE

AVC consists of the following hardware components: computer, camera with lens and housing, shield and IR-lights.

The camera has a grayscale 480x752 CCD sensor sensitive to visible and near-infrared range which operates at 40 fps. Horizontal view angle is 120 degrees, vertical view angle is 70 degrees. The camera is installed inside a heated housing with IP66 protection mark. The shield is a 300x2000 mm panel with black and white checkerboard coloring with cell size 150x100 mm. The LED IR-lights with software control are used to maintain sufficient luminance of the scene during night. An industrial production quality computer is used, it has the Intel Core i7-3610QE processor with peak productivity of 73.6 GFlops.

5. CONCLUSION

An automatic videostream based vehicle classifier was developed and is currently installed at toll roads in Russia. It can operate with only video signal present as a single input but also uses induction loop if available. Computer vision algorithms and special heuristics are used to detect vehicle passages and characteristics which determine vehicle type: number of wheel axles, full height, height over the first axle and length. Additionally motion direction and speed are determined. Classification errors rate is less than 1% with 24-hour all season realtime operation mode.

REFERENCES

[1] Khanipov T.M., Postnikov V.V., Grigoryev A.S., Usilin S.A., Nikolaev D.P., “Sposob avtomaticheskoi klassifikatsii transportnykh sredstv” // Patent RF № 2486597

[2] Kotov A.A., Usilin S.A., Nikolaev D.P., “Postroeniye ustoichivykh priznakov dlya algoritma Violy i Dzhonsa v zadache klassifikatsii transportnykh sredstv”, 35-ya konferentsiya molodykh uchenykh i spetsialistov “Informatsionniye tehnologii i sistemy – 2012”, 19-25 avgusta, Petrozavodsk, 2012;

[3] Grigoryev A.S., Usilin S.A., Nikolaev D.P., “Uskoreniye poiska ob’ektov v videopotoke metodom Violy-Dzhonsa putyem adaptivnogo vybora raspoznayushikh kaskadov”, Trudy 55-i nauchnoi konferentsii MFTI “Sovemenniye problemy fundamentalnykh i prikladnykh nauk”: chast IX. Innovatsii I visokiye tehnologii. – M.: MFTI, 2012

[4] Kotov A.A., Nikolaev D.P., “Proslezhivaniye v videopotoke ob’ektov, soderzhashikh mnozhestvo kontsentricheskikh dug”, 35-ya konferentsiya molodykh uchenykh i spetsialistov “Informatsionniye tehnologii i sistemy – 2012”, 19-25 avgusta, Petrozavodsk, 2012

[5] A.Grigoryev, T.Khanipov, D.Nikolaev., “Determination of axle count for vehicle recognition and classification”, 8th Open German-Russian Workshop «Pattern Recognition and Image Understanding»: Workshop proceedings. — Nizhny Novgorod, 2011. — p. 89–91.

[6] B. D. Lucas and T. Kanade (1981), “An iterative image registration technique with an application to stereo vision”, Proceedings of Imaging Understanding Workshop, pages 121 – 130

[7] D. Nikolaev, S. Karpenko, I. Nikolaev, P. Nikolayev., “Hough Transform: Underestimated Tool in the Computer Vision Field”, Proceedings of 22th European Conference on Modelling and Simulation, 2008, pages 238-243.