Intelligent traffic management through MPEG-7 vehicle flow surveillance

C. Anagnostopoulos
Department of Cultural Technology and Communication
University of the Aegean

T. Alexandropoulos, V. Loumos, E. Kayafas
School of Electrical and Computer Engineering
National Technical University of Athens

***Abstract*—Recent development in the aspects of low-level image processing for feature extraction, as well as in the standardization of description schemes for video description, provide means for the development of Intelligent Transportation Systems. ITS may support a wide range of abilities, including vehicle identification and reidentification, event detection and optimum resource management. This paper analyzes the structure operation principles and the structure of an integrated highway management system. Its operation is based on the combination of sensor networks, low-level image processing algorithms and high-level description schemes. Low-level algorithms perform the tasks of license plate recognition for vehicle identification, as well as feature extraction through change detection, for event detection purposes. High level schemes are based on the recently developed MPEG-7 Description Schemes and are utilized as an information processing framework, which formulates identification and event representation procedures.**

***Index Terms*—highway surveillance, traffic management, vehicle re-identification, license plate recognition, change detection, event detection.**

I. INTRODUCTION

Intelligent Transportation Systems (ITS) have emerged as a means to improve transportation safety and mobility and to enhance productivity through the use of advanced technologies. ITS are made of 16 types of technology-based systems. These systems are divided into intelligent infrastructure systems and intelligent vehicle systems. In this paper, a computer vision sensor network, which is based on specialized algorithms, is suggested as a core for an intelligent integrated system for vehicle surveillance and traffic management.

In recent years, cameras have been increasingly employed as image sensors for traffic control. In addition, video and image processing algorithms can effectively exploit video images to perform easy tasks, such as vehicle count and presence detection. Lately, video images can be used to extract further traffic information that is unavailable from loops or other conventional traffic detectors.
Specifically, advanced image processing routines may accomplish the following tasks:

- Dynamically update threshold for background extraction and perform change detection
- Estimate mean speed of vehicles per minute, density/capacity per direction and lane
- Estimate average travel time and queue length
- Recover vehicle parameters such as location, length, and speed
- Identify unsafe driving behaviour in individual vehicles
- Detect hazardous events, including accidents, stopped vehicles and obstacles
- Assess the severity/nature of an incident
- Identify the presence of pedestrians or animals in highways

The tracking of high-level features needed for event detection may be performed by integrating low-level techniques and high level description languages, such as the recently developed MPEG-7 standard, into a single framework.
This paper is structured as follows. Section II analyses the application of license plate recognition as a means for the implementation of analog toll billing and vehicle re-identification. Section III presents early vision change detection algorithms which are used for feature extraction in video streams prior to the estimation of high-level definitions. Section IV describes the utilization of high-level descriptors of the MPEG-7 standard, in order to address event interpretation and detection in highway surveillance. The architecture of an integrated highway management model is analyzed in Section V.

II. ANALOG TOLL BILLING VIA LICENSE PLATE RECOGNITION

ITS may be incorporated in the entrance and exit motorway gates in order to bill the appropriate amount, according to the route traversed by a vehicle. This problem is well known in the literature as vehicle re-identification (VrI). VrI is the process of matching vehicles from one point on the roadway (e.g., motorway entrance) to another (e.g., motorway exit). By performing vehicle re-identification, important traffic parameters, including travel time, travel time variability and analog toll billing in a motorway, can be obtained, as shown in Figure 1.

IEEE John Vincent Atanasoff 2006 International Symposium on Modern Computing (JVA'06)
0-7695-2643-8/06 $20.00 © 2006 IEEE

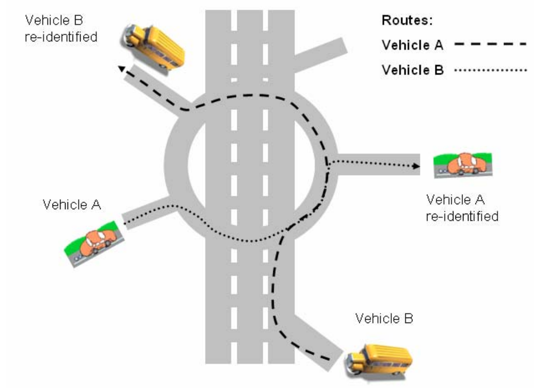

Fig. 1. Analog toll billing through license plate recognition

The practical traffic applications of VrI are many. The derivation of section travel times (time taken by a vehicle to go from one point to another) is useful to transportation engineers for traffic operations, planning and control. Accurate travel times and densities can be instrumental in travel reliability, feedback control, vehicle routing, traffic assignment and traveller information systems. If vehicles are tracked along consecutive points, then partial origin/destination demands can even be measured instead of estimated [1]. In addition, the problem of analog toll billing is safely addressed with a robust VrI system.

License plate remains the principal vehicle identifier and therefore, the vehicle re-identification process relies heavily on robust License Plate Recognition systems (LPR). One of the goals of this paper is the presentation of a novel segmentation technique implemented in a LPR system able to cope with outdoor conditions if parameterized properly.

In this paper, a new algorithm for vehicle license plate identification is proposed as a major component for the vehicle re-identification task. The LPR algorithm is implemented on the basis of a novel adaptive image segmentation technique (Sliding Concentric Windows-SCW) and connected component analysis in conjunction with a character recognition Neural Network.[2] The algorithm was tested with 1334 natural scene gray level vehicle images of different backgrounds and ambient illumination. The camera focused on the plate, while the angle of view and the distance from the vehicle varied according to the experimental setup. The license plates properly segmented were 1287 over 1334 input images (96.5%). The Optical Character Recognition (OCR) system is a two layer Probabilistic Neural Network (PNN) with topology 108-180-36, whose performance for entire plate recognition reached 89.1%. The PNN is trained to identify alphanumeric characters from car license plates based on data obtained from algorithmic image processing. Combining the above two rates, the overall rate of success for the proposed LPR algorithm is 86.0%. However, the success ratio is increased when the LPR software is accompanied with auxiliary infrared illumination units.

This latter method is motivated from the nature of the license plate surface and is already tested in many similar applications. License plates in Greece, as in many other countries in Europe and America, are made from a retro-reflective material, which causes them to shine when the sun or even a bright light is directed towards them. This attribute makes license plates perfect candidates for cameras that are sensitive to infrared illumination. Such cameras, employing a narrow band-pass filter and coupled with infrared illuminators make ideal license plate capture devices. Illumination in the area of 950nm is ideal for capturing retro-reflective license plates while avoiding distraction to drivers, since it is unobtrusive, invisible and eye-safe.[3] The illumination unit is an array of Light Emitting Diodes (LEDs), which were incorporated into a secure housing in a manner that allows them to operate in any outdoor condition. The camera incorporates a filter eliminating any light that falls outside the range of 920nm to 980nm.

The license plate recognition sequence, which is proposed in [2], consists of two distinct parts. The first one deals with the detection of the Region of Interest (RoI), i.e. the license plate. The second part includes operations for the successful segmentation of the license plate characters along with an Artificial Neural Network which performs the Optical Character Recognition task.

a) License plate segmentation: This part consists of pre-processing, in order to detect the Region of Interest (RoI) in ambient illumination conditions. The pre-processing component consists of four tasks: implementation of the SCW segmentation method [2], image masking, binarization with Sauvola method, and finally connected component labeling and binary measurements, which are arranged in sequence. The license plate segmentation flowchart is highlighted in Fig. 2.

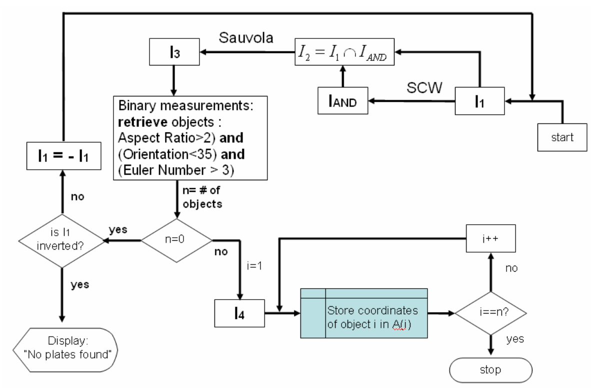

Fig. 2. Flowchart indicating the license plate location algorithm.

b) License plate processing: This part consists of pre-processing, in order to isolate the characters in the license plate image and then forward them in the Neural Network for identification. The license plate processing flowchart is shown in Fig. 4.

To summarize, the results and the characteristics of our implementation place it as a valid competitor among other non-commercial systems. Specifically, the proposed algorithmic structure was evaluated on a fixed location as well as in images acquired in various views and illumination conditions. In addition, it was tested on vehicle

0-7695-2643-8/06 $20.00 © 2006 IEEE

images from various European countries. Further details may be found in [2].

Fig. 3. Steps for license plate segmentation.

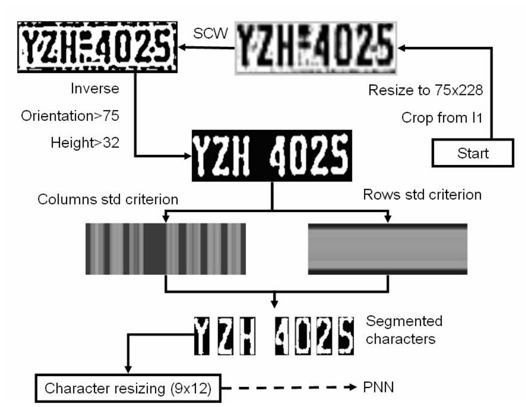

Fig. 4. Flowchart indicating the license plate processing and character recognition algorithm.

Fig 5. Successful segmentation of characters following the SCW method. Original segmented plate (left), SCW result (middle) and character segmentation after the binary measurements (height, orientation).

III. Feature extraction in highway surveillance

Feature extraction in highway surveillance focuses on the application of image processing algorithms at the early stage of image acquisition. It is useful as a refinement operation for the isolation of regions of interest, which will be fed to subsequent algorithms of the surveillance process chain for analysis. The term “regions of interest” refers to entities, which are labeled as candidate regions for the emergence of events.

It is clear that the presence of virtually every object on a highway scene may lead to the emergence of events. For example, the presence of a pedestrian or a stopped vehicle may be perceived as potentially hazardous behavior. Therefore, the process of feature extraction can be treated as occlusion detection analysis, where the reference background scene is the surveyed highway. Thus, feature extraction is satisfied through the application of change detection methods, i.e. the comparison of each image frame against a predetermined reference background.

Change detection, like all early vision methods, needs to cope with the challenges, which are encountered in “real scenes”. Some challenges include sensor noise and illumination fluctuations. A reliable change detection should extract content changes accurately, regardless of the presence of these factors. This procedure relies on the formulation of a decision rule, which achieves the optimum separation of content changes from noise level fluctuations. Change detection is frequently based on statistic approaches,[4] such as the significance test proposed by Aach et. Al. [5][6] However this test assumes a certain noise probability distribution and relies on the validity of the null hypothesis (no content change) in the examined region.[7]

The adopted noise model estimation strategy depends on the availability of a background video sequence. When a background video sequence is available, the noise model is estimated through statistic measurements on the image difference of subsequent background frames. When there is only a background frame available, noise model estimation is performed through block clustering, as described in [8]. In this case, image blocks of the absolute difference are grouped in clusters, depending on their statistic similarity. Based on the assumption that the dominant cluster carries noise model information, a statistic decision rule is formulated and applied at block level.

The effective operation of the change detection also depends on the infrastructure utilized for the video surveillance task. When grayscale cameras are used, the process is restricted to the comparison of intensity fluctuations in a single colour channel. Therefore, the system lacks the necessary data to discriminate content changes from local illumination changes and shadows. Variations in ambient illumination can be dealt with through histogram measurements as described in [8]. When colour images are employed, change detection algorithms facilitate this separation through the application of block clustering on the chromaticity components defined in colour spaces, such as *YCbCr* or *Lab* . In this scenario, it is assumed that the colour is a distinguishing characteristic of occlusions, thus the suppression of localized illumination changes is achieved.

However, despite the availability of colour cameras, image acquisition conditions in highway scenes may require the restriction of change detection on the luminance component as the best solution. This is usually the case when monochromatic light sources are used for the scene illumination. A sample frame is presented in Figure 2a. In this case, regardless of the image acquisition device type, the system is unable to extract colour information from the captured frames. The designer may either suggest the use of proper lighting or restrict the algorithm operation to intensity variation tracking. Another case that requires restriction of change detection to the intensity component is when old analog CCTV infrastructure is used. CCTV analog equipment is frequently characterized by increased chroma noise, which disorients chromaticity-based change detection criteria.

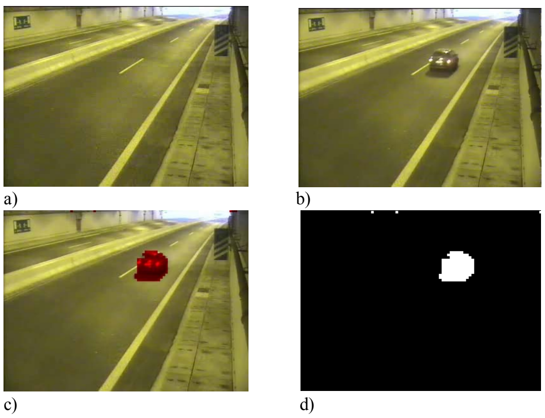

Fig 6 a) Reference highway scene frame, b) Occluded highway frame, c) Content changes detected through block-based change detection on the luminance component in the Lab color space d) Binary change mask

Regardless of the selected change detection algorithm variant, a factor that causes the emergence of false positives is the existence of time-variant light sources with different light temperature characteristics. In this case, localized illumination changes result in chrominance shifting and cause false alarms. Change detection algorithms, which focus on chromatic alterations, ignore localized illumination changes, provided that their colour temperature remains constant. However, changes in illumination colour temperature are common in highway environments, especially when switching from daylight to night shots. In order to compensate with temporal variations of mixed light sources there is the need to update the reference background through background estimation.

An important distinction in the background updating procedure used in highway surveillance lies in the demand that, stationery occlusions should not be incorporated to the background. This behavior is common in background update methods which perform background modeling through pixel-wise temporal tracking of intensity values. Temporal median,[9] linear prediction methods,[10] as well as the Mixture of Gaussians [4] are grouped in this method category. The demand that occlusions are not added to the background exists, because static changed regions may correspond to potentially hazardous events, such as immobilized vehicles or accident scenes. In order to avoid stationery target addition to the background, the respective algorithm may focus on edge detection rather than pixel-wise intensity tracking as suggested in [11].

Figures 6a, 6b represent a sample image pair subjected to comparison in order to isolate the regions of interest. The changes, which are detected by the block-clustering algorithm, are shown in Fig. 6c, while Fig 6d displays the binary change mask, which is fed to subsequent algorithms for scene interpretation purposes.

IV. Event detection

Upon completion of low-level processing algorithms, further processing of the extracted regions of interest focuses on the detection of events of significance in highway surveillance. Event detection lies in the interpretation of the temporal behavior of the extracted regions.

Temporal behavior of image regions can be described through the observation and tracking of high-level attributes. Most attributes of interest in highway surveillance are motion-related and have been incorporated as standardized descriptor schemas in the MPEG-7 standard. A detailed analysis on MPEG-7 Visual Motion Descriptors is available in [12]-[14]. Motion descriptors encapsulate information concerning the speed and trajectories of individual regions. Additionally, the overall activity in the surveyed speed, in terms of either overall observed speed or the number of objects detected, can serve as a means to identify events essential for traffic management applications. Additional visual attributes, such as the spatial distribution of motion activity, also provide usable information.

In the following paragraphs, a list of events, which can be expressed by certain combinations of visual descriptors, is presented:

a) Vehicle count: Vehicle counting is easily performed, by detecting the number of occlusions, which are present in the highway scene.

b) Speed estimation: Vehicle speed is estimated through measurements of motion vector magnitudes.

c) Stopped vehicles - Accidents: The emergence of a stationery region can be interpreted as an event that compromises the drivers' safety, such an immobilized vehicles, debris or an accident scene.

d) Highway crossing: The detection of motion vectors, which are perpendicular to the lane direction, can trigger the emergence of hazardous events, such as the presence of a pedestrian or animal crossing.

e) Traffic congestion: Highway scenes with high traffic congestion present an increased amount of occlusions (vehicles) and are characterized by slow motion. Therefore motion vectors will be characterized by an extended and uniform spatial distribution. Therefore the concurrent presence of an increased amount of blobs in the binary change mask as well as the extended presence or the absence of motion can be selected to represent the respective scenario.

f) Unsafe driver behavior: Unsafe driving can be detected through comparison of vehicle trajectories against a predetermined set of trajectory patterns, which represent safe driving profiles. Based on these profiles, the operator can establish tolerance parameters, such as the maximum acceptable curvature deviation in a trajectory. Deviations, which exceed the predetermined limit, will issue a warning to the operator.

It is clear that the template trajectory descriptors vary, depending on the highway curvature of each scene and the positioning of the image acquisition device with respect to the scene. Therefore, prior to the initialization of a video surveillance unit, a training procedure, which adopts the unit to the specific monitoring conditions of the surveyed environment, is mandatory.

g) Lane specific measurements: By separating the surveyed highway scene into distinct regions of interest in advance, the operator may apply the aforementioned event detection tasks separately for each highway lane, thus maintaining a separate record of statistics for each one. [15]

V. Proposed Architecture

This Section analyses the architecture and operation principles of a traffic management system. The respective diagram is shown in Figure 7.

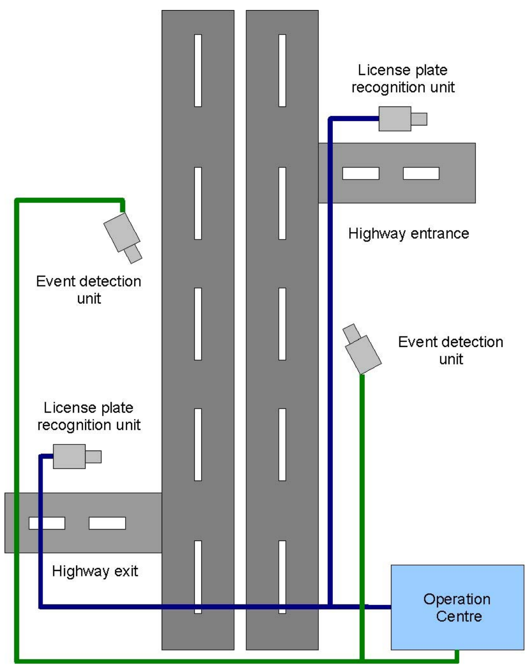

Fig. 7 Architecture of an integrated highway surveillance system

The proposed system incorporates IP/MPEG7 cameras, which combine the advantages of sensor networks and MPEG-7 features, in order to achieve video sensing, video processing, metadata production and communication within a single device. These cameras capture a video stream, assess traffic information, identify events and transfer compressed video stream along with useful metadata to the network. It is obvious that, in order to avoid transmission or storage of irrelevant information present in the traffic video sequence, the employment of feature extraction in the early stage of image acquisition is proposed.

The proposed system utilizes cameras for two purposes: license plate recognition and event detection. Cameras, which deal with license plate recognition, are placed in highway entrance and exit points, while cameras, which perform event detection, are spread across the highway network. All cameras are connected to an Operation Centre through a sensor network.

License plate recognition cameras provide information regarding the entrance and exit point of each vehicle, thus providing the necessary data for the application of route-based toll policy, as well as the extraction of highway usage profiles.

Event detection cameras are based on the MPEG7 annotation scheme. Their image sequences are used to track vehicles on a video sequence and produce metadata related to the traffic flow. Upon completion of the process, the results are stored locally as an XML file.

The Operation Centre periodically issues queries, which relate to the current traffic load, to each event detector unit and receives the respective response. More complicated queries can be issued by the utilization of MPEG-7 Descriptors. For example, a query for the detection of interchanges where traffic congestion exceeds a threshold can be issued by searching for MPEG-7 descriptor profiles that corresponds to the specific criterion.

When a significant event, such as an accident, is captured, the Operation Centre is automatically notified by the respective surveillance station and a live video feed is transmitted. Depending on the availability of feedback mechanisms, the functionality of the system may extend to the suggestion of an appropriate course of action. In an accident scenario, the suggested action will include the blocking of the respective highway lane and the intervention of emergency medical services. In a traffic congestion scenario, appropriate actions may include the provision of extra lanes, the automatic adjustment of top speed limits or the driver notification via display boards.

Additionally, the event detection module can extend its awareness to the level where it verifies that a suggested course of action is actually followed. For example, it may assure that the suggested traffic policies and safety practices are actually followed by the drivers.

Observed patterns and trends can also serve as a means to assure the optimum resource allocation and the optimum coordination of the services provided. Statistic analysis can be utilized, in order to improve quality of service through preventive measures. For example, the detection of weather-dependent or time-dependent patterns in traffic congestion statistics could be used as a feedback mechanism and therefore enforce congestion avoidance scenarios *before the event appears*.

VI. Conclusions

The structural components of an integrated intelligent traffic management system have been analysed. The operation is divided into two phases: early vision procedures and high-level metadata processing. Early vision algorithms bridge the gap between the intelligent transportation system and the highway environment and extract features which are fed to content management procedures. The latter provide an information processing framework which facilitates the formulation of data-related procedures such as the definition of queries regarding vehicle tracking, traffic status estimation and event representation. This representation schemes facilitate the selection of the appropriate course of action for each traffic scenario.

The objective and quantitative analysis of the above traffic scenarios can be extensively exploited for scientific purposes, improving the safety measures and regulations. Providing large amount of data to the scientific community on events detected by image processing sensors, it would be possible for traffic engineers to develop further preventive safety measures. Furthermore, it is possible to observe events that do not result in accidents and identify the conditions behind the occurrences of close shaves.

REFERENCES

[1] C.C. Sun, G.S. Arr,, R.P Ramachandran, S.G. Ritchie, "Vehicle Reidentification using multidetector fusion", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 5, Issue 3, Sep. 2004, pp 155–164

[2] C.N. Anagnostopoulos, I. Anagnostopoulos, V. Loumos and E. Kayafas, "A license plate recognition algorithm for Intelligent Transportation Systems applications", *accepted for publication in IEEE Transactions on Intelligent Transportation Systems*, 2006

[3] J. Kennedy, "License Plate Capture Systems And The Application Of Mid-Infrared Illumination" [Online]. Available: http://www.ettm.com/news/lpr_infrared.html

[4] R. J. Radke, S. Andra, O. Al-Kofahi, B. Roysam, "Image change detection methods: A systematic survey", *IEEE Transactions on Image Processing*, Vol. 14, Issue 3, 294-307, 2005

[5] T. Aach, A. Kaup, R. Mester, "Statistical model based change detection in moving video", *Signal Processing*, Vol. 31, Issue 2, pp 165-180, 1993

[6] T. Aach, A. Kaup, R. Mester, "Change detection in image sequences using Gibbs random fields: A Bayesian approach", *International Workshop on Intelligent Signal Processing and Communication Systems*, Sendai, Japan, pp 56-61, 1993

[7] A. Cavallaro, T. Ebrahimi, "Video object extraction based on adaptive background and statistical change detection", *Proc. SPIE Visual Communications and Image Processing*, pp 465-475, 2001

[8] T. Alexandropoulos, S. Boutas, V. Loumos and E. Kayafas, "Real-time change detection for surveillance in public transpartation", *IEEE International Conference on Advanced Video and Signal-Based Surveillance*, Como, Italy, September 15-16, pp 58-63, 2005

[9] M. Massey, W. Bender, "Salient Stills: Process and Practice", *IBM Systems Journal*, Vol. 35, No. 3 & 4, pp557-573, 1996

[10] C. Ridder, O. Munkelt, H. Kirchner, "Adaptive background estimation and foreground detection using Kalman filtering", *International Conference on Recent Advances in Mechatronics*, pp 193-199, 1995

[11] K. Kiratiratanapruk, P. Dubey, S. Siddhichai, "A gradient-based foreground detection technique for object tracking in a traffic monitoring system", *IEEE International Conference on Advanced Video and Signal-Based Surveillance*, Como, Italy, September 15-16, pp 377-381, 2005

[12] Shih-Fu Chang, Thomas Sikora, Atul Puri, "Overview of the MPEG-7 Standard", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No.6, pp 688-695, June 2001

[13] Thomas Sikora, "The MPEG-7 Visual Standard for Content Description", *IEEE Transactions on Circuits and Systems for Video Technology,* Vol. 11, No. 6, pp 696-702, June 2001

[14] Sylvie Jeannin and Ajay Divakaran, "MPEG-7 Visual Motion Descriptors", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 11, No. 6, pp 720-724, June 2001

[15] J. Melo, A. Naftel, A. Bernardino, J. Santos-Victor, "Detection and classification of highway lanes using vehicle motion trajectories", IEEE Transaction on Intelligent Transportation Systems, Volume 7, Issue 2, pp 188 – 200, June 2006

0-7695-2643-8/06 $20.00 © 2006 IEEE