Research, Development and Application on
Information & Communication Technology

CÁC CÔNG TRÌNH NGHIÊN CỨU, PHÁT TRIỂN VÀ ỨNG DỤNG CÔNG NGHỆ THÔNG TIN VÀ TRUYỀN THÔNG

Journal on INFORMATION TECHNOLOGIES & COMMUNICATIONS

NEW SERIES

Contents

A Literature Review on License Plate Recognition and Vehicle Classification

Trung-Kien Dao, Thi-Lan Le, Thanh-Hai Tran, Ngoc-Thuy Le

Int. Research Center in Multimedia Information, Communication and Applications

Hanoi University of Science and Technology

1 rd. Dai Co Viet, Hanoi, Vietnam

{trung-kien.dao, thi-lan.le, thanh-hai.tran, ngoc-thuy.le}@mica.edu.vn

Abstract – Automatic license plate recognition and vehicle classification based on computer vision technique have been deeply studied and developed. Certain commercial systems have been deployed. This paper aims at making a review of the approaches for license plate recognition and vehicle classification in the literature. Advantages and weaknesses of each approach are also pointed out. This paper can be used as a reference while developing a license plate recognition and vehicle classification system.

Keywords – *license plate recognition, vehicle classification, computer vision*

I. INTRODUCTION

Digital cameras have become ubiquitous in the last two decades, with increasing image quality and capturing speed, motivating attraction to automated visual traffic surveillance from the community of computer vision. The most notable applications of a surveillance system, among others, that can be mentioned are automatic toll payment, vehicle trajectory tracking, traffic statistics and monitoring, stolen vehicle detection, driver navigation support, and parking access control. An alternative method of license plate recognition (LPR) and classification is based on OBU (on-board unit). This device can communicate with a roadside unit to transfer declared information demanded by the station for the identification including license plate number as well as vehicle class. However, this method is not reliable, as the stored information can be alternated, and the OBU can be replaced by a fraudulent one.

This study aims to make a survey on the literature of license plate recognition and vehicle classification in the case of a toll payment system. From a general view, such a system can be decoupled into two main modules, as depicted in Figure 1. The first module is to identify on-road vehicles by recognizing their license numbers from a sequence of camera images, and to track them as they pass through the field of view of the camera. The tracking is important because it ensures that one vehicle can not be confused with others, so that each one is counted only once. The second module is to determine the vehicles' geometry parameters (i.e., width, length, height...) and to classify vehicles into predefined categories. The classification in this phase may or may not use the estimated geometry parameters. The final goal of this system is to determine the traffic toll tariff for each vehicle passing the station on the basis of gathered information consisting of country code of origin, vehicle size and weight. Beside the two main modules, the system might also be equipped by some other sensors installed at toll station, on or beside road such as camera, laser, and strain gauge to collect necessary information that can be taken into account in the classification for accuracy improvement.

II. LICENSE PLATE RECOGNITION

The license plate recognition in most studies is in general divided into three phases: (i) plate localization, (ii) character segmentation, and (iii) optical character recognition (OCR). In the first phase, the system needs to detect the areas within a single image which may contain a number plate. In the second one, each candidate area is first segmented into smaller rectangular ones corresponding to individual characters, which are finally identified by an OCR method. Nevertheless, there exist also a few studies in which the texts are identified directly without detecting the license plate areas. That means the first

phase is skipped in those studies. Besides, for many methods, obtained images for plate candidate areas might need to be preprocessed, including skewing, rotating, resizing and brightness normalizing, to improve recognition performance.
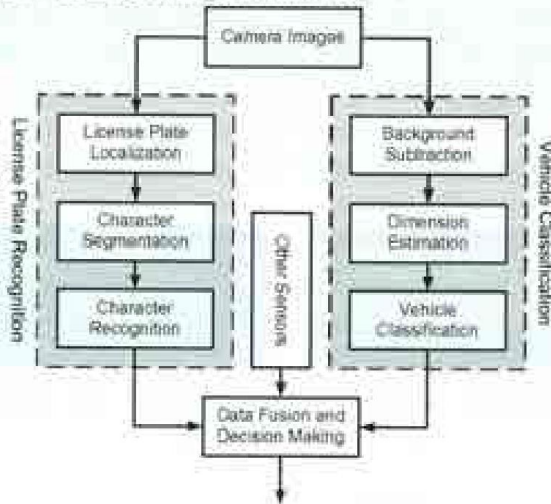

Figure 1. General structure of a typical toll payment system

A. Plate Localization

From the human point of view, the license plate is in general defined as a *"small rectangular plate attached in front of or behind a vehicle, and contains a registration number for identification purposes"*. However, unfortunately, this definition cannot be interpreted by a machine, as it involves other undefined objects such as plate, vehicle... and abstract notions like registration and identification. Therefore, alternative definitions based on descriptors comprehensible for machines are required, leading to different principles of detection. It is necessary to notice that in this phase, it is not strictly required to determine exactly the license plate areas, but only to find out the candidate areas, as the images might contain regions with alphanumeric characters or similar in shape to the license plate. The candidate areas will be analyzed later by character recognition, and together with different factors, to reject the irrelevant ones. The important point is that several potential areas can be given as candidate, but the real plate areas must not be ignored.

By a remark that a license plate is a *"rectangular area with increased occurrence of horizontal and vertical edges"*, an algorithm was introduced by Martinsky [1] on the basis of horizontal and vertical edge projections. Mathematically, the horizontal and vertical edge projections are defined as

$$p_x = \sum_{j=0}^{h-1} f(x, j), \text{ and } p_y = \sum_{i=0}^{w-1} f(i, y),$$

Figure 2. Vertical edge detection filter followed by a vertical projection [1]

where w and h are the width and height of the image, $f(x, y)$ is the gray level of a pixel in the image. From the projections, the license plate areas can be localized horizontally and vertically regarding the places the peaks occur. This analysis is simple to implement, but the stability might be problematical, as the remark mentioned above is usually but not always true. In the same study, different filters were proposed to reinforce the license plate features, such as rank filter and edge detection, because the brightness variation in the license-plate region is more remarkable and more frequent than elsewhere in the captured image of vehicles. As illustrated in Figure 2, a vertical edge detection filter followed by a vertical projection reveals clearly the plate area vertically. Related techniques can be found in other works [2][3][4]. In the study of Wroblewski [5], some indices were further defined for the analysis and rating of the projections.

An interesting detection algorithm proposed by Król and Pamuła [6] took consideration of several characteristics: regions containing characters and contrasting content, probability distribution, brightest and darkest regions, little colored regions, regions

bounded by edged. The decision is made by a fuzzy inference system (FIS). Unlike artificial neural network (ANN) systems which are highly independent on domain-specific knowledge, a fuzzy logic one is knowledge based. The membership functions and fuzzy rules are built empirically. However, similarly to ANN, the ability of learning and adaptation is advantageous, but a training phase is also required to tune the internal parameters of a FIS.

Anagnostopoulos et al. [7] proposed an adaptive segmentation technique named sliding concentric windows. In their study, license plates are viewed as irregularities in the texture of the image and therefore abrupt changes in the local characteristics of the image. This method aims to find the local irregularity in the image using image statistics such as standard deviation and mean value. The algorithm was implemented by calculating statistical measurements of two cascading concentric windows with different sizes moving along the image. If the measurements exceed a manual threshold, then the common central point of these windows is considered as belong to a region of interest.

Another class of vehicle license plate (VLP) detection is based on pattern recognition, where support vector machine (SVM) is used in a significant number of studies. Kim et al. [8] used SVM to analyze the color and texture properties of license plates without any feature extractor, and to locate their bounding boxes applied by a continuous adaptive mean shift algorithm (CAMShift). Beside the advantages such as working directly on color images, robustness against shape distortion, the shortcoming of this method is the adaptation ability in texture scale variation. As a workaround, the training patterns need to be sampled at various scales, making the size of training sample set increased. The frequency-domain techniques are also used to detect textures, such as Fourier transform on short scanning line [9], discrete cosine transform (DCT) [10], Gabor transform [11], wavelet decomposition [12], etc.

Haar-like features were widely applied for object detection [13][14], gaining over other schemes by their robustness against variance of illumination, color, size and position of objects to be identified.

However, one problem of these algorithms is that too many features are included in the classifiers, making the system complex and unstable. Zhang et al. [15] overcame this problem in their detection algorithm by proposing a cascade classifier. First, classifiers using global statistical features (e.g. gradient density, density variance), were constructed through simple learning procedures, excluding 70% of background area. Then the AdaBoost learning algorithm [16] was used to build up the other classifiers based on selected local Haar-like features.
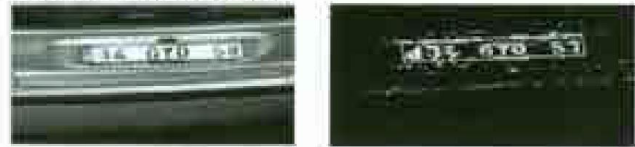
Figure 3. Classifier for text detection using ANN technique [17]
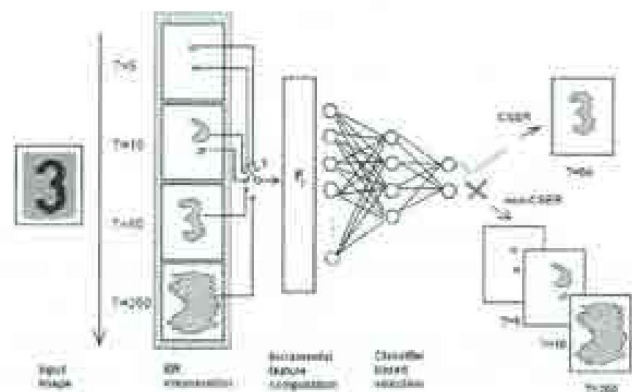
Figure 4. Interleaved enumeration of extremal regions and classification [22]

A closed approach to pattern recognition is text detection, which can be found in [17]. An ANN was trained to classify gray-level pixels as text or non-text (see Figure 3). By this ANN, the original grayscale image is mapped into a BW image with white pixels corresponding to text. After removing noise elements, the VLP zone is identified as a consecutive vertical zone in which the histogram values are above a certain threshold. Nevertheless, the authors did not mention about the detection performance with complicated backgrounds. Another text detector using ANN can be found in [12], where the classification is based on the mean, second-order and third-order central moment features. Other common schemes for text detection are

connected-component based (CC-based) [18][19][20], texture based [18][21].

(a) Input frame (b) MSER+ (c) MSER-

Figure 5. Two variants of MSER used for object detection [25]

Matas and Zimmermann [22] developed an unconstrained method based on extremal regions [23], which are defined as regions of connected components of an image after threshold. Extremal regions have two superior properties, i.e., the set is closed under continuous (and thus perspective) transformations of image coordinates, and under monotonic transformations of image intensities. For implementation, the objects of interest are decomposed into the spatial configuration of category-specific extremal regions; that is, a subset of extremal regions selected by machine learning methods and is likely to correspond to a letter or digit. In the

detection stage, all character-like extremal regions are first independently detected and the longest linear spatial configuration of these regions is labeled as text or license plate. For each enumerated extremal region, a description is computed so that the region can be classified as relevant or irrelevant. The principle of this method is presented in Figure 4. Since this method does not require any restrictive constraint such as text orientation, plate and text colors…, it is robust against geometric changes (scaling, rotation, deformation…), illumination conditions, and even partial occlusions. Also based on region labeling and analysis, the color segmentation in [24] is done in HSI (hue, saturation, intensity) color space.

Inspired by the study of Matas and Zimmermann [22], Donoser et al. [25] developed an effective method based on analyzing the detection results of Maximally Stable Extremal Regions (MSERs) [23][26], but without need of the learning stage. MSERs denote a set of distinguished regions and have proven to be one of the best interest point detectors in computer vision [27]. There are two variants of MSER detection, MSER+ detects bright regions within darker ones, while MSER- detects dark regions within brighter ones. As illustrated by an example in

Table 1. Summary of VLP Detection Approaches

Approach	Advantages	Disadvantages	References
Vertical and horizontal projections	Simplicity	Instability (but can be improved with the help of certain filters)	Martinsky [1], Kwasnicka et al. [2], Kamat and Ganesan [3], Zweng and Kampel [4], Wroblewski [5]
Statistical characteristics	Simplicity, flexibility	Stability dependant on the choice of characteristics in use	Król and Pamula [6], Anagnostopoulos et al. [7]
Pattern recognition	Ability to work with color images, robustness against distortion	Adaptation ability in texture scale variation, complexity, requirement of training	Kim et al. [8], Chen et al. [9], Zhong et al. [10], Ferreira et al. [11], Li et al. [12]
Object detection using Haar-like features	Robustness against variance of illumination, color, size and object position	Complexity, instability	Papageorgiou et al. [13], Viola and Jones [14], Zhang et al. [15], Freund [16]
Text detection	Robustness against distortion	Requirement of training, complexity for good performance	Oz and Ercal [17], Li et al. [12], Jain and Yu [18], Zhou and Lopresti [19], Zhu et al. [20], Wu et al. [21], Jain and Bhattacharjee [28]
Region extraction	Robustness	Complexity, time consuming	Matas and Zimmermann [22], Matas et al. [23], Deb et al. [24], Donoser et al. [25]

Figure 5, the license plate itself can be detected as MSER+, whereas the license characters can be detected as MSER−. Therefore, the algorithm of VLP detection can be achieved by looking for a MSER+ region that contains smaller MSER− ones with nearly equal sizes. Another advantage of this method is that MSER detection results can be further used in the character segmentation phase.

Table 1 summarizes the principal VLP detection approaches, together with their advantages as well as disadvantages.

B. Character Segmentation

There are a number of approaches of plate segmentation that are totally different from each other in nature. For a simplest way, the system can partition the plate into areas regarding a priori known dimensional scale [29]. However, this method needs several strict requirements: the characters are well spatially arranged within the plate, and the plate boundaries are accurately identified.
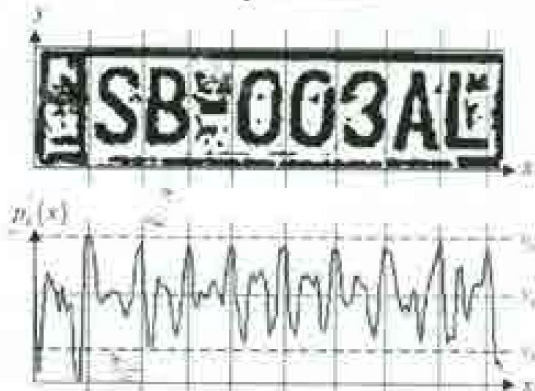

Figure 6. Plate character segmentation based on horizontal projection [1]

A more advanced method that is used in many studies [1][2][17][30][31] is using horizontal projection if the text has only one row, either of a grayscale image or a binary one (after an adaptive threshold), as shown in Figure 6. In the case that the text is composed of two rows, a vertical projection can be applied first to find the separating band. A necessary condition of this method is that the license plate is well aligned horizontally and vertically. More sophisticated studies of segmentation based on connected component (CC) extraction can be found in

[5][7][25]. CC-based methods usually need an additional step to filter out dirty regions that contain noise or other objects placed on the plate. In this step, the number of characters, which is generally known, can be taken into consideration. Alternatively, this can be performed after the recognition phase by rejecting regions in which no valid character is identified. However, it would be more time consuming.

In [32], an adaptive segmentation and extraction approach was proposed for degraded images based on mathematical morphology operators in conjunction with heuristics that determine the potential segmentation points. Morphological thickening and thinning algorithms automatically locate reference lines for separating the overlapped characters, and the baseline for segmenting the connected characters, respectively. This approach can detect fragmented, overlapping, or connected characters and adaptively apply one of three algorithms without manual fine tuning. The results are very promising but, since the algorithm is computationally complex, it is difficult to be proposed for real-time systems.

C. Character Recognition

Once the potential vehicle license plate candidate areas are detected and segmented into smaller regions containing individual characters, the system needs to identify the alphanumeric characters contained in these areas by an OCR method. In general, license plates usually contain a limited number of character compositions following a pre-defined format specific to the country or region of registration.
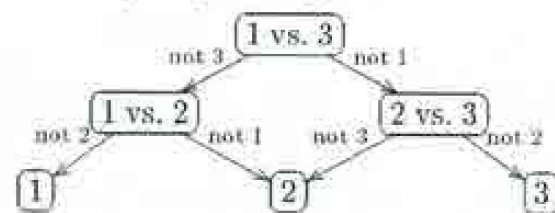

Figure 7. DAG combination for a three-class SVM [36]

Traditionally, as a supervised machine learning mechanism, ANN has largely been used in object classification or pattern recognition in general, and in character categorization in particular. A significant

number of studies [2][5][17] have been devoted for this approach. Multi-layer perceptron neural networks need to be trained by using back-propagation with a large set of samples in order to reach a good performance. The number of neurons in the input layer corresponds to the number of points in the described matrix, and each neuron in the output layer represents a single character recognized by this network. In most studies, images need to be normalized in size, brightness and contrast (even binarized) before training or identified. Besides, the numbers of hidden layers and neurons have to be specified by trial and error, and may be different from this study to another. In general, an ANN with larger size will have a better classification performance, but the computational time will also increase dramatically. Therefore, a compromise should be taken in real-time systems. As mentioned previously, the syntax of vehicle license plate is generally well defined. Therefore, at each position, only a subset of alphanumeric characters is used. From this consideration, a number of ANNs can be established specifically to positions. Beside the

traditional ANN, different variations have also been proposed, such as probabilistic neural networks [7], discrete time cellular neural networks [33], pulse coupled neural networks [34].

In the same class of supervised learning scheme, SVM [25][35] is a more novel alternative method to ANN in character recognition. It gains over ANN by the global optimality. As originally a linear classifier, in order to deal with nonlinear separable data sets, a so-called kernel function is needed to map these data into a linear separable domain. Also, unlike ANNs, an SVM is basically a binary classifier. As a consequence, usually multiple SVMs are combined to build a multi-class one. The organization of this combination is also an interesting topic that has been investigated in the literature. Tapia and Rojas [36] proposed an efficient algorithm on the basis of directed acyclic graph (DAG). An example for the case of three classes is illustrated in Figure 7. Apart from these studies, Huang et al. [31] introduced another machine-learning based classifier that made use of artificial immune clonal selection.

Table 2. Performance of VLP Detection and Recognition Systems in the Literature

Reference	Method D: VLP detection R: VLP recognition	Test dataset size	Accuracy (%) D: detection O: overall
Martín et al. [30]	D: mathematical morphology R: Hausdorff distance	105 (parking at constant distance)	O: 80.4
Kim et al. [8]	D: support vector machine	450 (300 parking, 150 on road)	D: 89
Kwasnicka and Wawrzyniak [2]	D: horizontal projection R: artificial neural network	233	D: 99 O: 47
Oz and Ercal [17]	D: artificial neural network R: artificial neural network	40 (under different weather conditions)	O: 95
Matas and Zimmermann [22]	D: extremal region extraction R: support vector machine	70	D: 94.3
Zhang et al. [15]	D: cascade classifier of global and local features	169	D: 93.5
Anagnostopoulos et al. [7]	D: statistical measurements R: probabilistic neural network	1287	O: 89.1
Deb et al. [24]	D: region extraction	150 (under very different conditions)	D: 94
Arth et al. [35]	D: Haar-like features R: support vector machine	(n/a)	O: 96
Huang et al. [31]	D: horizontal projection R: artificial immune clonal selection	600 (parking, on road, street and petrol stations)	O: 86.5
Zweng and Kampel [4]	D: horizontal projection R: character structural characteristics	1760	D: 100 O: 97.95

Different from machine learning, the recognition by heuristic analysis of character characteristics [1] is fulfilled by comparing the statistical features, such as average brightness, contrast, hue and their deviation. Such an analysis is fundamental and only reliable in recognizing a small character set. Structural analysis [1] is another technique used for character classification. The binary image is first skeletonized, i.e., reduce the character regions in the image to thickness of one pixel, while keeping the connectivity topology. The structural analysis is taken on the so-called chain code that is generated from the skeletonized image. Structural characteristics consist of junction, line end and loop. In the study of Zweng and Kampel [4], a decision tree with leaves representing characters was introduced on the basis of character structural characteristics. A difficulty of this approach is that different characters may have a same structural topology. For example, characters C, G, I, J, M, N, S, U, V, W, Z, 1, 2, 3, 5 and 7 have all two ends, no junction and no loop. Therefore, consideration of additional criteria is necessary to distinguish such characters, such as spatial structure.

D. Summary

Table 2 highlights several important remarks of various vehicle license plate detection and recognition studies in the literature. The projection approach is preferred in many studies and appears to be a good choice regarding the compromise between detection accuracy, system performance and simplicity in implementation. For the recognition task, expert systems are widely accepted, especially ANN and SVM.

III. VEHICLE CLASSIFICATION

As mentioned earlier, the classification can be achieved on the basis either on declared (on OBU) or measured information (e.g., length, height, width, weight) by aid of some sensors installed at toll station, on or beside the road such as camera, laser, and strain gauge. Declared information is highly accurate but unreliable, as the driver could use another OBU that is not for his vehicle. In this case, confirmation by the measured information can help to detect this kind of fraud.

A. Vehicle Detection

To implement the vehicle classification, the ambience background must be subtracted from camera images first to detect moving objects (see Figure 8). As the camera capturing the scene is static, the background of the scene will be analyzed to simplify the next steps in the process of vehicle classification. The background is composed of objects like barrier, road, tree which remain at the same position during time. Even so, due to changes in lighting, wind, weather… conditions, the background might not be stationary. The modeling of background needs to handle quasi-stationary properties and to be as fast as possible for real-time applications. Once the background is modeled, the foreground objects will be able to be segmented, so that the time for vehicle detection can be reduced.

In the study of Ridder et al. [37], a moving object recognition method is described that uses an adaptive background subtraction technique to separate vehicles from the background (see Figure 8). The background is modeled as a slow time-varying image sequence, which allows it to adapt to changes in lighting and weather conditions. In a related work [39], pedestrians are tracked and counted using a single camera. The images from the input image sequence are segmented using background subtraction. The resulting connected regions are then grouped together into pedestrians and tracked. Merging and splitting of regions is treated as a graph optimization problem. Similar algorithms can be found in many other studies [40][41][42][43]. In another study, Buch et al. [44] used Gaussian mixture model (GMM) for motion estimation from video sequences to detect moving vehicles.

Figure 8. Sample camera image and moving-vehicle mask [37]

B. Classification and Dimension Estimation

Once the regions containing moving objects are extracted, each object needs to be identified to be vehicle or not, and in the case it is, classified into predefined categories. Vehicle classification consists of classifying a vehicle into a category as function of some parameters.

One important requirement of vehicle classification is that a same vehicle, whatever the country of origin, must always be classified into the same tariff class. This requirement ensures that there will be no discrimination between users on the basis of nationality. Figure 9 shows an example of vehicle classification based on length and the number of axles. This is a vehicle classification standard of AUSTROAD proposed since 1998. It shows that to classify vehicles, the systems must be able to measure vehicle length and its number of axles.

Figure 9. AUSTROAD classification [38]

There are two approaches of vehicle classification using image processing. One approach is based on vehicle dimension estimation while the other is based on matching extracted features. In the first approach, the vehicle dimensions are estimated from 2D images captured by one or multiple cameras by applying camera calibration. The 3D frame of the vehicle is then reconstructed. After that, the classification can be fulfilled by regarding these dimensional parameters. In the second approach, the classification is done directly on the image by matching the features against those of template models. When needed, the

dimensions can be deduced by a scaling operation from the template regarding the extracted region sizes and its position within the image.
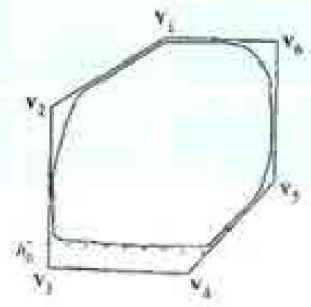

Figure 10.   Edges and vertices of the vehicle bounding box [39]

Since a point in space has three degrees of freedom (dof) but its projection on an image has only two, one algorithm following the first approach needs at least two images of an object in order to determine its position in principle. However, in the case that the point is constrained to a plane, it has only two dof, and a single image is enough for the calibration. For the traffic tracking, the vehicles always run on the road, which can be considered as planar. By this consideration, the bottom vertices ($v_{1..}$ in Figure 10) of the vehicle bounding box can be determined only by interpolation from the coordinates of fixed points on the ground, and without need of camera parameters as well as its model [40][41]. The dimensions are then determined by point-to-point distance regarding the region size and the vehicle position.

In the second approach, several techniques have been developed. In [42], contour matching was used to scoring the similarity between the vehicle image's contour and that of the wireframe model. A related work [44] took consideration of the so-called motion silhouettes detected by GMM instead. Alternatively, local features were also used, such as edge-based features [45] (see Figure 11). Beside fixed-rule matching techniques, general template matching ones like supervised learning [46] are another classification scheme. Dalka and Czyzewski [43] performed a comparison on four frequently-used classification algorithms: K nearest neighbors (KNN), ANN, decision tree (DT) and random forest (RF), on the basis of various vehicle image descriptors.
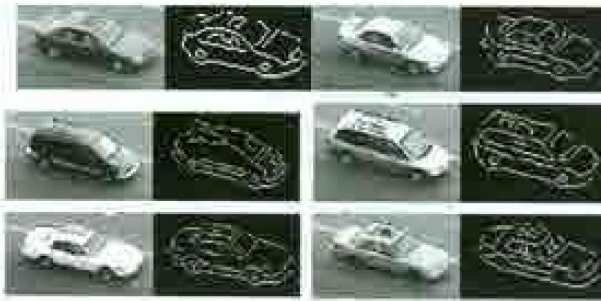
Figure 11. Extracted edge-based features for vehicle classification [45]

C. Summary

Table 3 highlights several important remarks of various vehicle classification studies in the literature. Background extraction is widely chosen for the detection in the most of cases, and all the references reported acceptable results.

IV. RELATED WORKS IN VIETNAM

In Vietnam, there have been a number of related works recently, not only in research projects, but also in commercial deployment. Tran et al. [47][48] introduced a system developed to be implemented at traffic tolls for detection and recognition of license plates of motorcycles, cars and trucks in Vietnam (see Figure 12). In their study, the VLP detection is based on the boundary line detection by combining Hough transform and contour algorithm, projections are used

for the character segmentation, and hidden Markov model for the OCR. The overall performance of the system was reported with accuracy of 92.85%.

Futech Ltd. [49] was one of pioneers in developing a group of systems based on vehicle surveillance using camera, including:

- iParking, a parking access control system: the license number of each vehicle (car, motorcycle or bicycle) is recognized when it comes into the parking, saved and compared when it goes out. The system works only with stationary vehicles; therefore, the vehicles are required to stop for 15 to 30 seconds when they pass the entrance and the exit.

Figure 12. Types of VLP in Vietnam [47][48]

- Speed violation detection system: this system aims to detect the vehicles violating the speed restrictions on roads or at crossing areas, and give information (including recognized license

Table 3. Performance of Vehicle Classifiation Systems in the Literature

Reference	Classification method D: detection C: classification	Test dataset size	Accuracy (%) D: detection C: classification
Lai et al. [41]	D: background subtraction C: calibration	(n/a)	C: 92.5
Gupte et al. [40]	D: background subtraction C: calibration	(n/a)	D: 90 C: 70
Ma et al. [45]	D: background subtraction C: edge-based features	530 (200 sedans, 200 minivans, 130 taxies)	C: 98.5
Ambardekar et al. [42]	D: background subtraction C: contour matching	(n/a)	D: 90
Buch et al. [44]	D: Gaussian mixture model C: model matching	144	D: 94 C: 96.1
Dalka and Czyzewski [43]	D: background subtraction C: feature-based classification	1210 (936 sedans, 154 vans, 120 trucks)	C: 95

numbers) about those vehicles.

- Red-light passing detection system: this system aims to detect the vehicles passing the red light at the crossroads, and give information about those vehicles.

- Black-listed vehicle detection system: this system detects and recognizes the license numbers of the vehicles passing a port, and checks if the vehicle is in a black list given by the police.

Beside a parking access control system that is very similar in functioning to the iParking system of Futech, the Sintech group of Le Quy Don Technical University [50] introduced a toll payment system. Two cameras are used: a front one to detect and identify the license number, and a top one to determine the dimensions and type without requiring the vehicles to stop.

V. CONCLUSIONS

In this paper, different aspects of the two principal common vision-based modules, license plate recognition and vehicle classification, of toll payment systems are discussed, including plate localization, character segmentation and recognition, vehicle detection and classification. This is a preliminary step to setting up a highway toll payment system in further studies. Such a system is normally composed of two main modules, and each module of smaller tasks, each one is addressed by very distinct approaches by different authors. Moreover, not all the tasks were considered in every works, as each system aimed to solve fairly different problems with its own restrictions; and the experiment of each system was carried out in a different way and on a different data set. Hence, it is impossible to make a comparison to assess the features of each system individually. However, in this paper, we would like to give some remarks and recommendations through the considered papers.

For plate localization, horizontal and vertical projections are exposed to be a preferable choice due to their simplicity and good performance. Moreover, it is possible to enhance the robustness by simple filters. In reality, horizontal and vertical projections were

widely applied not only for VLP detection, but they are also a native and effective approach for segmenting the VLP into areas containing individual characters. However, a necessary condition to apply horizontal and vertical projections is that the VLPs are well aligned horizontally and vertically, otherwise more sophisticated methods will be needed to be taken into consideration. For the last phase, OCR, of VLP recognition, machine learning schemes are dominating, especially the well-known ANN and SVM. A major advantage of these schemes is that knowledge on the target patterns is not required in the implementation of these schemes, and the performance can be adjusted via the network size. However, these schemes can differ a lot in implementation and functioning due to the training process, as well as the preprocessing of input images.

Detecting moving objects for vehicle classification is in general done by applying background subtraction. This technique is native, straightforward, and effective, hence widely accepted by authors. For the classification, each of the two approaches, calibration and template matching, has its own advantages and weaknesses. Camera calibration treats the dimensions of a vehicle individually, hence does not require knowledge on the form of vehicle classes. This is important as some kinds of vehicle can vary a lot in form and length. In contrast, calibration-based approach does not take pattern-related features into account, thus might miss useful information. Nevertheless, both approaches can be implemented together in an advanced system to overcome the weaknesses of each one.

REFERENCES

[1] Martinsky, O., 2007. Algorithmic and mathematical principles of automatic number plate recognition systems, B.Sc Thesis, Brno.

[2] Kwasnicka, H., and Wawrzyniak, B., 2002. License plate localization and recognition in camera pictures,

3rd Symposium on Methods of Artificial Intelligence, 243-246.

[3] Kamat, V., and Ganesan, S., 2006. An efficient implementation of the Hough transform for detecting vehicle license plates using DSP's, Real-time technology and applications symposium, pp. 58–59.

[4] Zweng, A., and Kampel, M., 2009. High performance implementation of license plate recognition in image sequences, 5th Int. Symp. on Advances in Visual Computing: Part II, 598-607, Nevada.

[5] Wroblewski, P., 2005. Recognition of Polish car license plates, Computer Recognition Systems, 30, 893-900.

[6] Król, A. and Pamuła, T., 2009. Localization of vehicle license plates in images using fuzzy logic and morphological operations, Transport Problems, 4(1), 95-101.

[7] Anagnostopoulos, C.N., Anagnostopoulos, I., Loumos, V., and Kayafas, E., 2006. A license plate recognition algorithm for intelligent transportation system applications, IEEE Trans. on Intelligent Transportation Systems, 7(3).

[8] Kim, K.I., Jung, K., and Kim, J.H., 2002. Color texture-based object detection: An application to license plate localization, 1st Int. Workshop on Pattern Recognition with Support Vector Machines, 2388, 293-309, Niagara Falls, Canada.

[9] Chen, B.T., Bae, Y., and Kim, T.Y., 1999. Automatic text extraction in digital videos using FFT and neural networks, IEEE Int. Fuzzy Systems Conf., 22-25, Seoul, Korea.

[10] Zhong, Y., Zhang, H.J., and Jain, A.K., 2000. Automatic caption localization in compressed video, IEEE Trans. on Pattern Analysis and Machine, 22(4), 385-392.

[11] Ferreira, S., Mancas-Thillou, C., and Gosselin, B., 2003. From picture to speech: an innovative OCR application for embedded environment, 14th ProRISC workshop on Circuits, Systems and Signal Processing, Veldhoven, Netherland.

[12] Li, H., Doermann, D., and Kia, O., 2000. Automatic text detection and tracking in digital video, IEEE Trans. on Image Processing, 9, 47-56.

[13] Papageorgiou, C.P., Oren, M., and Poggio, T., 1998. A general framework for object detection, 6th Int. Conf. on Computer Vision, 555-562.

[14] Viola, P., and Jones, M., 2001. Rapid object detection using boosted cascade of simple features, IEEE Conf. on Computer Vision and Pattern Recognition, 1, 511-520, Hawaii.

[15] Zhang, H.F., Jia, W.J., He, X.J., and Wu, Q., 2006. Learning-based license plate detection using global and local features, 18th Int. Conf. on Pattern Recognition, 2, 1102-1105.

[16] Freund, Y., 2001. An adaptive version of the boost by majority algorithm, Machine Learning, 43(3), 293-318.

[17] Oz, C., and Ercal, F., 2005. A practical license plate recognition system for real-time environments, Computational Intelligence and Bioinspired Systems, 3512, 881-888.

[18] Jain, A. K., and Yu, B., 1998. Automatic text location in images and video frames, 14th Int. Conf. on Pattern Recognition, 1499-1498.

[19] Zhou, J., and Lopresti, D., 1997. Extracting text from www images, 4th Int. Conf. on Document Analysis and Recognition, 48-50.

[20] Zhu, K., Qi, F., Jiang, R., and Xu, L., 2005. Using Adaboost to detect and segment characters from natural scenes, 1st Int. Workshop on Camera-Based Document Analysis and Recognition, 52-59, Seoul, Korea.

[21] Wu, V., Manmatha, R., and Riseman, E. N., 1997. Finding text in images, 2nd ACM Int. Conf. on Digital Libraries, Images and Multimedia, 3-12.

[22] Matas, J., and Zimmermann, K., 2005. Unconstrained licence plate and text localization and recognition, 8th Int. IEEE Conf. on Intelligent Transportation Systems, 572-577, Wien, Austria.

[23] Matas, J., Chum, O., Urban, M., and Pajdla, T., 2002. Robust wide baseline stereo from maximally stable extremal regions, 2nd Int. Workshop on Biologically Motivated Computer Vision, 1, 384-393, Tübingen, Germany.

[24] Deb, K., Gubarev, V.V., and Jo, K.H., 2007. Vehicle license plate detection algorithm based on color space and geometrical properties, J. of Network and Computer Applications, 30(4), 1324-1333.

[25] Donoser, M., Arth, C., and Bischof, H., 2007. Detecting, tracking and recognizing license plates, 8th Asian Conf. on Computer Vision, 4844, 447-456, Tokyo, Japan.

[26] Forssen, P.-E., and Lowe, D.G., 2007. Shape descriptors for maximally stable extremal regions, IEEE 11th Int. Conf. on Computer Vision, 1-8, Rio de Janeiro.

[27] Mikolajczyk, K., and Schmid, C., 2005. A performance evaluation of local descriptors, IEEE Trans. on Pattern Analysis and Machine Intelligence, 27(10), 1615-1630.

[28] Jain, A. K., and Bhattacharjee, S., 1992. Text segmentation using gabor filters for automatic document processing, Machine Vision Applications, 5, 69-84.

[29] Naito, T., Tsukada, T., Yamada, K., Kozuka, K., and Yamamoto, S., 2000. Robust license-plate recognition method for passing vehicles under outside environment, IEEE Trans. on Vehicular Technology, 49, 2309–2319.

[30] Martin, F., Garcia, M., and Alba, J.L., 2002. New methods for automatic reading of VLP's (vehicle license plates), Int. Conf. Signal Processing, Pattern Recognition, and Applications.

[31] Huang, R., Tawfik, H., and Nagar, A., 2008. Licence plate character recognition using artificial immune technique, Int. Conf. on Computational Science, 5102, 823-832.

[32] Nomuraa, S., Yamanakab, K., Kataia, O., Kawakamia, H., and Shiose, T., 2005. A novel adaptive morphological approach for degraded character image segmentation, Pattern Recognition, 38(11), 1961-1975.

[33] ter Brugge, M.H., Stevens, J.H., Nijhuis, J.A.G., and Spaanenburg, L., 1998. License plate recognition using DTCNNs, IEEE Int. Workshop on Cellular Neural Networks and their Applications, 212-217.

[34] Chacon, M.I., and Zimmerman, A., 2003. License plate location based on a dynamic PCNN scheme, Int. Joint Conf. on Neural Networks, 2, 1195-1200.

[35] Arth, C., Limberger, F., and Bischof, H., Real-time license plate recognition on an embedded DSP-platform, IEEE Conf. on Computer Vision and Pattern Recognition, 1-8, 2007.

[36] Tapia, E., and Rojas, R., 2005. Recognition of on-line handwritten mathematical expressions in the E-Chalk system - an extension, 8th Int. Conf. on Document Analysis and Recognition, 2, 1206-1210.

[37] Ridder, C., Munkelt, O., and Kirchner, H., 1995. Adaptive background estimation and foreground detection using Kalman-filtering, Int. Conf. on Recent Advances in Mechatronics, 193-199.

[38] http://www.transport.sa.gov.au/transport_network/facts_figures/traffic_pdfs/austroads_classes.pdf

[39] Masoud, O., and Papanikolopoulos, N.P., 1997. Robust pedestrian tracking using a model-based approach, IEEE Conf. on Intelligent Transportation Systems, 338-343.

[40] Gupte, S., Masoud, O., R.F.K. Martin, and Papanikolopoulos, N.P., 2002. Detection and classification of vehicles, IEEE Trans. on Intelligent Transportation Systems, 3(1), 37-47.

[41] Lai, A.H.S., Fung, G.S.K., and Yung, N.H.C., 2001. Vehicle type classification from visual-based dimension estimation, IEEE Conf. on Intelligent Transportation Systems, 201-206.

[42] Ambardekar, A., Nicolescu, M., and Bebis, G., 2008. Efficient vehicle tracking and classification for an automated traffic surveillance system, Signal and Image Processing, Hawaii.

[43] Dalka, P., and Czyzewski, A., 2010. Vehicle classification based on soft computing algorithms, 7th Int. Conf. on Rough Sets and Current Trends in Computing, 70-79, Warsaw, Poland.

[44] Buch, N., Yin, F., Orwell, J., Makris, D., and Velastin, S.A., 2009. Urban vehicle tracking using a combined 3D model detector and classifier, 13th Int. Conf. on Knowledge-Based and Intelligent Information & Engineering Systems, 170-177, Santiago, Chile.

[45] Ma, X.X., Eric, W., and Grimson, L., 2005. Edge-based rich representation for vehicle classification, 10th IEEE Int. Conf. on Computer Vision, 2, 1185-1192, Beijing.

[46] Li, Z.C., Jiang, J.M., and Xiao, G.Q., 2009. SVM-based classification of moving objects, 6th Int. Symp. on Multispectral Image Processing & Pattern Recognition, China.

[47] Tran, D.D., Duong, A.D., and Tran, L.H.D., 2004. Combining Hough transform and contour algorithm for detecting vehicles' license-plates, IEEE Int. Symp. on Intelligent Multimedia, Video and Speech Processing, 747-750, Hong Kong, China.

[48] Tran, D.D., Tran, L.H.D., Tran, V.P., and Nguyen V.H., 2005. Building an automatic vehicle license-plate recognition system, Int. Conf. in Computer Science, 59-63, Can Tho, Vietnam.

[49] Homepage: http://www.futech.com.vn/

[50] Homepage: http://www.mta.edu.vn/

AUTHORS' BIOGRAPHIES

Trung-Kien Dao received his BS and MS degrees in Information Technology from University of Pierre and Marie Curie (Paris VI), France in 2004 and 2006, respectively. He received his PhD degree Mechanical and Automation Engineering from Da-Yeh University, Taiwan in 2010. His current research interests are in automation, advanced instrumentation, perceptive environment and computer vision.

Thi-Lan Le graduated in Information Technology from Hanoi University of Science ACnd Technology. She has MS degree in Signal Processing and Communication from Hanoi University of Science and Technology. She received her PhD degree from Nice University, France in 2009. She is currently lecturer/researcher at Computer Vision group, International centre MICA, Hanoi University of Science and Technology. Her main research interests are in content-based indexing and retrieval, video understanding, human-robot interaction.

Thi-Thanh-Hai Tran graduated in Information Technology from Hanoi University of Science and Technology in 2001. She got her MS degree in Imagery Vision and Robotic at Grenoble Institute of Technology in 2002. She received her PhD degree from Grenoble Institute of Technology, France in 2006. She is currently lecturer/researcher at Computer Vision group, International centre MICA, Hanoi University of Science and Technology. Her main research interests are visual object recognition, video understanding, and human-robot interaction.

Ngoc-Thuy Le graduated in Mechatronics from Ho-Chi-Minh City University of Technology in 2004. She got her MS degree in Automatic Control from Grenoble Institute of Technology, France in 2005. She received her PhD degree in Mechanical Engineering from National University of Singapore in 2011. Her current research interests are in image processing.