A Vision-based Serial Number Recognition Algorithm for HSR Trains by Nearest Neighbor Chains of Connected Components

Bo Li*, Bin Tian*, Ye Li*, Gang Xiong*†, and Fenghua Zhu*†
*The State Key Laboratory of Management and Control for Complex Systems
Institute of Automation, Chinese Academy of Sciences
Beijing, China
Email: bo.li@ia.ac.cn
†Dongguan Research Institute of CASIA
Cloud Computing Center, Chinese Academy of Sciences
Dongguan, China

***Abstract*—With the rapid development of High Speed Railway (HSR) in recent years, its research become one hot academic topic. Serial numbers of HSR trains are unique identifications, which play an important role in railway management and operation. In this paper, we present a vision-based algorithm to automatically recognize the serial numbers of HSR trains by image sensors. Firstly, according to the fixed character layout of serial numbers, the serial number regions in the image are located by a combination of connected components. Maximally Stable Extremal Region (MSER) detector is introduced to extract reliable connected components as candidate characters. Based on the pairwise geometry relation between candidates, these connected components constitute the nearest neighbor chain. The accurate location of the serial number is obtained with the nearest neighbor chain. Meanwhile, character images can be segmented simultaneously. Finally, the characters are recognized with Histogram of Gradient (HoG) features and simple similarity-based classifiers. In the experiments, our recognition performance is evaluated by the test images which are collected from real application scenes. The experimental results show the reliability and effectiveness of our serial number recognition algorithm.**

Keywords*—automatic serial number recognition; computer vision; high speed railway

I. Introduction

Owing to the particular advantages of speed and transport capacity, railway is one primary and important means of transportation for people traveling and goods delivery. With the rapid economic and social development of China, construction of the railway network, especially the High Speed Railway (HSR), has become a significant and long-standing issue in recent years. HSR is one new generation of traffic mode, which utilizes advanced technologies in mechanism, control, electronics, and so on. They provide more fast, safety, reliable, and comfortable services for passengers compared with traditional railways. The extensive construction and expansion of HSR will also greatly enhance the sustainable development of China. Under these circumstances, the combination of HSR and some advanced technologies in various fields has attracted the attentions of researchers [1] [2] [3].

This work is supported in part by NSFC project 71232006,61233001, 90920305; and CAS 2F11N01, 2F11N07.

In order to increase the information and intelligence level of HSR management, image sensors are brought into use in the ordinary operation and maintenance of HSR [4] [5]. In this paper, we focus on one specific application which uses image sensors to automatically collect information of HSR trains. The background of our application is described as follows. To ensure the operation security of trains, formal routine check and examination on trains must be done each time before they start to work. When a train enters into the inspection station, the serial numbers on both engine and coaches should be recorded for reference purpose. These serial numbers are similar to the license plate number of vehicles, which represent their unique identification. The previous method is to manually record these numbers. It consumes a lot of time and manpower. Therefore, with the help of technologies on computer vision, image processing, and pattern recognition, we aim to accurately recognize these serial numbers by computers.

The architecture of this automatic serial number recognition system is illustrated in Fig. 1. The monitoring camera is installed along the railway track that leads to the inspection station. Since the serial numbers are all explicitly printed on the body of the engine and coaches in the similar positions, the camera is fixed on a platform with comparable height in a horizontal viewpoint. Thus, when one train passes by the camera site, the sequences of images that contain serial numbers in engine and different coaches are captured. These images are transferred to computers for serial number detection and recognition.

The key module of this system is the algorithm for serial number recognition. We analyze that the algorithm should contain three steps: serial number region detection, character segmentation, and character recognition. Among these procedures, the serial number detection is the most basic one. The detection accuracy will directly influence the overall recognition performance. However, accurate detection is actually one difficult problem in natural and outdoor environment, which have several disturbances like various environmental illuminations, background clutter, and serial number appearance variations caused by different captured viewpoints. Moreover, other texts on the coaches should not be detected as false

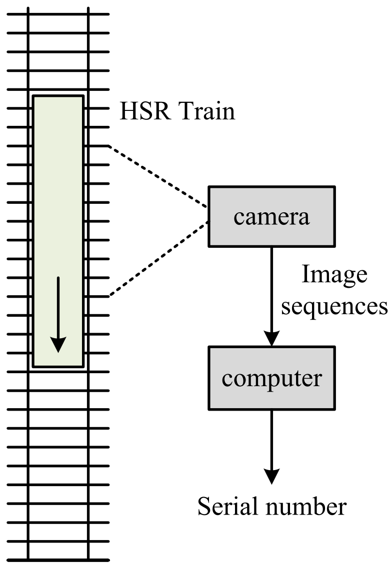


Fig. 1. Architecture of the automatic serial number recognition system.

Fig. 2. Examples of two kinds of serial number. (a) Engine serial number. (b) Coach serial number.

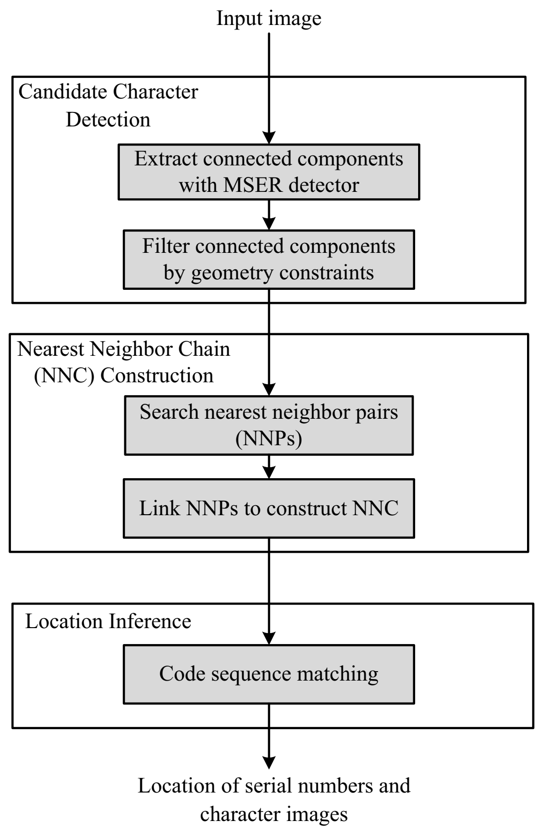


Fig. 3. Flowchart of the serial number region detection.

alarms. As a result, it is a challenging task to achieve effective and reliable serial number detection and recognition.

In this paper, we present a serial number recognition algorithm applied to HSR trains. The flowchart is similar with the license plate recognition algorithms. First of all, the serial number region is located with the prior knowledge about the character layout. Characters are detected and assembled to form the serial numbers by their geometry attributes. Through our component-based detection process, all the character images in one serial number are all segmented. Finally, characters are recognized based on shape features. The remainder of this paper is organized as follows. Section II explains the serial number region detection approach in detail. Then, the character recognition is presented in the Section III. Section IV describes the experiments which are performed to evaluate the performance of the proposed algorithm in practical applications. Finally, we briefly make the conclusion in the Section V.

II. Serial Number Region Localization

A. Overview

Our serial number region detection approach is motivated by the component-based license plate detection algorithms [6] [7] [8] [9] [10] and text detection algorithms [11] [12]. Specially, the method in this paper is based on some idea from [7] and [8]. We define the serial numbers as regions composed of several characters that are arranged in fixed layouts. Two kinds of character arrangement for serial numbers on the engine and coaches are respectively shown in Fig. 2. In this figure, it is observed that serial numbers consist of English letters and numerals with similar height and width (except for the numeral 1 and the small numeral after the string "CRH" in engine serial number). Thus, ideally, the characters can be segmented as connected components. However, similar-sized connected components may be detected as other objects in the image. Hence, the contextual information among characters is utilized to seek the collection of connected components, whose spatial arrangement is in accord with the standard serial number. This detection method can overcome most of the impact from background clutter and appearance variation of serial numbers.

The flowchart of our detection approach is shown in Fig. 3. Firstly, connected components with reasonable size are detected as candidate characters. Next, according to the geometry attributes between each pair of connected components, the nearest neighbor chain of connected components is defined and constructed. Finally, the inference on this chain is performed to obtain the location, as well as the compositional character images of the serial number. The character segmentation is finished simultaneously with the detection.

B. Candidate Character Detection

Generally, connected components are extracted from binary images by simple threshold segmentation. However, simple threshold segmentation method is sensitive to the variation of environmental illumination and it restricts the capability in practical applications. Therefore, we adopt Maximally Stable

Extremal Region (MSER) detector [13], which is a regional feature extraction algorithm. It has several particular advantages comparing with many other feature detectors [14]. In our task, the main superiority is its invariance to illumination variation. The MSERs are highly robust with monotonic illumination changes in the environment. For images captured both at day and night, MSER detector can extract reliable connected components without adjusting any parameters.

MSER detection is similar to the watershed algorithm. A sequence of intensity thresholds ranging from 0 to 255 is applied to obtain a series of binary images. In this sequential binarization, MSERs are defined as the connected components at the same location with local minimum area variations. Since the threshold sequence can go in the opposite directions, two kinds of MSERs are obtained: bright MSERs and dark MSERs. Bright MSERs mean that intensities of pixels inside the blobs are higher than those boundary pixels. Dark MSERs are reverse. In our task, since all the serial numbers are black, only the dark MSERs are remained for subsequent process.

After MSER extraction, all the connected components are filtered by certain geometric constraints, such as height, width, area, and aspect ratio. The remaining blob regions are considered as candidate characters in serial numbers. Results of candidate character detection are shown in Fig. 5(a) and Fig. 6(a). Bounding boxes of the MSERs are marked by red rectangular boxes. It is noted that most of the characters in serial number regions are extracted, but many character-like regions in the background are also detected. Hence, further character selection is necessary for serial number detection.

C. Nearest Neighbor Chain Construction

In order to find a set of candidate characters with similar layout of serial numbers, we define the Nearest Neighbor Chain (NNC), which is constructed by several connected components with similar size in neighborhood [6]. One NNC determines a hypothesis of serial number in the next inference procedure.

We give the definition of the nearest neighbor pair (NNP) on candidate characters C_i and C_j in [8] at first. Important attributes for describing C_i and its relationship with C_j are enumerated in Table I and intuitively illustrated in Fig. 4. In consideration of the possible numeral "1" in license plates, connected components with an elongated shape (aspect ratio over 2.5) are also included as candidate characters. However, their widths are not consistent with regular characters. As a result, mean height h rather than width is used as a relative measurement of distance. Constraints and parameters for the neighboring pair (C_i, C_j) are detailed in Table II. The five constraints are respectively defined in the aspect of Euclidean distance, x-coordinates, vertical distance, height difference, and angle of inclination. The parameters are designed based on the character layout prior in serial number to ensure that the two candidates belong to one serial number. For each candidate character, other candidates are searched on the right side to form the NNP. The order in x-coordinates is also recorded as $(C_i \rightarrow C_j)$ in NNC with C_i in the left. If more than one candidate character satisfies the conditions in Table II, only the nearest one is kept.

TABLE I. MATHEMATICAL ITEMS AND NOTIONS FOR CANDIDATE CHARACTERS

Item	Notion
C_i	One candidate character
(cx_i, cy_i)	Coordinates of the center for C_i
h_i	Height of the bounding box for C_i
$d(C_i, C_j)$	Euclidean distance between C_i and C_j
$d_x(C_i, C_j)$	Horizontal distance between C_i and C_j
$d_y(C_i, C_j)$	Vertical distance between C_i and C_j
$\bar{h}$	Mean height of C_i and C_j

TABLE II. CONSTRAINTS FOR NEAREST NEIGHBOR PAIR

No.	Constraints	Parameters
1	$k_1 \cdot \bar{h} \leq d(C_i, C_j) \leq k_2 \cdot \bar{h}$	$k_1 = 0.3$, $k_2 = 2.5$
2	$cx_i < cx_j$	-
3	$d_y(C_i, C_j) < k_3 \cdot \bar{h}$	$k_3 = 0.5$
4	$\|h_i - h_j\| < k_4 \cdot \bar{h}$	$k_4 = 0.3$
5	$tan^{-1} \frac{d_y(C_i, C_j)}{d_x(C_i, C_j)} < \theta$	$\theta = \pi/4$

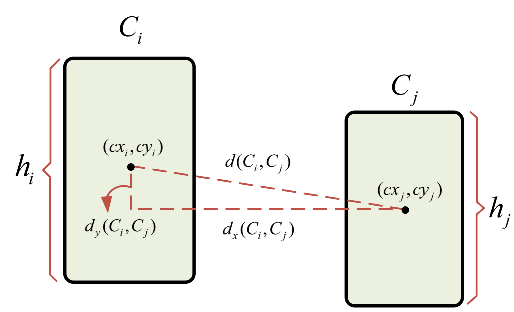


Fig. 4. Geometric attributes for a pair of candidate characters.

After all NNPs in the image are searched out, the matched start and end components are emerged and connected together. In this way, the K connected components construct a NNC in length K. For example, if (C_i, C_j), (C_j, C_k) and (C_i, C_m) are three NNPs, then they construct the NNC $\{C_i, C_j, C_k, C_m\}$. Eventually, several NNCs are chosen from the image. Each chain contains a cluster of similar-sized candidate characters. Actually, the shortest one is a NNP with two connected components. To remove the effects of noise in the background, we ignore the NNC with less than 4 characters. As a result, the NNCs primarily decide the locations of serial number regions.

Results of NNC construction are illustrated in Fig. 5(b) and Fig. 6(b). Yellow dots and lines denote connected components in the NNC and links between NNPs respectively. It is observed that after NNC construction, the remaining candidate characters are all located around the serial number regions. Most of the connected components in the background are successfully removed.

D. Location Inference on Nearest Neighbor Chain

The inference on NNC is to obtain the accurate location and specific character combination of the serial number. We represent one NNC as a special code sequence, which is taken to compared with the standard code sequence. As a result, the correct sequence matching means the presence of a serial number.

Fig. 5. Example for engine serial number detection. (a) Candidate character detection result. (b) Nearest neighbor chain construction result. (c) Final detection result.

Fig. 6. Example for coach serial number detection. (a) Candidate character detection result. (b) Nearest neighbor chain construction result. (c) Final detection result.

The NNC is encoded according to the geometry relation between pairwise connected components. At the beginning, the code sequence cq is set as 1. Then, for each pair of connected components in NNC, we assign a flag f based on Euclidean distance between them, as in (1). $f = 1$ indicates that the two connected components correspond to adjacent characters, while $f = 01$ means there is a missing character in the middle. With the flag assignment in order, the value of f is put at the end of cq in turn. Finally, we obtain a code sequence cq to describe the NNC.

$$f = \begin{cases} 1 & d(C_i, C_j) < 1.2\bar{h} \\ 01 & otherwise \end{cases} \quad (1)$$

Since the fixed arrangements of serial number characters are known, we can also use (1) to transform them into benchmark code sequences for matching. The matching strategy is simple. If the benchmark code sequence has the same length with the new one, they are considered as a correct match. If a test code sequence is successfully matched, the region of the serial number is located. Connected components in this NNC are taken as characters of a serial number. Meanwhile, the missing detected characters will be estimated by the standard character layout.

Specially, for the engine serial number, the numeral followed the character string "CRH" is much smaller than the other characters. This character will be ignored in above processes. In order to capture this particular character image, connected component extraction is performed once again by MSER detector on a local Region Of Interest (ROI) in the image. We set the ROI as the region between the third and the fourth character. Considering that the ROI is quite small, the detected connected component that satisfy geometry constraints is directly utilized as the supplementary character image of the engine serial number.

The final detection results are marked with green boxes in Fig. 5(c) and Fig. 6(c). All the serial number regions are exactly detected with proper segmentation of character images. Character recognition can be directly performed on these segmented images.

III. Character Recognition

Character recognition is a typical pattern recognition problem. Therefore, our solution for this problem is along the procedure of general multi-category classification system. The character recognition is implemented following the way of [7]. Firstly, the character image is mapped to the feature vector by certain feature descriptor. Then, the proper character category is assigned to the feature vector by classifier that is trained in advance.

Before feature extraction, the segmented character images are normalized to uniform size of 20×40. Then, we use the Histogram of Gradients (HOG) [15] feature descriptors to describe the contour of characters. HoG computes histogram of gradient orientation on a dense grid of uniformly spaced cells. Local contrast normalization is also conducted in overlapped blocks to improve its robustness to illumination. Due to the representation of dense orientation field map, HoG can effectively describe the local appearance and shape of an object. Since the HoG descriptor computes statistical histograms on local cells, the method keeps nice invariance to geometric and photometric transformations. In our implementation, we use 5×10 pixel cells and 2×2 cell blocks with 9 histogram channels. As a result, one feature vector of 324 dimensions is obtained for each normalized character image.

Before character recognition, a large amount of character image examples are collected to train classifiers at first. In this paper, we use a simple multi-category classification model based on feature centers. For each character category, we extract all the feature vectors for images and compute their mean vector μ_i. Thus, we can get the collection of mean vectors $\{\mu_1, \mu_2, ..., \mu_N\}$, where N denotes the number of the complete character categories. These mean vectors are noted as model files for character recognition.

If a new character image is to be recognized, its HoG feature descriptor μ is computed firstly. Then, the category label c is assigned as the category whose mean vector is the closest one to μ using the similarity metric of Euclidean distance, as in (2).

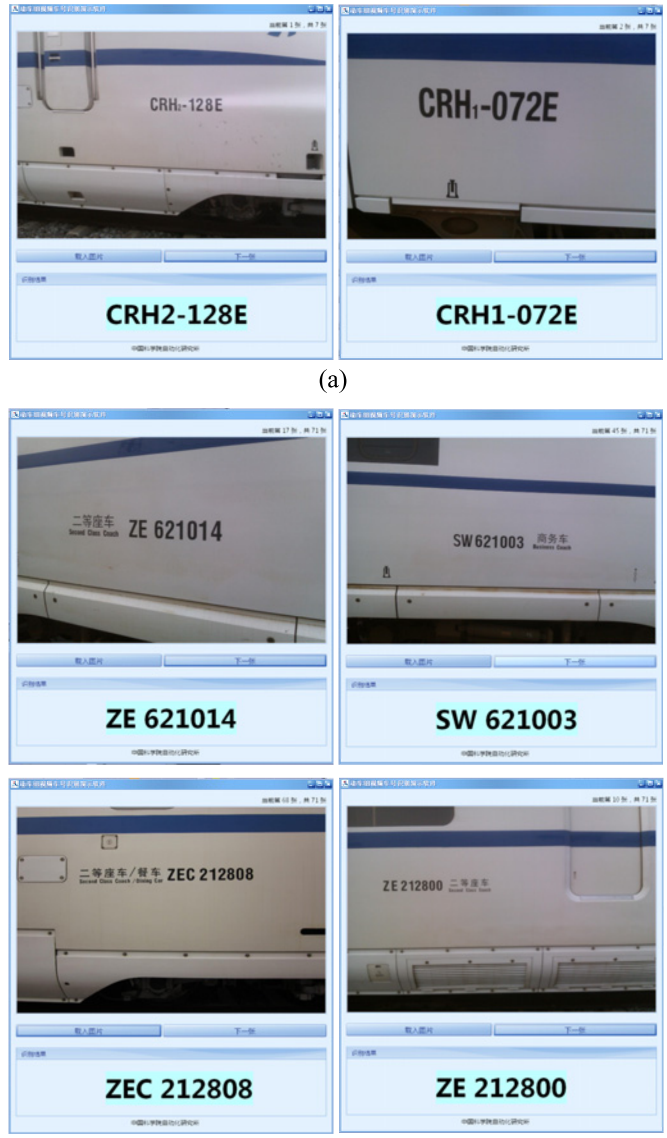
$$c = \underset{i=1,2,...,C}{\arg\min} ||\mu_i - \mu||. \quad (2)$$

It is noted that the elongated connected components are directly recognized as the numeral "1" without feature extraction and matching. The parameter C in (2) denotes the available search scope of the character categories. $C = N$ is set for ordinary character images. However, characters in certain places of the serial number may have the fixed choices to represent some particular meanings. For example, the last character in engine serial number must be A, B, or E. Hence, the scope for available character category is greatly refined. Owing to the regular appearances of these printed characters, this simple classification approach can achieve reliable recognition results as well as low computational complexity in applications.

IV. Experiments

The performance of this serial number recognition algorithm is evaluated on our own collected image dataset. Test images in this dataset are captured by high-resolution digital cameras installed at real application environment. In order to evaluate the robustness of our approach in different environmental conditions, the images are captured in various viewpoints and illumination. The dataset contains 184 images with resolution of 2592×1728. Only one serial number exists per image. Meanwhile, the width of all the serial numbers is limited to 400 pixels to 600 pixels. In order to perform a comprehensive evaluation, all test images are fully annotated. We manually label the ground truth bounding boxes and contents of every serial numbers. The evaluation criterion for serial number detection is the overlap area between our detection bounding box and the marked benchmark region. If the overlap area is over 90% area of their unions, we can say the serial number is correctly localized. The recognition performance is directly tested by comparing the recognized character string with the benchmark one. In our experiments, no restriction is on the prior locations of the serial numbers and the parameters need not to be changed during operation.

All of the functional modules of this system are embedded in a dynamic link library (DLL), which can be easily imported by other projects. To ensure the realtime operation efficiency, the DLL is developed using C++ language and OpenCV 2.3. The experiment is performed on a PC with Intel Core2 Quad Q8400 CPU 2.66GHz and 3GB RAM. In experiments



(a) (b)

Fig. 7. Examples of serial number recognition results on the interactive software platform. (a) Engine serial number recognition. (b) Coach serial number recognition.

on our collected dataset, 98.9% overall recognition accuracy is achieved with no false positives. In practical application, we have developed an interactive software platform with C# language based on .Net 2.0 framework. The DLL is imported by the platform for better result visualization. Some examples of recognition results on this platform are shown in Fig. 7. Most of the time is consumed on the MSERs extraction. The experimental results show that our proposed approach is reliable and effective to be applied in practical HSR train management.

V. Conclusion

In this paper, a vision-based serial numbers recognition algorithm is proposed for intelligent management of HSR trains. Firstly, regions of serial numbers are located based on their certain character layouts. In detection, the connected components in the image are extracted by MSER detector as candidate characters. Pairwise geometry relations between candidates are analyzed to constitute a nearest neighbor chain, which indicates the accurate location of the serial number. The character images are segmented simultaneously with the serial number detection. Then, they are recognized using HoG features and simple similarity-based classification model. In the experimental section, our recognition performance is evaluated by dataset which are collected from real application scenes. We achieve approximately 99% overall recognition accuracy in this dataset, which proves the reliability and effectiveness of our serial number recognition algorithm. Based on the component-based serial number detection, our approach has the advantages of robustness in diverse viewpoints,background clutter, and variant environmental illumination.In the future, some hardware acceleration strategies such as parallel computation can be considered to improve the efficiency.

ACKNOWLEDGMENT

We would like to thank professor Fei-Yue Wang for his instruction and encouragement to our work.

REFERENCES

[1] H. Dong, B. Ning, B. Cai, and Z. Hou, "Automatic train control system development and simulation for high-speed railways," *IEEE Circuits Syst. Mag.*, vol. 10, no. 2, pp. 6–18, 2010.

[2] S. Leva, A. P. Morando, and P. Colombaioni, "Dynamic analysis of a high-speed train," *IEEE Trans. Veh. Technol.*, vol. 57, no. 1, pp. 107–119, 2008.

[3] B. Ning, T. Tang, H. Dong, D. Wen, D. Liu, S. Gao, and J. Wang, "An introduction to parallel control and management for high-speed railway systems," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 4, pp. 1473–1483, 2011.

[4] F. Marino, A. Distante, P. L. Mazzeo, and E. Stella, "A real-time visual inspection system for railway maintenance: automatic hexagonal-headed bolts detection," *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.*, vol. 37, no. 3, pp. 418–428, 2007.

[5] S. Piva, M. Zara, G. Gera, and C. S. Regazzoni, "Color-based video stabilization for real-time on-board object detection on high-speed trains," in *Proc. IEEE Conf. Adv. Video Signal Based Surveill.*, 2003, pp. 299–304.

[6] L.-G. Miao, "License plate detection algorithm based on nearest neighbor chains," *Acta Autom. Sin.*, vol. 37, no. 10, pp. 1272–1280, 2011.

[7] B. Li, B. Tian, Q. Yao, and K. Wang, "A vehicle license plate recognition system based on analysis of maximally stable extremal regions," in *Proc. IEEE Int. Conf. Networking Sens. Control*, 2012, pp. 399–404.

[8] B. Li, B. Tian, Y. Li, and D. Wen, "Component-based license plate detection using conditional random field model," *IEEE Trans. Intell. Transp. Syst.*, in press.

[9] W. Zhou, H. Li, Y. Lu, and Q. Tian, "Principal visual word discovery for automatic license plate detection," *IEEE Trans. Image Process.*, vol. 21, no. 9, pp. 4269–4279, 2012.

[10] Y. Wen, Y. Lu, J. Yan, Z. Zhou, K. M. v. Deneen, and P. Shi, "An algorithm for license plate recognition applied to intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 830–845, 2011.

[11] D.-Q. Zhang and S.-F. Chang, "Learning to detect scene text using a higher-order mrf with belief propagation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2004, pp. 101–106.

[12] H. Chen, S. S. Tsai, G. Schroth, D. M. Chen, R. Grzeszczuk, and B. Girod, "Robust text detection in natural images with edge-enhanced maximally stable extremal regions," in *Proc. IEEE Int. Conf. Image Process.*, 2011, pp. 2609–2612.

[13] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide baseline stereo from maximally regions," in *Proc. Br. Mach. Vis. Conf.*, 2002, pp. 384–396.

[14] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. V. Gool, "A comparison of affine region detectors," *Int. J. Comput. Vis.*, vol. 65, no. 1-2, pp. 43–72, 2005.

[15] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 886– 893.