Vehicle License Plate Recognition Based on Extremal Regions and Restricted Boltzmann Machines

Chao Gou, Kunfeng Wang, Yanjie Yao, and Zhengxi Li

***Abstract*—This paper presents a vehicle license plate recognition method based on character-specific extremal regions (ERs) and hybrid discriminative restricted Boltzmann machines (HDRBMs). First, coarse license plate detection (LPD) is performed by top-hat transformation, vertical edge detection, morphological operations, and various validations. Then, character-specific ERs are extracted as character regions in license plate candidates. Followed by suitable selection of ERs, the segmentation of characters and coarse-to-fine LPD are achieved simultaneously. Finally, an offline trained pattern classifier of HDRBM is applied to recognize the characters. The proposed method is robust to illumination changes and weather conditions during 24 h or one day. Experimental results on thorough data sets are reported to demonstrate the effectiveness of the proposed approach in complex traffic environments.**

***Index Terms*—License plate detection, license plate recognition, hybrid discriminative restricted Boltzmann machines, extremal regions.**

I. Introduction

LICENSE plate recognition (LPR) is an important research topic in intelligent transportation systems (ITS) and becomes more and more useful in many applications during the past decades. All vehicles around the world should have a license number as their principal identifier. With the rapid development of computer vision technology, more and more vision-based license plate recognition methods are applied in ITS such as electronic payment systems, traffic activity monitoring and automatic vehicle ticketing.

Although significant progress of LPR techniques has been made in the last decade and various commercial products are reliable under some ideal environments, it is still a challenging task to recognize license plates from complex images. A robust system should work effectively under a variety of conditions such as sunny day, night time or with different colors and complex backgrounds as shown in Fig. 12 [1].

To our knowledge, there is a lack of researches on both vehicle license plates detection and recognition in these challenging traffic scenes. And many methods are proposed to deal with images with low resolution in local view [2], [3]. Many other works have been done only to locate the license plates [4]–[7]. In general, LPR consists of three parts: license plate detection (LPD), character segmentation, and character recognition.

Accurate license plate localization and segmentation are crucial for the entire LPR system. There are two major methods for the localization of vehicle license plates: one method is based on color information [2], [7]–[9], and another is based on textures or edges of the license plates [4], [6]. The color combination of a license plate and its characters is specific, and this combination occurs almost only in the license plate region [2]. In [9], Ashtari *et al.* proposed to detect the Iranian license plate by using a geometric template on connected target pixels with the same color. However, the methods that use color features to localize license plates may become invalid when there are regions in the image whose color information is similar to that of the license plate. Moreover, license plate detection based on color information is sensitive to adverse illumination conditions and camera settings. On the other hand, texture-based methods use high edge-density areas where color or intensity transition occurs dramatically. In [6], the sliding concentric window is applied to find vertical and horizontal edges from vehicle region on the basis of statistical measurement of standard deviation. These methods can detect license plate regions in relatively simple environments, but can easily be affected by noises and are computationally complex when there are many edges in the image.

Characters are segmented after license plate detection. Jiao *et al.* [10] propose to use grey-level quantization and morphology analysis to obtain candidate characters. In [11], the extracted license plate is rescaled to a template size, while in the template all the character positions are known. This method is incapable of dealing with any shift in the extracted license plates. Considering that the characters and license plate backgrounds have different colors, some methods [2], [3] project the extracted binary license plate vertically to determine the starting and the ending positions of characters, and then project the extracted characters horizontally to extract each character alone. In addition, the projection method is also applied when the extraction is not successful because of discontinuity and

Manuscript received February 6, 2015; revised May 31, 2015 and August 10, 2015; accepted October 19, 2015. Date of publication November 18, 2015; date of current version March 25, 2016. This work was supported in part by the National Natural Science Foundation of China under Grant 61304200 and in part by the MIIT Project of Internet of Things Development Fund under Grant 1F15E02. The Associate Editor for this paper was H. Huang. *(Corresponding author: Kunfeng Wang.)*

C. Gou is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with Qingdao Academy of Intelligent Industries, Qingdao 266109, China (e-mail: gouchao.cas@gmail.com).

K. Wang is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: kunfeng.wang@ia.ac.cn).

Y. Yao is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the Research Center for Computational Experiments and Parallel Systems, National University of Defense Technology, Changsha 410073, China (e-mail: yanjie.yao@ia.ac.cn).

Z. Li is with North China University of Technology, Beijing 100041, China (e-mail: lzx@ncut.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TITS.2015.2496545

connectivity. The projection method is common and simple, but is dependent of their accurate positions. It is obvious that these methods are reliable for clean plates but not robust to complex backgrounds with ambiguous characters.

The recognition process for classifying characters comes to the third step. It is difficult to recognize license plate characters because different camera zooms lead to different character sizes. Moreover, the extracted characters are very small and may be similar in their shapes, such as the pairs of S-5, C-G, and D-0. A large number of character recognition methods have been proposed, including neural networks [3], [10]–[12], support vector machine (SVM) [13], character templates [14], [15] and so on. The conventional neural networks and SVM methods have a high correct rate when used for character recognition, but they need a proper data set and a good set of features. Mostly used linear SVM is not so time-consuming but is not reliable for classifying Chinese characters which consists of more than 30 classes. The character templates method is unable to deal with the complexity of various positions and conditions of characters.

In this paper, an effective vehicle license plate recognition method is proposed, based on character-specific Extremal Regions (ERs) and hybrid discriminative restricted Boltzmann machines (HDRBM). The flowchart of the method is illustrated in Fig. 1. The method consists of four steps. Firstly, top-hat transformation is adopted to preprocess the input image and restrain background noises, followed by Sobel filter to find the vertical edges, simple background curve removing and morphological operations to remove blank spaces between two adjacent vertical edge lines. Then the coarse license plate (LP) detection is achieved by filtering different sizes of rectangular regions through geometrical validations. Secondly, character-specific ERs are selected as character regions through a Real AdaBoost classifier with decision trees. In the third step, accurate character segmentation and LP location are achieved based on the geometrical attributes of characters in standard license plates. Finally, features of the character regions, namely histogram of oriented gradients (HOG) descriptors, are extracted from the input image and the characters are recognized using an off-line trained classifier based on HDRBM.

The innovative contributions of this paper are summarized as follows: 1) a new coarse license plate detection method along with a sequence of operations including morphological operations, various filters, different contours and validations; 2) selection of suitable character-specific regions as license plate characters in color space; 3) integration of HDRBM for license plate character recognition; 4) thorough experiments and evaluations on a set of real images with low and high definitions, and day and nighttime illuminations. The conference version of part of this work was published in [16] as an oral presentation. Compared with [16], improvements are achieved by optimizing parameters and proposing a more effective recognition algorithm. More technical details of license plate localization, character recognition, and experimental evaluations are provided in this paper.

The remainder of this paper is arranged as follows. Section II explains coarse license plate detection. Character-specific ERs extraction and fine license plate detection come to Section III.

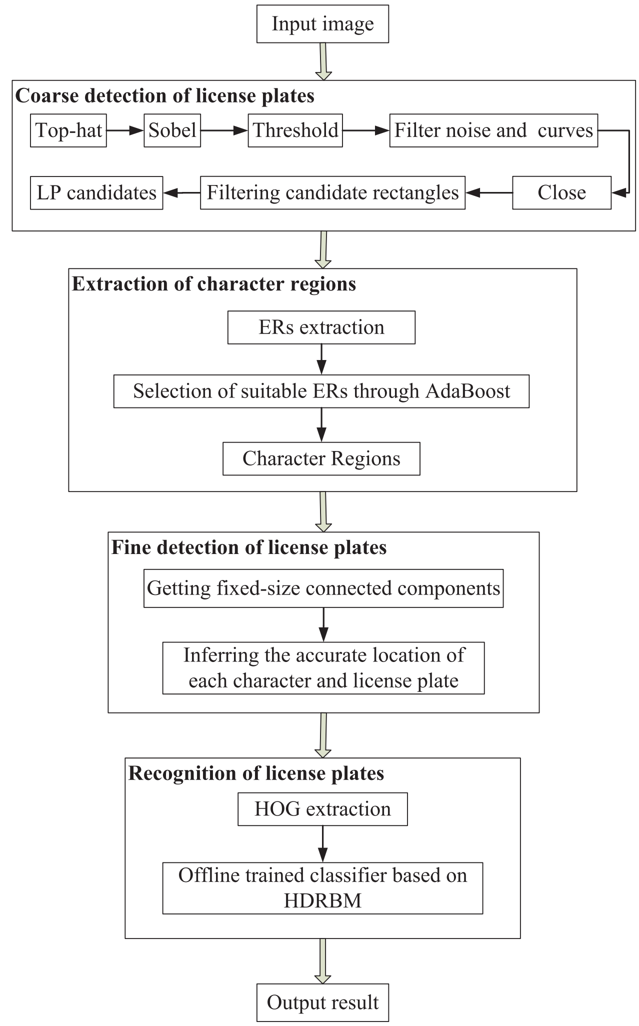


Fig. 1. The flowchart of the proposed approach.

License plate recognition based on HOG features and HDRBM are presented in Section IV. Experimental results are discussed in Section V. The conclusion is drawn in Section VI.

II. Coarse Detection of License Plates

License plate localization aims at distinguishing license plates from other regions of the image. The proposed framework for coarse localization of license plates is shown in Fig. 1. The details of the framework are given in the following subsections.

A. Top-Hat

Firstly we perform top-hat transformation in the gray image converted from the input image. Top-hat is a mathematical morphological transformation and is a non-linear filter [17]. It has multiple functions such as restraining noises, extracting features, segmenting images, etc. The top-hat transformation equation is formulated in

$$\text{Top\_hat}(A) = A - (A \circ B). \tag{1}$$

For aggregate A and structure B, $A \circ B$ is defined as an open arithmetic in mathematical morphology, which helps to remove

(a)

(b)

Fig. 2. An example of top-hat transformation (pixels with value above 5 are added by 30 for demonstration).

small objects and to smooth edges of the input image. The open arithmetic equation is formulated as $A \circ B = (A\Theta B) \oplus B$. $A\Theta B$ and $A \oplus B$ are both morphological arithmetic, called erosion and dilation, respectively. Erosion is used to eliminate isolated noises and its equation is shown in (2). Dilation is used to combine the object with the background points and its equation is shown in (3)

$$X = A\Theta B = X : B + X \subset A \tag{2}$$

$$X = A \oplus B = X : (-B + X) \cap A \neq \varnothing. \tag{3}$$

An efficient character region enhancement is crucial for the proposed character region-based license plate detection. Top-hat transformation could remove some noise for images with complex backgrounds and extract bright edges. In addition, top-hat not only enhances the image details especially the character regions in dim conditions but also enlarges the contrast between character regions and backgrounds. The result of top-hat transformation for a night time test image is shown in Fig. 2. Obviously, the character regions are enhanced. This is crucial to conduct license plate location (the final detection results is shown in Fig. 12). The performance of top-hat transformation mostly depends on a square structuring element whose width is 3 pixels.

B. Coarse Location of License Plates

As shown in Fig. 2 after top-hat transformation, a character region with dense vertical edges is more likely to be a plate candidate. Hence, the Sobel filter is used to get vertical edges as shown in Fig. 3(b). A binary method with a global threshold cannot generate satisfactory results because the brightness distribution of a LP may vary due to lighting environment. In light of that, the adaptive local binary method is applied. In this paper, similar to [18], the local Otsu method is used to obtain binary image and the result is shown in Fig. 3(c). But there are many short random noise and long bounding curves besides the license plate edges. According to [4], the too long (background curve) and too short (noise) edge length can be removed (see Fig. 3(d)). Close operation, which is one morphological operation, is applied to connect text pixels into plate regions. Once extracting the bounding rectangles of external contours (see Fig. 3(e)), a filtering method based on aspect ratio and area is applied to get regions that may contain license plates. The filtering method for coarse location is to set the thresholds of bounding rectangles. In this paper, a bounding rectangles with area between 1000 and 16000, aspect ratio (width/height) between 1.5 and 8 and width above 70 is selected as a LP candidate (see Figs. 3(f) and 7).

A high recall of license plates is essential to the whole LP recognition system. As described above, coarse location method is robust to dim and bright light conditions (see Figs. 2 and 3). This method is also effective in highlight conditions with noises as shown in Fig. 3. Moreover, the precision of location is high corresponding to a high recall. More experimental results are demonstrated in Section V.

III. Extraction of Character Regions and Fine Detection of License Plates

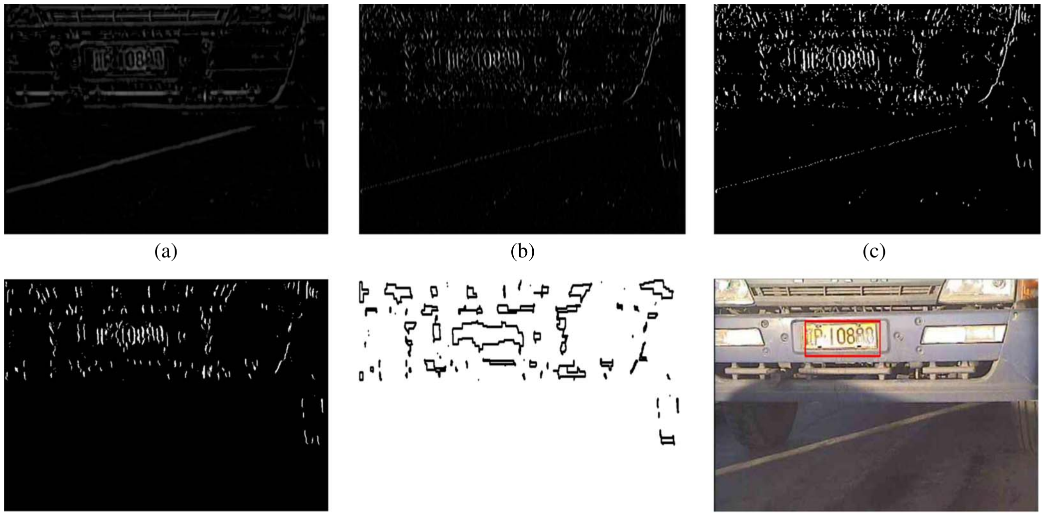
It is crucial to obtain the precise bounding box of each character region before recognition. In this paper, we propose to select suitable character-specific extremal regions (ERs) as character candidates in each channel of color space. After each character region is segmented, accurate LP location can be achieved. The details of coarse-to-fine license plate detection based on character-specific regions are given in this section.

A. Extraction of Character Regions

1) Extremal Regions: A color image can be considered as a mapping $I : D \subset M \times N \to V$ where V indicates three channels represented by $\{0, \ldots, 255\}^3$. A channel C of the image I is a mapping $C : D \to P$ where P is a totally ordered set and $f_c : V \to P$ is a projection of pixel values to an ordered set. Region R of an image I is a contiguous subset of D. The outer boundary ∂R of R is a set of pixels adjacent to R (4-connected pixels) but do not belong to R. Extremal Region (ER) is a region whose outer boundary pixels have strictly higher values than the region itself, i.e., $\forall\, m \in R, n \in \partial R : C(m) \leq t < C(n)$, where t denotes the threshold of the ER. As presented in [19], an ER r at threshold t is a union of one or more ERs R_{t-1} at threshold $t-1$ and pixels of value t formulated in (4)

$$r = \left(\bigcup u \in R_{t-1}\right) \cup \left(\bigcup m \in D : C(m) = t\right). \tag{4}$$

2) Detection of Character Regions: Maximally Stable Extremal Region (MSER) [20], whose size remains unchanged over a range of thresholds in grey level, is a special case of ER.



(a) (b) (c)

(d) (e) (f)

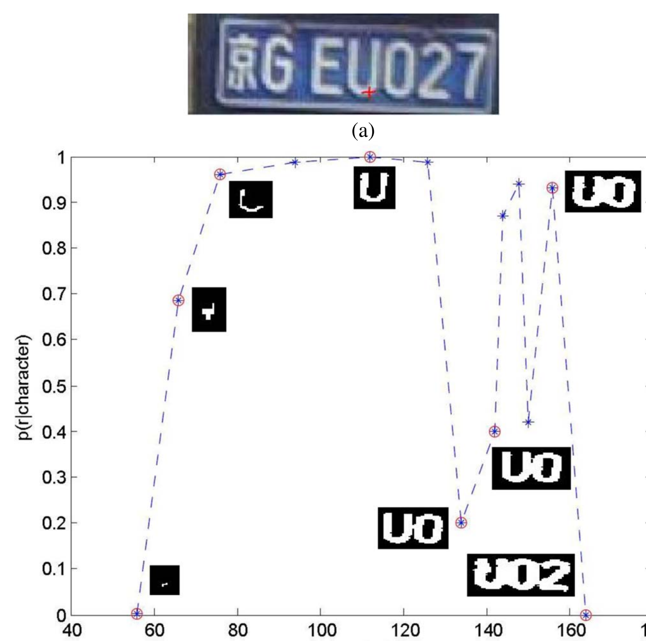
Fig. 3. Some intermediate results of coarse localization for a noisy image. (a) Top-hat. (b) Sobel. (c) Binarization. (d) Noise and bounding curves removing. (e) Outer contours after close operation. (f) Coarse location.

Some methods [5], [14], [21], [22] have been proposed to use the MSER detector to detect and segment license plate characters. However, these methods are not robust against low image contrast and low illumination in complex traffic environments. In light of that, Matas and Zimmermann [23] proposed to drop the stability requirement of MSERs and select class-specific ERs for object detection and categorization. Neumann *et al.* [19] proposed to select suitable ERs for scene text detection through a sequential classifier trained for character detection.

It is necessary to transform each color image into a single-channel image through a color space projection. Considering that each channel of the image is processed separately, the number of ERs for a megapixel image can reach the level of 10^6. Hence, a specific classifier is needed to select suitable ERs for character detection. Referring to Neumann *et al.* [19], we use the incrementally computable descriptors as features for the classifier. The incrementally computable descriptors of ERs are computed by sequentially increasing threshold t from 0 to 255, and include vectors as below:

- **Bounding box** $(x, y, \text{width}, \text{height})$: coordinate of top-right corner, width and height of the region.
- **Perimeter** p: length of the region boundary.
- **Area** a: number of pixels in the region.
- **Euler number** e: a topological feature based on changes of 2×2 pixel in the binary image which indicates the difference between the number of connected components and number of holes.
- **Horizontal crossings** h_i: a feature about transitions between pixels belonging to and not belonging to the region in a given row i of the region.

In this paper, we also compute the incrementally computable descriptors (compactness $\sqrt{a}/p$, number of holes $1 - e$, horizontal crossings feature h and aspect ratio width/height) for each region r and use the AdaBoost classifier presented in [19] to estimate the character-conditional probability $p(r|\text{character})$.



(a)

(b)

Fig. 4. Some estimated probabilities $p(r|\text{character})$ of each ER in different threshold in red channel of color space. (a) A source image with initial seed of the ER marked with a red cross. (b) The value of $p(r|\text{character})$ in red channel corresponding to ERs of different threshold.

The value of $p(r|\text{character})$ is tracked by the inclusion relation of ER across all thresholds (see Fig. 4). Then the local maximum of the character-conditional probability is selected only if the local maximum of the probability is greater than a threshold $p_{\min}$ and the difference between local maximum and local minimum is over a limit $\triangle_{\min}$.

According to the character detector aforementioned, we can select a set of ERs as character regions in each channel for each candidate license plate region. Only the ERs with local

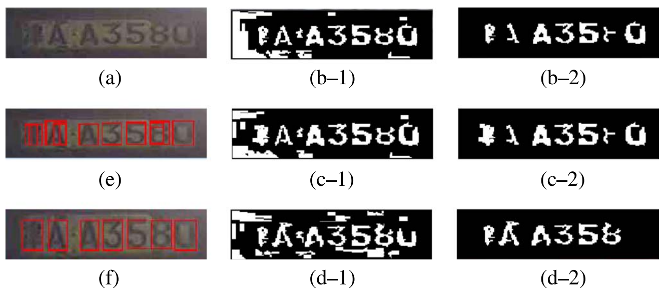


Fig. 5. Selected ERs for each channel and character-specific region segmentation. (a) Origin license plate. (b-1) Blue. (b-2) Filtration in blue. (c-1) Green. (c-2) Filtration in green. (d-1) Red. (d-2) Filtration in red. (e) All candidate rectangles in color space. (f) Final segmentation result.

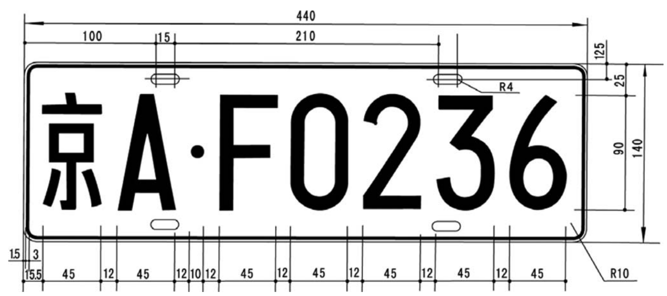


Fig. 6. Standard Chinese license plates.

maximum of the probability $p(r|\text{character})$ more than 0.75 are selected (see Fig. 5). Another filter method based on the standard license plate character bounding box's aspect ratio and area is applied to get the accurate character regions (see Fig. 5(b-2), (c-2), (d-2)). In our study, we set the threshold value of height-width ratio to 1–10, maximal width 25, and threshold value of area to 80–1100. Additionally, $p_{\min}$ is set to 0.5 and $\triangle_{\min}$ is set to 0.2.

B. Fine Detection of License Plates

It is easy to locate the LP if all the characters are accurately segmented. The detection of character regions (selected ERs) in each channel are combined after the steps aforementioned. As a result, a series of outer rectangles of character regions are obtained as shown in Fig. 5(e). The filter operation is carried on a pair of rectangles which are chosen from the set of outer rectangles. The one rectangle with height more similar to the median height is chosen as a character region. Until there are not any rectangles being overlapped, a sequence of candidate character regions are segmented simultaneously.

Considering the geometrical relationship among characters in standard Chinese license plates (see Figs. 6 and 8), the number of adjacent ERs (corresponding to characters, such as Fig. 8 $C1$ and $C2$ are called to be adjacent) allows a maximum of 7. In light of that, some geometrical parameters can be used to remove the disturbing regions and infer the locations of missed characters. A candidate license plate region is identified only if the number of characters detected in this region is between 5 and 10. In this paper, a simple but efficient location method is proposed based on character regions. Since some license plates are tilted, some thresholds are set as (5) to verify and infer the accurate locations of all 7 characters (see Fig. 5)

$$\begin{aligned} 0.8 \times D_{\text{ad}} &< \text{Dist\_1} \leq 1.3 \times D_{\text{ad}} \\ 1.3 \times D_{\text{ad}} &< \text{Dist\_2} \leq 1.8 \times D_{\text{ad}} \\ 1.8 \times D_{\text{ad}} &< \text{Dist\_3} \leq 2.8 \times D_{\text{ad}} \\ 2.8 \times D_{\text{ad}} &< \text{Dist\_4} \end{aligned} \tag{5}$$

where D_{ad} represents the average distance of all detected adjacent characters. As shown in Fig. 8, Dist_1 denotes that the adjacent rectangles are recognized as two adjacent characters, Dist_2 denotes the space mark of LP and $C2$ and $C3$ can be localized, Dist_3 means there is a missing character, and Dist_4 means there are two missing characters. As there are at least 5 detected characters, the missing characters can be inferred. As a result, as shown in Fig. 7, coarse-to-fine license plate localization is achieved, and meanwhile the character regions are segmented accurately.

IV. Recognition of License Plate Characters

After the accurate LP location and character segmentation, the character recognition process is conducted. As aforementioned in Section I, it is difficult to recognize license plate character because different camera zooms lead to different character sizes. Moreover, the Chinese characters with many strokes are more difficult to recognize especially in complex backgrounds. Here an effective recognition method based on hybrid discriminative restricted Boltzmann machines (HDRBM) is proposed.

A. Hybrid Discriminative Restricted Boltzmann Machines

A RBM is a Boltzmann machine that has a bipartite connectivity graph between hidden units h-vector and visible units x-vector. Units in each layer have no connections between them and are connected to all other units in other layers. As a result, the information flows in both directions and the weights W-vector are the same in both directions. The restricted Boltzmann machines (RBMs) are probabilistic models that use a layer of hidden variables to model a distribution over visible variables [24]. The RBMs have been developed for many learning problems and also used as feature extractors. Moreover, RBMs can be used as non-linear classifiers and can achieve better performance comparing with conventional neural networks and SVMs [25].

Hinton *et al.* [26] propose to model the joint distribution of an input $\mathbf{x} = (x_1, \ldots, x_d)$ and target class $y \in 1, \ldots, C$ using a hidden layer of binary stochastic units $\mathbf{h} = (h_1, \ldots, h_N)$. Larochelle *et al.* propose to add the class label y to the visible data to learn RBM using a discriminative approach. This model is illustrated in Fig. 9. The energy function is formulated as (6):

$$E(y, \mathbf{x}, \mathbf{h}) = -\mathbf{h}^{\mathbf{T}}\mathbf{W}\mathbf{x} - \mathbf{b}^{\mathbf{T}}\mathbf{x} - \mathbf{c}^{\mathbf{T}}\mathbf{h} - \mathbf{d}^{\mathbf{T}}\mathbf{e}_y - \mathbf{h}^{\mathbf{T}}\mathbf{U}\mathbf{e}_y \tag{6}$$

where $\mathbf{e}_y = (1_{k=y})_{k=1}^{C}$ represents one out of C for C classes, $\mathbf{U}, \mathbf{W}$ represent weights and vectors $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ represent

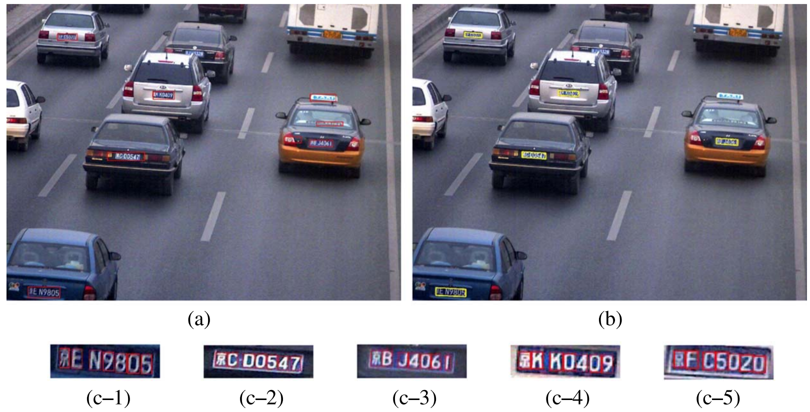


Fig. 7. Coarse-to-fine license plate detection based on character regions. (a) Coarse detection. (b) Fine detection. (c) Characters segmentation.

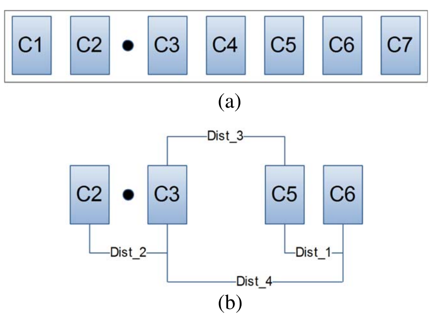


Fig. 8. Different distances for accurate segmentation. (a) Standard characters layout. (b) Example of four possible different distances of detected adjacent characters, Dist_1-4.

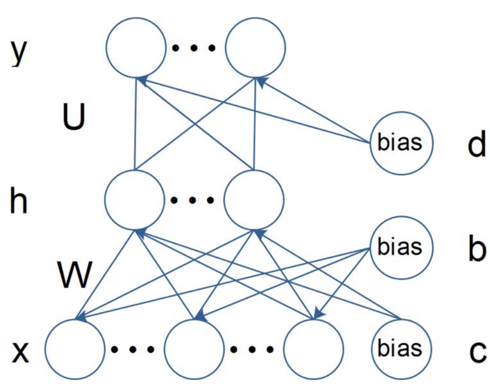


Fig. 9. RBM modeling the joint distribution of inputs and target classes.

the bias corresponding to input units, hidden units and target class units, respectively. The model parameters are denoted by $\theta = (\mathbf{W}, \mathbf{b}, \mathbf{c}, \mathbf{d}, \mathbf{U})$. From (6), the distribution probabilities to values of y, $\mathbf{x}$ and $\mathbf{h}$ is expressed as (7)

$$p(y, \mathbf{x}, \mathbf{h}) = \frac{e^{-E(y,\mathbf{x},\mathbf{h})}}{Z}. \tag{7}$$

In (7), Z is a partition function which is a normalization constant. For simple description, binary inputs of $\mathbf{x}$ are carried on in this paper and it is easily to be generalized to real-valued inputs [26].

Although $p(y, \mathbf{x}, \mathbf{h})$ is typically intractable, it is also possible to be estimated using Gibbs sampling, that is, alternating between sampling a value for the hidden layer given the current value of the visible layer [27]. Binary input variables are assumed for simplicity, but the model can be easily generalized to non-binary categories, integer-valued, and continuous-valued inputs. The conditional distributions are calculated as below:

$$p(\mathbf{x}|\mathbf{h}) = \prod_i p(x_i|\mathbf{h})$$

$$p(x_i = 1|\mathbf{h}) = \mathrm{sigm}\left(\sum_j W_{ji}h_j + b_i\right) \tag{8}$$

$$p(\mathbf{h}|y, \mathbf{x}) = \prod_j p(h_j|y, \mathbf{x})$$

$$p(h_j = 1|y, \mathbf{x}) = \mathrm{sigm}\left(\sum_i W_{ji}x_i + c_j + U_{jy}\right) \tag{9}$$

$$p(y|\mathbf{h}) = \frac{\exp\left(\sum_j U_{jy}h_j + d_y\right)}{\sum_{y^*}\exp\left(\sum_j U_{jy^*}h_j + d_{y^*}\right)}. \tag{10}$$

As described in [28], conditional distribution $p(y|\mathbf{x})$ can be computed as (11). Hence, the classification task can be achieved. One way to interpret (11) is that, when assigning probabilities to a particular class y (y^* denotes all classes) for some inputs $\mathbf{x}$, the probability $p(y|\mathbf{x})$ indicates how well the input fits with different filters denoted by rows $\mathbf{W}_j$ of $\mathbf{W}$. Different classes share the same filters (weights), and biases U_{jy} make it possible to distinguish different classes corresponding to different filters

$$p(y|\mathbf{x}) = \frac{\exp(d_y)\prod_j\left(\exp\left(\sum_i W_{ji}x_i + U_{jy} + c_j\right) + 1\right)}{\sum_{y^*}\exp(d_{y^*})\prod_j\left(\exp\left(\sum_i W_{ji}x_i + U_{jy^*} + c_j\right) + 1\right)}. \tag{11}$$

Given N training samples where $(y_k, \mathbf{x}_k)$ denotes the k-th sample. A widely used generative objective function for learning RBM is shown as (12) which concentrates on modeling the input only. While for the LP character recognition task, it is

Fig. 10. Some examples of training characters.

more appropriate to focus on supervised part of discriminative training as expressed in (13)

$$O_{\text{gen}} = -\sum_{k=1}^{N} \log p(y_k, \mathbf{x}_k) \tag{12}$$

$$O_{\text{disc}} = -\sum_{k=1}^{N} \log p(y_k|\mathbf{x}_k). \tag{13}$$

Referring to [27], adding the generative training objective to the discriminative training objective is a way to regularize the discriminative training objective, and as a result, the hybrid restricted Boltzmann machine (HDRBM) based on objective function is shown as below:

$$O_{\text{hybrid}} = O_{\text{disc}} + \lambda O_{\text{gen}} \tag{14}$$

where the weight λ can be optimized based on the training data.

B. Characters Recognition

For the LP character recognition, we use the probabilistic models (HDRBM) aforementioned to classify the characters. Since different character regions have different sizes, we rescale each character region into a fixed size of 20 $\times$ 40. In this paper, histogram of oriented gradients (HOG) [29] features are extracted to represent the character regions. HOG describes distributions of edges at different parts of an image, and is very effective in representing the license plate characters. In addition, we set the cell size, block size, block stride, and number of orientations to 10 $\times$ 10, 20 $\times$ 20, 5 $\times$ 5 and 9, respectively. As a result, a feature vector of 180 dimensions for a character region is obtained.

The standard Chinese license plate has seven characters, among which the first is a Chinese character and six are digits or alphabets. Moreover, there are 31 Chinese characters, 10 digits and 24 alphabets (excluding letters I and O). Different from [16] in classifying 65 classes in total, to achieve more accurate recognition, two classifiers based on HDRBM are trained for 31 Chinese characters and 34 letters respectively. Hence, 12366 labeled Chinese characters and 10408 labeled letters are collected (some training samples are cropped from real traffic scene with different light conditions and the others are artificially generated with rotation or adding noise). Some examples of training characters are shown in Fig. 10. For

TABLE I
COMPARISON RESULTS ON TESTING CHARACTERS

	SVM	SaE-ELM	HDRBM
Chinese character	97.7%	98.0%	98.5%
Letter	98.8%	98.6%	99.1%

	Ground truth	HDRBM	SVM	SaE-ELM
	京	京	皖	苏
	晋	晋	吉	鲁
	冀	冀	藏	青
	V	V	V	N
	1	1	J	J
	7	7	T	T

Fig. 11. Some recognition results of different classification methods.

the training step on the basis of objective function (14), the gradients of (12) and (13) with respect to parameters θ of a training example $(y_k, \mathbf{x}_k)$ are calculated as below:

$$\frac{\partial \log p(y_k, \mathbf{x}_k)}{\partial \theta} = -\left\langle \frac{\partial E(y_k, \mathbf{x}_k, \mathbf{h})}{\partial \theta} \right\rangle_{p(\mathbf{h}|y_k, \mathbf{x}_k)} + \left\langle \frac{\partial E(y, \mathbf{x}, \mathbf{h})}{\partial \theta} \right\rangle_{p(y, \mathbf{x}, \mathbf{h})}$$

$$\frac{\partial \log p(y_k|\mathbf{x}_k)}{\partial \theta} = -\left\langle \frac{\partial E(y_k, \mathbf{x}_k, \mathbf{h})}{\partial \theta} \right\rangle_{p(\mathbf{h}|y_k, \mathbf{x}_k)} + \left\langle \frac{\partial E(y, \mathbf{x}, \mathbf{h})}{\partial \theta} \right\rangle_{p(y, \mathbf{h}|\mathbf{x})}$$

where $\langle \cdot \rangle_p$ denotes the mathematical expectation corresponding to the distribution p. Indeed, as mentioned previously, the joint probability $p(y_k, \mathbf{x}_k)$ in the second expectation is generally intractable. In this paper, contrastive divergence estimator [30] (also called CD-k, and k is set to 1 in our experiment) is used to estimate the joint probability.

Similar to the training characters (see Fig. 10), manually cropped 1612 images with Chinese characters and 1609 images with letters are used as a testing set for HDRBM. Since there are 31 classes of Chinese characters and 34 classes of letters, each class contains from 16 to 70 test images. To evaluate the performance of HDRBM, we also trained SVM and self-adaptive evolutionary extreme learning machine (SaE-ELM) [16], [31] using the same training set. In the experiment, the number of hidden nodes is set to 150 and the weight λ of HDRBM is set to 0.01. The number of hidden nodes in SaE-ELM is set to 800. The kernel function of SVM is set to RBF, and the parameters of C and γ is set to 8 and 1, respectively. Finally, we get recognition rates of 98.5% and 99.1% for Chinese characters and letters, respectively. The comparison results with SVM and SaE-ELM are listed in Table I and some recognition results of different classification methods are listed in Fig. 11. From Table I, the proposed HDRBM outperforms other classification methods. Conventional classifiers like SVM and neural

TABLE II
IMAGE DATABASE COMPOSITIONS

Data set	Taken time	Resolution	No. of Images	No. of Plates	Plates width	Characters Height
Subset 1	Day	$1936 \times 2592, 1280 \times 736$	1390	2398	80~170	16~43
Subset 2	Day	720×280	1344	1344	90~150	20~30
Subset 3	Night	1936×2592	676	1247	80~170	16~43
Subset 4	Night	720×280	832	832	90~150	20~30

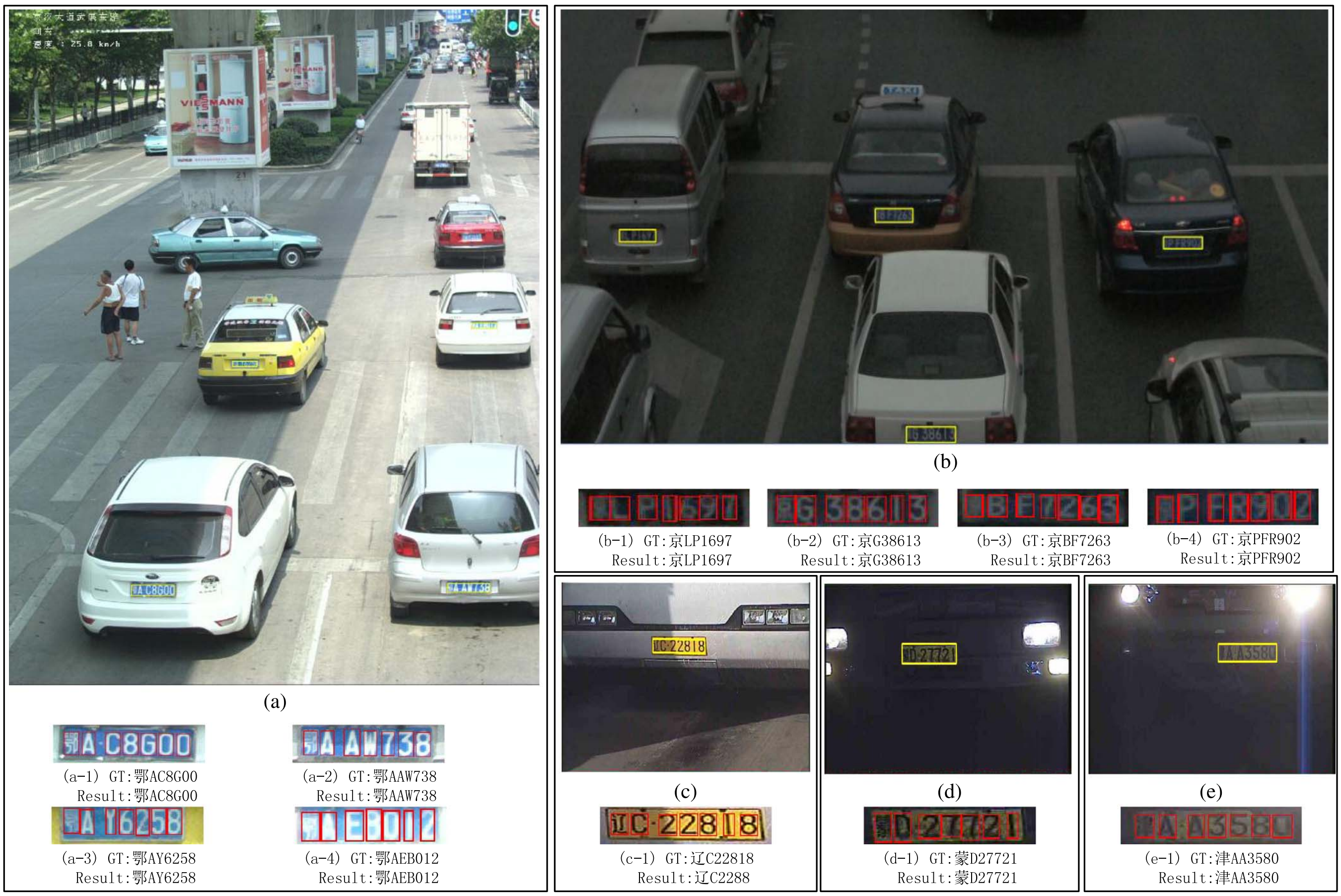


Fig. 12. Some detection and recognition results of the data set. For the LP detection, yellow bounding boxes are the detected locations of LPs. For the recognition, red bounding boxes are the detected characters. GT is short for ground truth. (a) From subset 1. (b) From subset 3. (c) From subset 2. (d) From subset 4. (e) From subset 4.

networks are trained through discriminative training objective. It is appropriate to focus on supervised part of discriminative training. But for the LPR task, to enlarge the training set, some samples can be artificially generated with rotation or adding noise. As a result, adding the generative training objective to discriminative training objective is an effective way to regularize the training objective for worse plate images conditions. In light of that, HDRBM is robust enough to the license plate recognition task. More experimental results are conducted and described in next section.

V. EXPERIMENTAL RESULTS

A. Data Set and Evaluation Criteria

To evaluate the effectiveness of the proposed method, we extend the conference version data set [16] to 4242 images which are taken from a wide variety of real traffic monitoring scenes under various illumination conditions. All the test images are separated into four subsets as listed in Table II. This data set is very challenging due to various illumination conditions and complex backgrounds. Some sample images are shown in Figs. 12 and 13. The ground truth of license plate positions and contents are manually labeled. In our experiment, the characters with height above 16 pixels are detected and recognized. In addition, in this paper, the significant license plates regions have heights from 16 to 43 and widths more than 80 pixels.

The performance of our proposed method is evaluated by calculating separately the percentage of correctly detected license plates, correctly recognized characters and overall performance. A license plate is correctly detected only if the overlap of the detected and ground truth bounding box is above 0.8. Hence, in this paper, LDR (license plate detection rate),

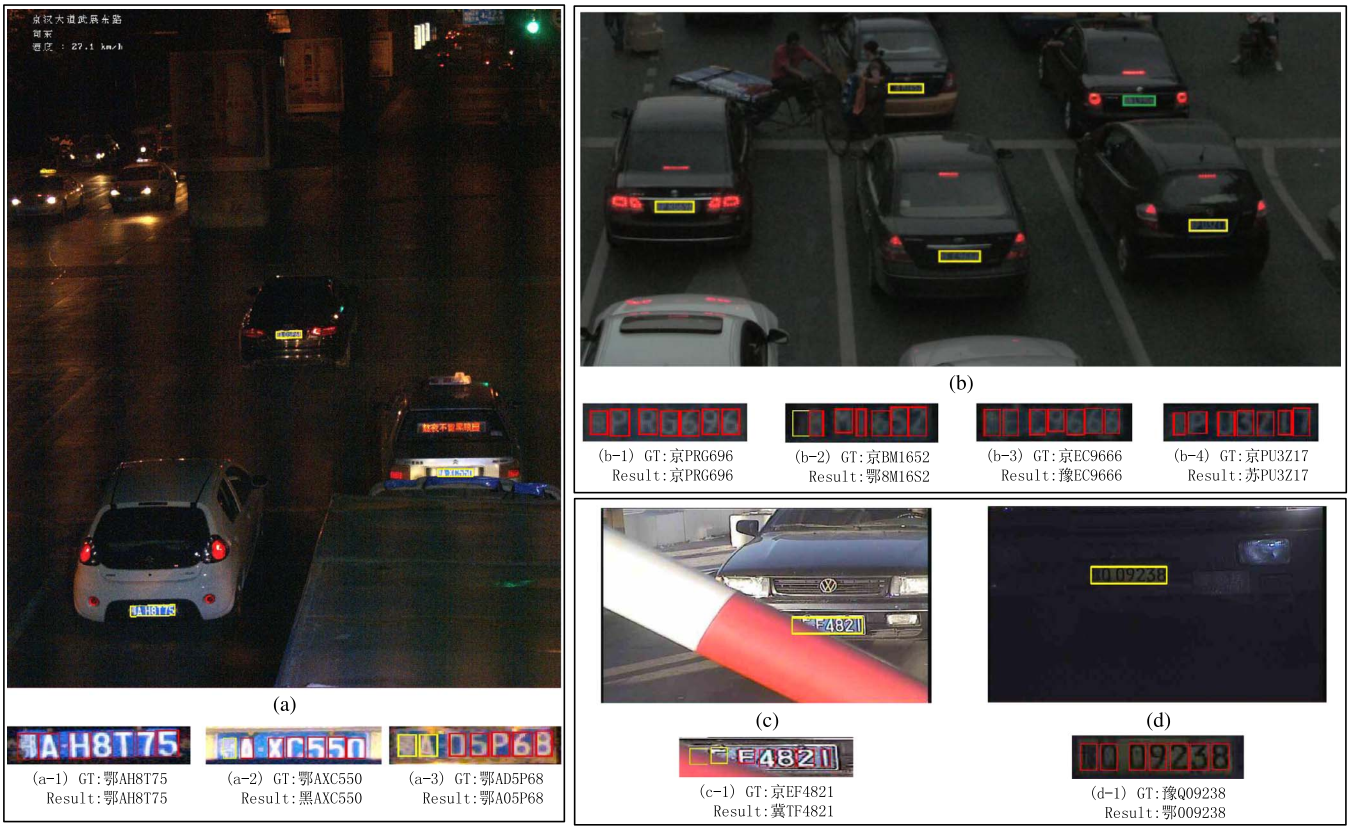


Fig. 13. Some wrong recognition results of the data set. For the LP detection, yellow bounding boxes are the detected locations of LPs, green bounding boxes are the undetected locations of LPs. For the recognition, red bounding boxes are the detected characters and yellow bounding boxes are the inferred locations of character regions. GT is short for ground truth. (a) From subset 3. (b) From subset 3. (c) From subset 2. (d) From subset 4.

CRR (character recognition rate), OVR1 and OVR2 (overall performance) are defined as below:

$$\mathrm{LDR} = \frac{\text{Number of Correctly Detected LPs}}{\text{Number of All Ground Truth LPs}}$$

$$\mathrm{CRR} = \frac{\text{Number of Correctly Recognized Characters}}{\text{Number of All Detected Characters}}$$

$$\mathrm{OVR1} = \frac{\text{Number of Correctly Detected and Recognized LPs}}{\text{Number of All Ground Truth LPs}}$$

$$\mathrm{OVR2} = \mathrm{LDR} \times \mathrm{CRR}$$

where OVR1 and OVR2 are two different criteria for evaluating the overall performance.

B. Experimental Results

In this paper, we consider blue (B), green (G) and red (R) channel in color space and their combination of channels to detect character-specific ERs. According to the experimental results (see Table III), the combination of G channel and R channel is adopted in day-time test images and B, G and R channel in night-time images. With more character-specific ERs extracted in subset 1, the real characters are more likely to be eliminated by the coarse-to-fine location methods described in Section III-B. In light of that, the combination of R, G and B channel in subset 1 taken in day-time results in not so good performance. Actually, after the coarse location step, a rough verification method based on local binarization is applied to eliminate most of the false positive license plates. It is efficient to decrease the processing time of the whole system because the extraction of ERs is time-consuming.

TABLE III
LICENSE PLATE DETECTION RESULTS FOR SUBSET 1 AND SUBSET 3

Channel	Subset 1		Subset 3	
	Recall	Precision	Recall	precision
B	38.1%	97.6%	28.3%	89.6%
G	87.5%	95.8%	65.7%	97.5%
R	90.2%	99.4%	87.3%	97.8%
G&R	**94.8%**	**98.8%**	92.5%	96.7%
B&G&R	93.5%	91.6%	**93.6%**	**98.9%**

TABLE IV
EXPERIMENTAL RESULTS ON OUR DATA SET

Data set	LDR	CRR	OVR1	OVR2
Subset 1	94.8% (2273/2398)	98.2% (15624/15911)	90.2% (2163/2398)	93.1%
Subset 2	98.2% (1320/1344)	98.5% (9101/9240)	94.3% (1267/1344)	96.7%
Subset 3	93.6% (1167/1247)	97.7% (7981/8169)	89.2% (1112/1247)	91.4%
Subset 4	96.9% (806/832)	98.3% (5546/5642)	93.8% (780/832)	95.3%
Average	95.9%	98.2%	91.9%	94.1%

TABLE V
PERFORMANCE COMPARISON WITH DIFFERENT METHODS

Data set	Proposed method			MSER based[14]			Grey quantization [10]			Text-line based [3]		
	LDR	CRR	OVR2	LDR	CRR	OVR2	LDR	CRR	OVR2	LDR	CRR	OVR2
Subset 1	94.8%	98.2%	93.1%	89.4%	94.1%	84.1%	92.4%	96.3%	89.0%	–	–	–
Subset 2	98.2%	98.5%	96.7%	95.8%	94.8%	90.8%	97.8%	97.8%	95.5%	95.4%	96.1%	91.7%
Subset 3	93.6%	97.7%	91.4%	87.7%	92.5%	81.1%	92.3%	96.1%	88.7%	–	–	–
Subset 4	96.9%	98.3%	95.3%	94.6%	95.3%	90.2%	96.6%	97.3%	94.0%	94.7%	95.8%	90.7%
Average	95.9%	98.2%	94.1%	91.9%	94.2%	86.6%	94.7%	96.9%	91.8%	95.1%	96.0%	91.2%

The experimental results of the proposed method for each data set are listed in Table IV. From this table, the proposed method achieves a promising performance in various conditions and can averagely get LDR, CRR, OVR1 and OVR2 of 95.9%, 98.2%, 91.9% and 94.1%, respectively. For the coarse location step, Gaussian blur in 5×5 grid is applied in test images with resolution of 1936×2592. Actually, the average recall rate of LDR is more than 97% after coarse location step. According to (11) in our experiment, almost all the probabilities of the correctly detected character regions are more than 0.98. Hence, the false positive license plates with average probabilities lower than 0.9 are eliminated. As a result, little false positive is detected in the experiment. The LDR listed in Table IV is final fine license plates detection recall rate. Actually, the proposed method results in an average license plate detection precision of 99.2% under the recall rate listed in Table IV. Some license plates detection, segmentation and recognition results are shown in Fig. 12. It is observed that the proposed method can detect multi-plates in each image and is robust enough to deal with various light conditions. In particular, for test images in subset 3, the detection and recognition task is quite more challenging due to the waning light and complex backgrounds. The proposed method can get OVR2 of 91.4%.

It is noted that 20 pixels is typically the lowest height of identifiable characters in most characters recognition methods [32]. While in our experiment, to detect more license plates, characters with height above 16 is detected and recognized. Actually, for the recognition step, the low height of characters is more likely to result in some recognition error especially for Chinese characters as shown in Figs. 11 and 13 (see Fig. 13, (a-2) and (b) characters with minimal height 16). The major reason is that the detected character regions are too small to be recognized especially the Chinese characters regions with many strokes. Another reason is that for the character detection step, there are some inferred locations of some character regions (see Fig. 13, extremal regions of Chinese characters with complex strokes are not detected as characters in (a-2) and (a-3), and the first character is occluded in (b-2) (c-1)).

C. Performance Comparison

Two existing Chinese license plate detection methods are implemented to compare with our proposed method (see Table VI). The edge-based [4] location method needs enhancing the local areas to intensify the texture of the plate region. It is not robust to wide-view images in subset 1 and subset 3. The component-based [5] method detected the character regions as MSERs,

TABLE VI
LDR COMPARISON WITH DIFFERENT METHODS

Data set	Our method	Edge-based [4]	Component-based [5]
Subset 1	94.8%	90.1%	93.5%
Subset 2	98.2%	96.3%	97.7%
Subset 3	93.6%	89.7%	92.4%
Subset 4	96.9%	95.8%	95.3%
Average	95.9%	93.0%	94.7%

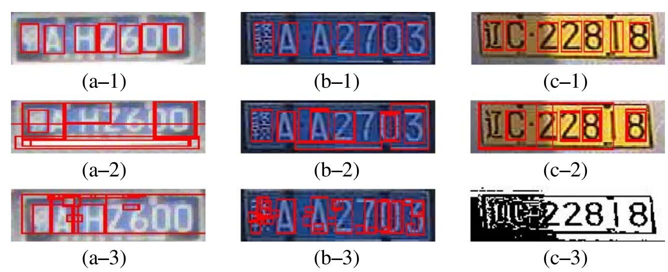


Fig. 14. Some detection results of character regions by other methods. (a-1), (b-1), (c-1) Final segmentation results of the proposed method. (a-2), (b-2), (c-2) MSER based method. (a-3), (b-3) Grey quantization. (c-3) Local adaptive binarization of text-line based.

then the conditional random field models were constructed on the candidate characters (MSER) in neighborhoods. This method's performance highly depends on the MSER detection. But the MSER is not so effective in license plates with low-contrast in gradient. Our method extract ERs in color space and MSER is a special case of ERs. Hence, our method is more robust to all-day images and performs better as listed in Table VI.

Performance comparisons between the proposed system and other Chinese LPR systems are given in Table V. Three existing detection methods are implemented. The approach in [14] is proposed to detect the character regions in grey level using the MSER descriptors, which actually is a special case of ERs. This MSER based method can hardly obtain the exact bounding boxes of characters in low resolution while the proposed method can (see Fig. 14, the first row is the results of the proposed method and the second row is the results of MSER based [14]). In addition, the recognition in [14] is achieved by matching templates represented by 180 dimensions HOGs. It is not robust because the Euclidean distance can be affected a lot in high dimensions. Jiao *et al.* [10] propose a configurable LPR method for multi-style license plate. Plate candidates (called ROI, region of interest) are also detected by

TABLE VII
OUR PROPOSED METHOD PROCESSING TIME OF IMAGES WITH DIFFERENT RESOLUTIONS

Processing steps	Cost time(s) 1936 × 2592	Cost time(s) 1280 × 736	Cost time(s) 720 × 280
Coarse LP detection	0.97	0.213	0.077
Characters extraction	0.39	0.215	0.187
Fine location	0.031	0.024	0.018
Characters recognition	0.015	0.016	0.016
Total cost	1.406	0.468	0.298

dense vertical edges. Local binarization method is used. The size of "close" operator is set as $(1, 0.2 \times \text{Height of ROI})$ pixels in X and Y orientations and the quantization parameter Q is adjusted from 5 to 10 to find all possible character regions. In our implementation, to achieve better performance in the high resolution images, the "close" operator is set as (1, 3). A three-layer ANN (artificial neural network) is trained to choose detected blocks and recognize characters. Many false positives in [10] are detected which affect the final method performance. Moreover, the quantization based method cannot detect character regions in low contrast as shown in Fig. 14(a-3) and (b-3). Obviously, this method can segment a character into two in low resolution. In [3], for the license plate location, local adaptive binarization method is applied to get the plate candidates. For the subset 1 and subset 3 with high resolution test images, the comparison is not carried on because this local binarization (Otsu) location method can hardly locate the license plates as shown in Fig. 14(c-3). After coarse license plate location, text-line is constructed to verify the location of plates based on the geometrical relationship in standard Chinese plates. Vertical projection is applied to segment the characters which is not robust enough to achieve accurate segmentation.

The experimental results in Table V suggest that the proposed LPR system outperforms other existing detection methods. In addition, the proposed method can achieve effective performance and is robust to deal with various light conditions with complex backgrounds.

D. Computation Cost

All the experiments are performed on a standard PC with 3.1-GHz Intel Core 2 Quad CPU and 4-GB RAM. The whole system is implemented using the Microsoft Visual C++ compiler. The computational cost highly depends on the test image size. Taking images with one plate in the data set as examples, the average computation time of each step (see Fig. 1) for daytime images has been assessed and the results are presented in Table VII. From Table VII, most time are spent on the first two steps. For high resolution images with complex backgrounds, most time is spent on coarse LP location because more false positives are likely to be detected. In addition, it costs about 0.4 seconds to detect and select the character-specific ERs as character regions. Actually, in real applications, regions of interest (ROI) can be set in high resolution images. In our experiment, it costs only 0.47 seconds to detect and recognize 2 plates in an image of 1280 × 736. As a result, our proposed method satisfies the demands of the practical application.

VI. CONCLUSION

An effective approach for vehicle license plate detection and recognition, based on character-specific regions, is proposed in this paper. Firstly, a sequence of morphological operations is applied to find plate candidates with dense vertical edges. Then the character-specific ERs are extracted and selected as character regions in color space. The recognition step is achieved by an effective classifier named HDRBM. As discussed in Table IV, the proposed method achieves better average performance with LDR $= 95.9\%$, CRR $= 98.2\%$, OVR1 $= 91.9\%$ and OVR2 $= 94.1\%$.

To our knowledge, HDRBM is the first to be integrated for license plate recognition. In addition, there is little literature dedicated to both license plate detection and recognition in wide view with high resolution images like subset 1 and subset 3. The experimental results show that the proposed method can achieve outstanding performance in all-day surveillance environment with various illumination conditions and complex backgrounds. The approach can be easily generalized to license plates from other countries as long as the character layouts are given.

However, our proposed method still has restrictions. For the recognition step, the recognition rate highly depends on accurate extractions of character regions. Some extracted or inferred character regions with low probabilities can be relocated. Actually, we have tried to extract features on raw pixel data by using deep architectures but have not got ideal results. A large scale of training data set for recognition task will be collected and the proposed framework will be generalized to license plates of other countries. In addition, deep architectures for location and recognition will be our future work.

ACKNOWLEDGMENT

Kunfeng Wang is the corresponding author. The authors would like to thank Professor Fei-Yue Wang at the Institute of Automation, Chinese Academy of Sciences for his supports and instructions to this work.

REFERENCES

[1] B. Tian *et al.*, "Hierarchical and networked vehicle surveillance in its: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 557–580, Apr. 2015.
[2] X. Shi, W. Zhao, and Y. Shen, *Automatic License Plate Recognition System Based on Color Image Processing*. Berlin, Germany: Springer-Verlag, 2005, pp. 1159–1168.
[3] B. Shan, "Vehicle license plate recognition based on text-line construction and multilevel RBF neural network," *J. Comput.*, vol. 6, no. 2, pp. 246–253, Feb. 2011.
[4] D. Zheng, Y. Zhao, and J. Wang, "An efficient method of license plate location," *Pattern Recog. Lett.*, vol. 26, no. 15, pp. 2431–2438, Nov. 2005.
[5] B. Li, B. Tian, Y. Li, and D. Wen, "Component-based license plate detection using conditional random field model," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 4, pp. 1690–1699, Dec. 2013.
[6] K. Deb, H.-U. Chae, and K.-H. Jo, "Vehicle license plate detection method based on sliding concentric windows and histogram," *J. Comput.*, vol. 4, pp. 771–7, Aug. 2009.
[7] J. Dun, S. Zhang, X. Ye, and Y. Zhang, "Chinese license plate localization in multi-lane with complex background based on concomitant colors," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 3, pp. 51–61, Fall 2015.

[8] K. Deb, V. V. Gubarev, and K.-H. Jo, *Vehicle License Plate Detection Algorithm Based on Color Space and Geometrical Properties*. Berlin, Germany: Springer-Verlag, 2009, pp. 555–564.

[9] A. Ashtari, M. Nordin, and M. Fathy, "An Iranian license plate recognition system based on color features," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1690–1705, Aug. 2014.

[10] J. Jiao, Q. Ye, and Q. Huang, "A configurable method for multi-style license plate recognition," *Pattern Recog.*, vol. 42, no. 3, pp. 358–369, Mar. 2009.

[11] I. Paliy, V. Turchenko, V. Koval, A. Sachenko, and G. Markowsky, "Approach to recognition of license plate numbers using neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, 2004, vol. 4, pp. 2965–2970.

[12] C.-N. Anagnostopoulos, I. Anagnostopoulos, I. Psoroulas, V. Loumos, and E. Kayafas, "License plate recognition from still images and video sequences: A survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 3, pp. 377–391, Sep. 2008.

[13] Y. Wen *et al.*, "An algorithm for license plate recognition applied to intelligent transportation system," *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 3, pp. 830–845, Sep. 2011.

[14] 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 9th ICNSC*, 2012, pp. 399–404.

[15] Y.-P. Huang, S.-Y. Lai, and W.-P. Chuang, "A template-based model for license plate recognition," in *Proc. IEEE Int. Conf. Netw., Sens. Control*, 2004, vol. 2, pp. 737–742.

[16] C. Gou, K. Wang, B. Li, and F.-Y. Wang, "Vehicle license plate recognition based on class-specific ERs and SaE-ELM," in *Proc. IEEE 17th ITSC*, Oct. 2014, pp. 2956–2961.

[17] J.-W. Hsieh, S.-H. Yu, and Y.-S. Chen, "Morphology-based license plate detection from complex scenes," in *Proc. IEEE 16th Int. Conf. Pattern Recog.*, 2002, vol. 3, pp. 176–179.

[18] D. L. Baggio, *Mastering OpenCV with Practical Computer Vision Projects*. Birmingham, U.K.: Packt Publishing, 2012.

[19] L. Neumann and J. Matas, "Real-time scene text localization and recognition," in *Proc. IEEE CVPR*, 2012, pp. 3538–3545.

[20] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image Vis. Comput.*, vol. 22, no. 10, pp. 761–767, Sep. 2004.

[21] W. Wang, Q. Jiang, X. Zhou, and W. Wan, "Car license plate detection based on MSER," in *Proc. IEEE Int. Conf. CECNet*, 2011, pp. 3973–3976.

[22] Q. Gu, J. Yang, L. Kong, and G. Cui, "Multi-scaled license plate detection based on the label-moveable maximal MSER clique," *Opt. Rev.*, vol. 22, no. 4, pp. 1–10, Aug. 2015.

[23] J. Matas and K. Zimmermann, *A New Class of Learnable Detectors for Categorisation*. Berlin, Germany: Springer-Verlag, 2005, pp. 541–550.

[24] P. Smolensky, "Parallel distributed processing: Explorations in the microstructure of cognition," in *Information Processing in Dynamical Systems: Foundations of Harmony Theory*, vol. 1, D. E. Rumelhart, J. L. McClelland, and P. R. Group, Eds. Cambridge, MA, USA: MIT Press, 1986, pp. 194–281.

[25] H. Larochelle and Y. Bengio, "Classification using discriminative restricted Boltzmann machines," in *Proc. 25th Int. Conf. Mach. Learn.*, 2008, pp. 536–543.

[26] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, May 2006.

[27] H. Larochelle, M. Mandel, R. Pascanu, and Y. Bengio, "Learning algorithms for the classification restricted Boltzmann machine," *J. Mach. Learn. Res.*, vol. 13, no. 1, pp. 643–669, 2012.

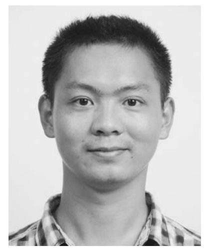
[28] R. Salakhutdinov, A. Mnih, and G. Hinton, "Restricted Boltzmann machines for collaborative filtering," in *Proc. 24th Int. Conf. Mach. Learn.*, 2007, pp. 791–798.

[29] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. CVPR*, 2005, vol. 1, pp. 886–893.

[30] G. E. Hinton, "Training products of experts by minimizing contrastive divergence," *Neural Comput.*, vol. 14, no. 8, pp. 1771–1800, 2002.

[31] J. Cao, Z. Lin, and G.-B. Huang, "Self-adaptive evolutionary extreme learning machine," *Neural Process. Lett.*, vol. 36, no. 3, pp. 285–305, Dec. 2012.

[32] C. Anagnostopoulos, "License plate recognition: A brief tutorial," *IEEE Intell. Transp. Syst. Mag.*, vol. 6, no. 1, pp. 59–67, Spring 2014.



Chao Gou received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012. He is currently working toward the Ph.D. degree in control theory and control engineering with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China.

In addition, he is currently an Intern at Qingdao Academy of Intelligent Industries, Qingdao, China. His research interests include intelligent transportation systems, pattern recognition, and image processing.

Kunfeng Wang received the bachelor's degree in materials science and engineering from Beihang University, Beijing, China, in 2003 and the Ph.D. degree in control theory and control engineering from the University of Chinese Academy of Sciences, Beijing, in 2008.

Since 2008, he has been an Assistant Professor with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences. His research interests include intelligent video analytics, intelligent transportation systems, and machine learning.

Yanjie Yao received the B.S. degree from China University of Geosciences, Beijing, China, in 2011. She is currently working toward the Ph.D. degree in control theory and control engineering with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing.

Currently, she is also with the Research Center for Computational Experiments and Parallel Systems, National University of Defense Technology, Changsha, China. Her research interests include intelligent transportation systems, image processing, and computer vision.

Zhengxi Li received the B.S., M.S., and Ph.D. degrees from the University of Science and Technology Beijing, Beijing, China, in 1984, 1990, and 2004, respectively, all in control science and engineering.

He is currently a Professor of electrical engineering with and the Vice President of North China University of Technology, Beijing. His research interests include motor drives and intelligent transportation.