

Visual saliency-based vehicle logo region detection

Yao Zhao^{1, a}, Fuqiang Zhou^{1, b} and XinmingWang^{1, c}

¹School of Instrumentation Science and Opto-electronics Engineering, Beihang University, Beijing 100191, China;

^ajoeasir0560@163.com, ^bzfq@buaa.edu.cn, ^cxmw19890815@163.com

Keywords: Visual saliency, intelligent transportation system, vehicle logo detection.

Abstract. Vehicle logo detection (VLD) is one of the crucial parts of intelligent transportation system (ITS). VLD methods are mostly based on learning progress or the relative position of vehicle logo and license plate. However, the learning progress is time-consuming, and the relative position above limits the application of VLD, especially when the license plate is removed. In this paper, a novel VLD method, based on the features of vehicle logo using saliency detection is proposed and solved the two problems above. Three dataset containing totally 3000 images is generated to assess the accuracy of this system. A detection rate of 92.27% is finally obtained, demonstrating the robustness and efficiency of our method.

1. Introduction

Followed by the development of modern transportation, intelligent transportation system (ITS) [1], which aims at the monitoring and feedback of the traffic flow, abnormal traffic events, and emergency circumstances, are becoming increasingly crucial to deal with the rising demand of transportation in the modern society. As a result, vehicle logo detection (VLD) is now paid increasing attention in the domain of pattern recognition.

Despite several works recently, VLD is still far from a mature domain as vehicle logos are in different shapes, textures and colors. Lu et al. [2] proposed a method based on sample matching with the information of surrounding objects near vehicle logos. While the learning process complicates the detecting system seriously. Mao et al. [3] used vehicle license plate detection (VLPD) to get the coarse localization of vehicle logo. While it doesn't work when the license plate is removed. Also, there are some methods using edge detection and projection [4], which are easily influenced by the background noisy.

Visual saliency detection, aiming to find the most salient parts in an image, is gaining intensive research attention, which becomes the fundamental of many fields in image process, such as, image segmentation, image resizing and visual recognition. Itti et al. [5] proposed the early saliency detection method, with some biological plausible features, including color, intensity and orientation, extracted on multiple scales. And many new and complex features are proposed in some novel methods.

Mao et al. used visual saliency detection in the VLD method, which shows good results. However, saliency detection contributed limited in Mao's method and the main method is still based on VLPD. Zhang et al. [6] proposed a VLD method, which also uses visual saliency to generate a coarse localization of logo region, and the method, is actually supervised, is mainly based on template matching process. In our work, a novel bottom-up visual saliency-based VLD method is proposed and presents good results. Compared with the methods of Mao and Zhang, our saliency-based VLD method concentrates more on the features of logo region, in spite of the relative position of vehicle logo and license plate or any template matching process. Fig.1 illustrates the main system configuration of our method.

The paper is organized as follows. The VLD algorithm is described in Section 2. In Section 3, the experiment results and discussion is presented. Finally, Section 4 concluded this paper.

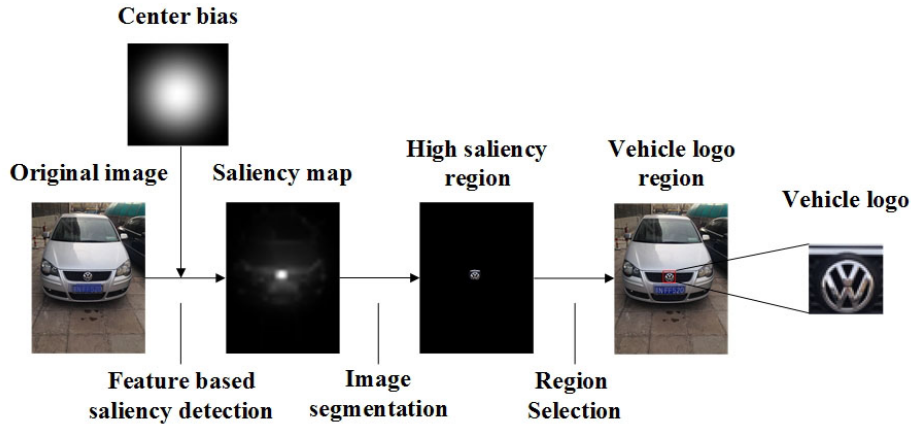


Fig. 1 Flowchart of visual saliency-based VLD.

2. Saliency-based VLD algorithm

2.1 Feature Extraction.

Considering the characteristics of vehicle logos, we can extract some benefitting features to describe the logo precisely. In our work, intensity and orientation features are chosen, which display a significant result. Vehicle logos are always in metallic colors, which have high brightness or luminance. As the input is a standard RGB image, an intensity image, defined as I , can be obtained from the separated r, g, and b channels; this image can be expressed mathematically as follows:

$$I = (r + g + b) / 3 \quad (1)$$

Fig. 2 shows the original input image and a three-dimensional intensity statistical graph of the input image, respectively. Despite other high-intensity regions in the image, the vehicle logo region presents a relatively high and isolated intensity in the local region.

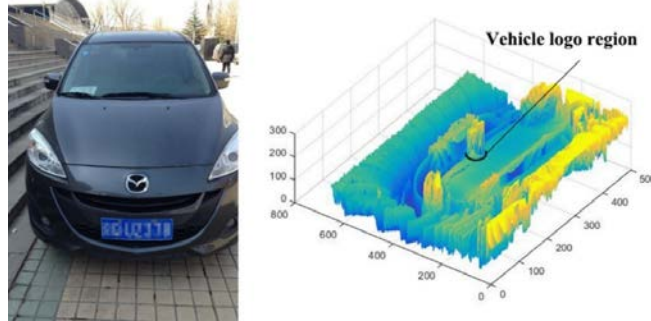


Fig. 2. Original image and the three-dimensional statistical graph of intensity.

Based on image intensity I , the local orientation information can be obtained with oriented Gabor filters of 0° , 45° , 90° , and 135° , which represent the different information in the different orientations, with the intensity image of Fig.2 (left) as the input image, and the vehicle logo regions have more noticeable results and at a higher-saliency level than any other regions in the image. To make it more intuitive, the three-dimensional statistical graphs of images with 0° , 45° , 90° and 135° Gaborfilters, shown in Fig. 3, are used.

2.2 Saliency Map Generation and Logo Detection.

In the method of Itti et al., by using a dyadic Gaussian pyramid, where $\sigma = \{0, 1, 2, \dots, 8\}$, nine spatial scale images are obtained. We firstly compute the intense of scale 2, 3, 6, and 7, which are defined as $I(3)$, $I(4)$, $I(6)$ and $I(7)$. Then, the $I(2,6)$ and $I(3,7)$, which are actually two raw feature maps, are obtained by interpolation to the finer scale and point-by-point subtraction. Based on the $I(2,6)$, and $I(3,7)$, the raw orientation maps are obtained with oriented Gabor filters of 0° , 45° , 90° , and 135° , which are $O(2,6,0^\circ)$, $O(2,6,45^\circ)$, $O(2,6,90^\circ)$, $O(2,6,135^\circ)$, $O(3,7,0^\circ)$, $O(3,7,45^\circ)$, $O(3,7,90^\circ)$ and $O(3,7,135^\circ)$. Totally, two raw intense maps and eight raw orientation maps are obtained.

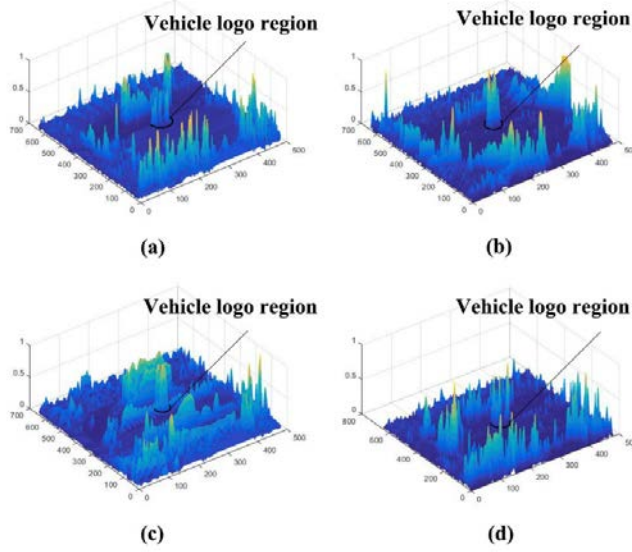


Fig. 3. Three-dimensional statistical graphs of results obtained using different Gabor filters. For node p and node q in a raw feature map F , the dissimilarity is defined as follows:

$$d(p \parallel q) = |M(p) - M(q)| \quad (2)$$

Assume that a fullyconnected directed graph G_A consists of all pixels in the feature image, which means every node in this region is connected to the others. The directed edge from node p to node q will be assigned a weight as follows:

$$\omega_A(p, q) = d(p \parallel q) \cdot F(p_x - q_x, p_y - q_y) \quad (3)$$

$$F(a, b) = \exp\left\{-\frac{a^2 + b^2}{2\sigma^2}\right\} \quad (4)$$

In (3), (p_x, p_y) and (q_x, q_y) denote the coordinates of nodes p and q , respectively, and $\omega_A(p, q)$ represents the proportionality to the dissimilarity and the closeness of p and q . By normalizing the weights of the outbound edge of each node to $[0, 1]$, we can define a Markov chain on G_A , which can reflect the saliency level of this region and generate the activation map for each raw feature map. The equilibrium of this Markov chain represents the period of time that a random walker spends at each node, and would accumulate more at nodes which have high dissimilarity with their surrounding nodes.

Further, we propose another Markov chain, similar to the one generated by the algorithm described above. A fullyconnected directed graph G_N is built on the activation map, and its directed edge can be expressed as follows:

$$\omega_N(p, q) = S(q) \cdot F(p_x - q_x, p_y - q_y) \quad (5)$$

In (5), $S(q)$ denotes the saliency value of node q . The normalized saliency map is represented by the equilibrium of this Markov chain. Mass is preferable to present at the nodes with high activation, by which could help us choose nodes with high activation and normalize the activation map.

After normalization, we combine these normalized activation maps into intense map $S(I)$ and orientation map $S(O)$ by point-by-point addition. With respect to the combination, feature maps should be put together to generate a saliency map S . Saliency map is finally obtained by a weight addition by intense map and orientation map.

$$S = 0.3S(I) + 0.7S(O) \quad (6)$$

With the saliency map generated, we incorporate a center bias to weaken the saliency value of the input image boundary and strengthen the saliency value of the center region. The center bias is just a supplementary means, which helps to reduce the system false drop and cannot be the dominating factor of saliency detection. In fact, the center bias that we use is only a simple 1-D measure of the distance of each pixel from the image center.

After the saliency map with center-bias obtained, we proposed the 99.4% most salient parts as the potential vehicle logo region as it is the most suitable for logo detection. For fear that the potential vehicle logo region is too small to be seen clearly, we capture it from the whole image and show it in Fig. 4, and the images in actual progress always have a static size.

With a low threshold binarization, the potential vehicle logo region is transformed to some connected regions. Next, we calculate the area of each connected region and consider the largest region to be the vehicle logo region, as the logo region is always the largest region in the potential logo region. After the printing of the minimum enclosing rectangle, the vehicle logo region is finally detected.

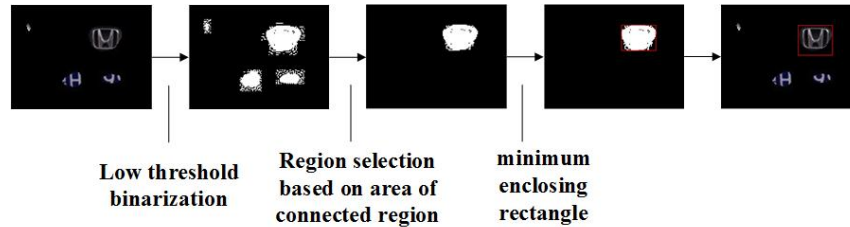


Fig. 4. Pipeline of vehicle logo region selection.

3. Experiment and discussion

In our work, the input is in the form of static RGB, and in 2448*3264 resolution. And our work is implemented by MATLAB R2014a on an Intel Core 2 Q6600 PC with 4.00GB memory. The dataset of our work consist of 3000 images, including different vehicles scales and different shooting time. Fig. 5 presents some examples of detected vehicle logo region in our experiment, which shows good effects, and our detailed result is summarized, which is analyzed in Tab.1. As presented in Fig.5, some samples with license plate removed can also be detected.



Fig. 5 Experiments result of vehicle logo region detection

There are 3000 images in our experiment, which is divided as Dataset A, B and C. Dataset A includes 1000 images, in which the vehicles are in direct sunlight. Dataset B includes 1000 images with the vehicles in poor illumination. And there are also 1000 images in Dataset C, in which the vehicles are in uniform illumination. Correct rate of Dataset A, B and C is 92.60%, 90.50% and 93.70%, respectively, and the total correct rate is 92.27%. Compared with Lu's [2], Mao's [3], Liu's [4] and Zhang's [6] methods, as is presents in Tab 2, our method doesn't need the VLPD as coarse

localization or any learning process for sample matching, which is a universal and bottom-up method with robustness and efficiency.

Table 1 Results of vehicle logo region detection

Dataset	Dataset A	Dataset B	Dataset C	Total
Number of image	1000	1000	1000	3000
Correct localization	926	905	937	2768
Failed localization	74	95	63	232
Correct rate	92.60%	90.50%	93.70%	92.27%
Failed rate	7.40%	9.50%	6.90%	7.73%

Table 2 Comparison of our methods and others'

Methods	Lu's	Mao's	Liu's	Zhang's	ours
Based on VLPD	✓	✓	✓	✗	✗
Based on learning process	✓	✗	✗	✓	✗
Detection rate	90%	92.6%	90%	94.37%	92.27%

4. Summary

A novel VLD method based on visual saliency detection is proposed in this paper. Intensity and orientation features are extracted to generate saliency map of input images, due to the high salient value that intensity and orientation of vehicle logo region presents. Computed the area of each connected region, vehicle logo region is then selected. Our method has two advantages. First, our method is based on visual saliency detection without considering about the location relationship of vehicle logo and license plate. Second, vehicle logo region can be detected directly without learning progress in our method. The result of our method is excellent in detection rate, which inspires us to focus on the following works, accurate recognition of vehicle logo for instance.

Acknowledgement

This work was supported under the National Key Scientific Instruments and Equipment Development Program (grant 2012YQ140032).

References

- [1]. Y. P. Huang, R. Wu, Y. Sun, et al. Vehicle Logo Recognition System Based on Convolutional Neural Networks with a Pretraining Strategy. IEEE Transactions on Intelligent Transportation Systems. Vol. 16 (2015) No. 4, p. 1951-1960.
- [2]. W. Lu, H. Zhang, K. Lan, et al. Detection of Vehicle Manufacture Logos Using Contextual Information. 9th Asian Conference on Computer Vision, Xi'an, China, September 23, 2009, p. 546-555.
- [3]. S. Mao, M. Ye, X. Li, et al. Rapid Vehicle Logo Region Detection Based on Information Theory. Computers & Electrical Engineering, Vol. 39 (2013), No. 3, p. 863-872.
- [4]. Y. Liu, S. Li, A Vehicle-logo Location Approach Based on Edge Detection and Projection. IEEE Int. Conf. Vehicular Electron Safety, Beijing, China, July, 2011, p. 165-168.
- [5]. L. Itti, C. Koch, E. Niebur, et al. A Model of Saliency-based Visual Attention for Rapid Scene Analysis. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 20 (1998), No. 11, p. 1254-1259.
- [6]. F. Zhang, Y. Shen, H. Chang, et al. A Visual Saliency Based Method for Vehicle Logo Detection, Fifth Int. Conf. Digital Image Processing. International Society for Optics and Photonics, Beijing, China, April, 2013, pp. 2714-2739.