ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 7, No. 3, 2012, pp. 235-240

Saliency and Active Contour based Traffic Sign Detection

Shangbing Gao [1+] and Yunyang Yan [1]

[1] The School Computer Engineering, Huaiyin Institute of Technology, Huai'an, 223003, P.R.China

(*Received January 25, 2012, accepted May 20, 2012*)

Abstract. In this paper, we propose a new approach to detect salient traffic signs, which is based on visual saliency and auto-generated strokes for image segmentation. The proposed algorithm deals with two tasks on detecting traffic signs: auto-location and auto extraction. Firstly, inspired by recent work of visual saliency detection, we obtain the location of traffic signs in a natural image by multi-scale principle component analysis (MPCA). Secondly, in order to extract traffic signs, auto-generated strokes are used instead of drawing the strokes by the users, the sign board area is extracted using localizing Region-Based Active Contour. Extensive experiments on public datasets show that our approach outperforms state-of-the-art methods remarkably in salient traffic sign detection. Moreover, the proposed detection method has higher accurate rate and robustness to different natural scenes.

Keywords: Multiscale; Image segmentation; active contour; traffic sign

1. Introduction

Automatic extraction of traffic sign is a hot topic of intelligent transportation systems [1]. It has been widely applied in many applications such as: driving safety and automatic vehicle guidance etc. Traffic sign detection has two key goals: location and extraction. Because of complex traffic sign images, color-based and shape-based methods can not rapidly find the object and deal with illumination, viewpoint change. For example, Loy et al. [2] applied the radial symmetry algorithm [3] to detect regular polygons, which needs to set parameters in advance and unfit for all the signs. Moreover, the fully automatic segmentation of traffic sign from the background is very difficult; and there is not a mature and integrated method for traffic sign detection system until now. Therefore, to develop a real-time and accurate road traffic sign detection system has been a challenging task in computer vision.

Humans can identify salient areas in their visual fields with surprising speed and accuracy before performing actual recognition. Computationally detecting such salient image regions remains a significant goal, as it allows preferential allocation of computational resources in subsequent image analysis and synthesis. A number of very inspiring and mature saliency models have been recently introduced in the literature. Itti et al. [4] introduced a saliency model which was biologically inspired. Specifically, they proposed the use of a set of feature maps from three complementary channels as intensity, color, and orientation. The normalized feature maps from each channel were then linearly combined to generate the overall saliency map. Based on Itti's algorithm, many saliency models have appeared, such as, SR [5] , Duan's method[6].

Our work is inspired by [6] which defined the saliency in three elements: dissimilarity, spatial distance and central bias. Motivated by Duan's model, a new saliency model is proposed based on multi-scale principle component analysis (MPCA).

We use first two information to measuring the patch's saliency value. The central bias, which based on the principle that dominant objects often raise to the center of the image, is proposed by [7]. This underlying hypothesis brings two problems. First, background near the center of image maybe more salient than the foreground which is far away from the center. Second, for a salient object, the part near the center is more salient than that far away from the center. To diminish this effect, we give up the central bias and use the multiple scales to decrease the saliency of background patches, improving the contrast between salient and non-salient.

We get traffic sign's location using the proposed algorithm. Based on an interactive segmentation method proposed by Shawn et al. [8], which is semi-automatic image segmentation. In order to extract traffic

+ Corresponding author. Tel.: +86 25 83591046
E-mail address: luxiaofen_2002@126.com

sign region, we achieve automatic segmentation by auto-generating strokes after location. The proposed algorithm can make salient region more significant than Duan's algorithm, as shown Fig.1, and has good robustness to adverse environment.

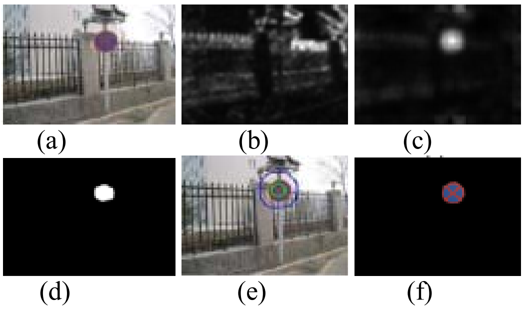


Figure 1. An example of traffic sign detection. (a) Original image; (b)Saliency map by Duan's method; (c) Saliency map by MPCA; (d) Binary map; (e) Auto-generated strokes; (f) Extraction result.

This paper is organized as follows: section 2 describes the proposed traffic sign detection method; section 3 performs extensive experiments to verify the proposed method; and section 4 shows the conclusion of this paper.

2. THE PROPOSED TRAFFIC SIGNS DETECTION METHOD

Visual saliency often appears in all kinds of visual scales, different regions also have different visual attention. In order to comprehensively consider locality and integrity for the image saliency region, in this paper, we present an improved algorithm (MPCA) on the basis of Duan's model to compute traffic sign saliency map. And then we extract traffic sign part from original image using automatic segmentation method obtained by loading auto-generated strokes.

2.1. Multi-scale PCA method

Given an image I with dimension $H \times W$, non-overlapping patches with the size of $n \times n$ pixels are drawn from it. The total number of patches is $L = \lfloor H / n \rfloor \cdot \lfloor W / n \rfloor$. Denote the patch as $p_i, i = 1,2,\cdots,L$. Then each patch is represented as a column vector x_i of pixel values. The length of the vector is $3k^2$ since the color space has three components. Finally, we get a sample matrix $X = [x_1, x_2, \cdots, x_L]$, L is the total number of patches as stated above.

To effectively describe patches in a relatively low dimensional space, we used an equivalent method to PCA to reduce data dimension. Each column in the matrix X subtracts the average along the columns. Then, we calculated the co-similarity matrix $A = (X^T X) / L^2$, so the size of the matrix A is $L \times L$ .The eigenvalues and eigenvectors were calculated based on the matrix A selected with their eigenvector $U = [u_1, u_2, \cdots, u_d]^T$ according to the biggest d eigenvalues, where u_i is an eigenvector. The size of the matrix U is $d \times L$.

New algorithm considers two factors for evaluating the saliency: the dissimilarities of color between image patches in a reduced dimensional space, and their spatial distance.

A patch is salient if the color of its pixels is unique. We should not, however, look at an isolated patch, but rather at its surrounding patches, which lead to a center-surrounding contrast. Thus, a patch p_i is considered salient if the appearance of the patch p_i is distinctive with respect to all other image patches.

Specifically, let $dist_{color}(p_i, p_j)$ be the distance between the patches p_i and p_j in the reduced dimensional space. Patch p_i is considered salient when $dist_{color}(p_i, p_j)$ is high for $\forall j$.

$$dist_{color}(p_i, p_j) = \sum_{n=1}^{d} \left| u_{ni} - u_{nj} \right| \tag{1}$$

The positional distance between patches is also an important factor. Generally speaking, background patches are likely to have many similar patches both near and far-away in the image. It is in contrast to salient patches that the latter tend to be grouped together. This implies that a patch p_i is salient when the patches similar to it are nearby, and it is less salient when the resembling patches are far away.

Let $dist(p_i, p_j)$ be the Euclidean distance between the positions of patches p_i and p_j, which is represented by the two centers of patches p_i and p_j in the image, normalized by the larger image dimension. Based on the observations above we define a dissimilarity measure between a pair of patches p_i and p_j as:

$$dissimilarity(p_i, p_j) = \frac{dist_{color}(p_i, p_j)}{1 + dist(p_i, p_j)} \tag{2}$$

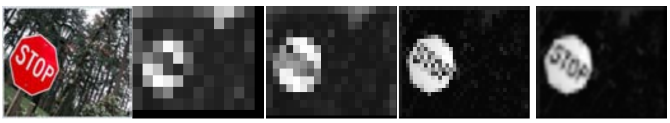
This dissimilarity measure is proportional to the difference in appearance and inverse proportional to the positional distance.

To evaluate a patch's uniqueness, we can compute the dissimilarity between the patch and all of other patches and take the sum of these dissimilarities as the saliency of related patch. In practice, there is no need to incorporate its dissimilarity to all other image patches. It suffices to consider the K most similar patches that if the most similar patches are highly different from p_i, then clearly all image patches are highly different from p_i. Hence, for every patch p_i, we search for the K most similar patches $\{q_i\}, i = 1,2,\cdots,K$ in the image, according to (2). Under this definition, our algorithm that measures the saliency value from the perspective of global information and local information is different from global saliency detection [20] and Duan's method [6]. A patch p_i is salient when $dissimilarity(p_i, q_k)$ is high for $\forall k \in [1, K]$. The saliency of patch p_i is defined as (we choose $K = 100$ in our experiments):

$$S_i = 1 - \exp\{-\frac{1}{K}\sum_{k=1}^{K} dissimilarity(p_i, q_k)\} \tag{3}$$

Based on the observation that patches in background are likely to have similar patches at multiple scales, which is in contrast to more salient patches that could have similar patches at a few scales but not at all of them(It is equal to the principle proposed by [10] that salient object always smaller than the background). Therefore, we wish to incorporate multiple scales to further decrease the saliency of background patches, improving the contrast between salient and non-salient regions.

In addition, the patch with large scale can not describe the boundary of small salient object. So we hope to use different scales that large scale to detect the whole information and the small scale to describe the salient object in details. Last, we compile all saliency value into final saliency. The results of different scales and the final result illustrated in Fig.1. The number of PCs set to 4.



a b c d e

Figure 2.Our method using different scales. a.input, b.scale=30, c. sacle=20, d.scale=10, e. final result

For a patch p_i of scale r, the saliency value according to (3) is defined as

$$S_i^r = 1 - \exp\{-\frac{1}{K}\sum_{k=1}^{K} dissimilarity(p_i^r, q_k^r)\} \tag{4}$$

We consider the scales $R_c = \{r_1, r_2, \cdots, r_M\}$, using (4) to calculate the saliency of patch i as $\{S_i^{r_1}, S_i^{r_2}, \cdots, S_i^{r_M}\}$. The final saliency is computing as

$$S_i = \frac{1}{M}\sum_{r \in R_c} S_i^r. \tag{5}$$

2.2. Traffic signs extraction

Saliency map only provides coarse information about where traffic signs locate in the original image. Traffic sign extraction is also an important task and necessary for road traffic sign recognition work. Image segmentation can be applied to extract the prominent regions. In very complex images the fully automatic segmentation of the object from the background is very difficult.

In [4], Shawn et.al proposed a natural framework that allows any region-based segmentation energy to be re-formulated in a local way. For example, we choose the global region-based energy that uses mean intensities is the one proposed by Yezzi et al. [9] which we refer to as mean separation energy:

$$E_{MS} = \int_{\Omega_y} (u - v)^2 \tag{6}$$

where u and v represents the global mean intensities of the interior and exterior regions, respectively. Optimizing this energy causes that the interior and exterior means have the largest difference possible.

$B(x,y)$ is introduced to mask local regions. This function $B(x,y)$ will be 1 when the point y is within a ball of radius r centered at x, and 0 otherwise.

Accordingly, the corresponding F is formed by localizing the global energy with local mean equivalents as shown:

$$F_{MS} = (u_x - v_x)^2 \tag{7}$$

We can get the following local region-based flow:

$$\frac{\partial \phi}{\partial t} = \delta\phi(x)\int_{\Omega_y} B(x,y)\delta\phi(y)\cdot(\frac{(I(y)-u_x))^2}{A_u} - \frac{(I(y)-v_x))^2}{A_v})dy + \lambda\delta\phi(x)div(\frac{\nabla\phi(x)}{|\phi(x)|}) \tag{8}$$

where A_u and A_v are the areas of the local interior and local exterior regions respectively given by

$$A_u = \int_{\Omega_y} B(x,y)\cdot H\phi(y)dy \tag{9}$$

$$A_v = \int_{\Omega_y} B(x,y)\cdot H(1-\phi(y))dy \tag{10}$$

Two closed curve lines are marked in an original image as the object marker and the background marker respectively. As shown Fig.1 (e), the green line is the object marker and the blue line is the background marker.

3. Experimental results

Our algorithm is implemented in Matlab 7.0 on a 2.8-GHz Intel Pentium IV PC. Extensive experiments are performed to validate the proposed method. First, we compare the proposed MPCA with Itti's model, SR and Duan's models in saliency evaluation, as shown Fig.3. Finally, we test the proposed algorithm under different surroundings, as shown Fig.4. all the images for test are gained from website and public datasets in our experiments, including different scene conditions, different colors and shapes.

3.1. Evaluation and comparison with other saliency models

To compare our results with [6], we chose 11 as reduced dimension which is the best value to maximize saliency predictions. For the patch size, we choose $\{30,20,10\}$ because better results are easy to obtaining in these values [6]. We obtained an overall saliency map by using YCbCr color space in all experiments. Some visual results of our algorithm are compared with state-of-art methods in Fig. 3.

Note that our method is much less sensitive to background texture, which is different from Itti's method, Duan's method and SR. However, our visual saliency model can accurately find traffic sign area and keep complete area information, making the road traffic region more prominent than other models.

3.2. Results by using the proposed method in different scenes

We apply our method to test its performance on some adverse conditions, such as some traffic signs affected by illumination changes, occlusion, rotation, and shadow, different sizes.

Fig.4 (a) shows the detection results using the proposed model on some unfavorable conditions. Experimental results show the proposed method can make traffic sign region "popup" in an image of even bad environment and robust to viewpoint change and insufficient illumination. Fig.4 (b) shows the experimental results of the proposed traffic sign detection model in various suburban and highway scenes. The simple background on highway scene can make the detection more convenient. However, in urban areas and campus as shown Fig.4 (c), there may be some complex visual objects, such as trees and buildings, pedestrians and motor vehicles. Traffic sign detection becomes more difficult if only using color-based and shape-based detection method. From experimental results, we can see the proposed model is also invariant to traffic sign scale change, and our model does not rely on the size and shape of traffic sign, and can still accurately detect traffic sign area in complex scene.

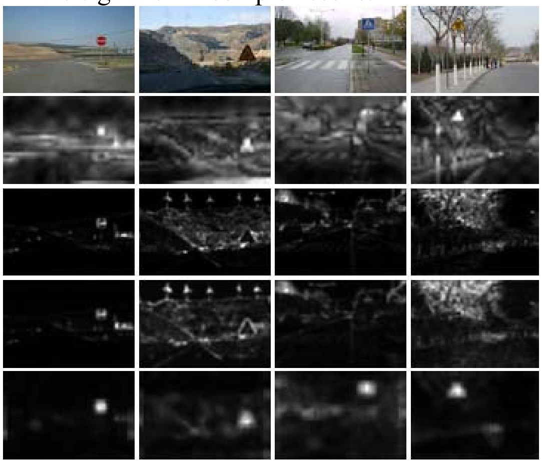
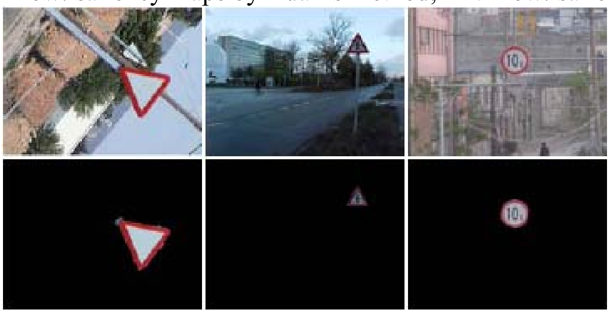
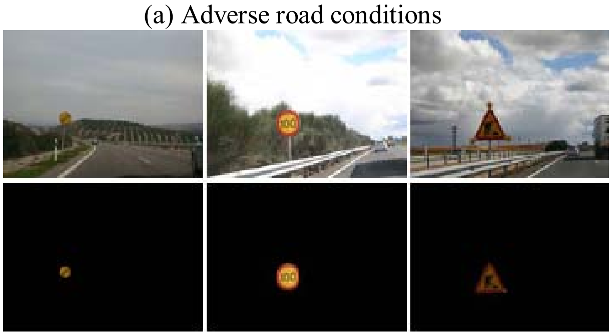


Figure 3. The comparison of traffic signs saliency maps. Second row: saliency maps by Itti's model [4]; Third row: saliency maps by SR; Fourth row: saliency maps by Duan's method; Fifth row: saliency maps by MPCA model.



(a) Adverse road conditions



(b) Highway scene

(c) Urban area and campus road scenes

Figure 4. Results of traffic sign detection using the proposed model in different conditions. (a) Experimental results of traffic signs affected by adverse external factors containing tilted, insufficient illumination and dustsandstorm weather; (b) Experimental results of highway signs in different sizes and simple background; (c) Experimental results of road signs with complex background in urban area and campus.

4. Conclusion

In this paper we introduce a multiscale detection method of visual perception based on Duan's model to obtain more salient area and lessen redundant information. Firstly, we obtain multiscale images using PCA, then compute saliency maps and label in original images, finally extract traffic signs area by automatic image segmentation. Through computing auto-generated strokes image in advance, we can realize automatic segmentation for traffic signs. The proposed algorithm could deal with traffic signs in the presence of light, scale and viewpoint change, experimental results test our model can achieve a high detection rate. But the proposed algorithm has some limitations for quite a few complicated pictures, in the future work, we will consider improving the algorithm, and make it more universal for all the traffic signs.

Acknowledgements

This work was supported by the grant from Natural Science Research Project of Jiangsu Provincial Colleges and Universities, No. 11KJD520003, Natural Science Research Key Project of Jiangsu Provincial Colleges and Universities, No. 11KJA460001, the Jiang Su Qing Lan project and Huai'an produce-learn-research projects, No. HAC201113， HASZ2011024、HAG2010030、 HAG2011041, fund of Huai'an 533 project

5. References

[1] Kastrinaki, V. and Zervakis, M. and Kalaitzakis, K, *A survey of video processing techniques for traffic applications*, Journal Image and Vision Computing, 21(4)(2004), pp.359-381.

[2] G. Loy and N. Barnes, *Fast shape-based road sign detection for a driver assistance system*, in Pro. IEEE/RSJ Intelligent Robots and Systems, 1(2004) ,pp.70-75.

[3] G. Loy and A. Zelinsky, *Fast radial symmetry for detecting points of interest*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(8) (2003), pp. 959–973.

[4] L. Itti, C. Koch, and E. Niebur, *A model of saliency-based visual attention for rapid scene analysis*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(11)(1998), pp.1254–1259.

[5] Hou X.D, Zhang L.Q, *Saliency detection: A spectral residual approach*, In IEEE Conf Computer Vision and Pattern Recognition, (2007),pp.1-8.

[6] L. Duan, C.Wu, J. Miao, L. Qing and Y. Fu, *Visual Saliency Detection by Spatially Weighted Dissimilarity*, IEEE Conference on Computer Vision and Pattern Recognition (CVPR),(2011),pages 21-23.

[7] B. W. Tatler, *The central fixation bias in scene viewing: Selecting an optimal viewing position independently of motor biases and image feature distributions*, Journal of Vision, 7(14)(2007), pages 1-17.

[8] Shawn Lankton,Allen Tannenbaum, *Localizing Region-Based Active Contours*, IEEE Transactions on Image processing, 17(11)(2008),pp.2029-2039.

[9] J. A. Yezzi, A. Tsai, and A.Willsky, *A fully global approach to image segmentation via coupled curve evolution equations*, Journal of Visual Communication and Image Representation, 13(1) (2002),pp.195–216.