ACADEMIC
World Academic Union

ISSN 1746-7659, England, UK
Journal of Information and Computing Science
Vol. 9, No. 3, 2014, *pp*. 233-240

The Traffic Sign Detection based on the visual saliency

Yan Zhang[1] ,Shangbing Gao[1+], Shiliang Xu[1], Yue Zhang[1]

[1] Faculty of Computer Engineering, Huaiyin Institute of Technology, Huai'an, 223003, P.R.China
(*Received November 10, 2013, accepted January 12, 2014*)

Abstract. In this paper, we propose a new approach to detect salient traffic signs, which is based on visual saliency and the auto segmentation method based on region merging. The proposed algorithm introduces the visual saliency model, and puts forward the top-down and bottom-up two-way fusion mechanism, tried to consciously take the initiative to find and simulate human visual unconscious passive visual process by attracting interact. Then the segmentation method based on region merging is used to extract the traffic sign. Experiments show that the two-way fusion mechanism in target detection has fast processing speed and high accuracy. 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: visual saliency; Image segmentation; region merging; 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].

Our work is based on our previous work [6]. The extraction method based on MPCA [6] surprisingly has a performance superior to the other popular methods, but is not the most effective. To remedy such shortcoming, bi-directional integration mechanisms have been proposed. First of all, bi-directional integration mechanism is used to quickly locate the target in the image, and then the segmentation method based on region merging by maximal similarity was adopted in order to completely detect the traffic signs.
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

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

proposed by Ning et al. [8], which is semi-automatic image segmentation. In order to extract traffic sign region, we achieve automatic segmentation by auto-generating strokes after location.

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

Bi - directional integration mechanism is actually the fusion of two methods the "top-down" active search and "bottom-up" salient area detection. The "top-down" active search uses partition - first search strategy, that is, the most likely object area is the first class search area. If the object is not searched, then turn to the next level may be the regional search, or to a centre for the search area to expand your search scope

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