A NEW PARADIGM FOR THE COMPUTATION OF CONSPICUITY OF TRAFFIC SIGNS IN ROAD IMAGES

Ludovic Simon, Jean-Philippe Tarel, Roland Brémond

Laboratoire Central des Ponts et Chaussées (LCPC)
58 Bd Lefebvre, 75015 Paris, France

ABSTRACT

The estimation of conspicuity is of importance for engineers who aim at making traffic signs conspicuous enough to attract attention regardless of drivers' preoccupation. Unfortunately, conspicuity remains a poorly understood attribute due to the relatively limited - although growing - knowledge about the human visual processing system. Our goal is to develop a system which estimates the conspicuity of traffic signs based on the processing of images acquired with a camera onboard a vehicle, in order to be able to make a diagnosis regarding their conspicuity. Aside from specific feature known to be of importance for road signs, there is currently no complete model for conspicuity. However, a computational model for attentional conspicuity was proposed in computer science. This model is based on vision science knowledge of the low levels of the human visual processing system, and we show that it is not suitable for sign detection tasks. We thus propose a new paradigm for conspicuity estimation in search tasks based on statistical learning of the features of the searched object.

Keywords: road safety, conspicuity, saliency, object detection, image processing, statistical learning, visibility, visual performance, quality level.

1. INTRODUCTION

A driver's attention is directed at some particular stimuli among those available in the road environment, and what attracts his attention is related to both psychological factors and to the photometrical and geometrical characteristics of the road scene. The degree to which an object attracts attention compared to its scene background is the so-called conspicuity.

Traffic signs are designed to attract driver’s attention regardless of their preoccupation. The estimation of the conspicuity of traffic signs from digital images is therefore of practical importance. Despite this need, as explained in [1], we only know some of the characteristics which account in the resulting conspicuity. The question of how to combine these characteristics is still open. This means that a complete conspicuity model and a method to estimate the conspicuity of an object from an image remain unsolved problems. This is partially due to our relatively limited knowledge about the human visual processing system and also to the fact that measuring driver’ attention is subject to many difficulties, even with eye-tracking. One of the main difficulties in developing a computational model of the conspicuity is to take into account both the photometrical and geometrical aspects of the problem.

Our long term goal is to develop a system to automatically estimate the conspicuity of traffic signs from digital images. A major application will be to use this system with a camera onboard a vehicle, in order to be able to perform diagnosis of traffic signs along a road network.

2. ATTENTIONAL CONSPICUITY

In [2], one of the most used computational model for attentional conspicuity (or saliency) was proposed, based on vision science knowledge of the low levels of the human visual processing system. From a digital image, the algorithm computes the associated saliency map by linear combination of saliency maps of different elementary vision patterns: corners, luminance contrasts, red-green and blue-yellow color contrasts, at several spatial scales. In [3], the saliency model was tested with success when the task is to memorize an image. But when the task is to search for a particular object, it is no longer valid. With a simple experimental protocol, we have tested this saliency model in a driving context for the specific case of traffic lights. On 26 road images, we computed the traffic lights saliency according to [2], then we presented the images to 13 subjects in a psychophysical experiments. The participants were asked a question which told us whether they had seen the traffic lights or not, without focus only on it. Results show no correlation between performances of the subjects and the saliency estimated from the computational model. In [4], we detail this experiment and we discuss results. We propose as an explanation that in the driving context, traffic sign detection is not a pure attentional task but also a search task. Indeed in the saliency model, no prior information about the object or the class of objects of interest is used. That is to say that this computational model is too generic for application in driving tasks, see also more complex models of objects conspicuity in images [5, 6].

3. A NEW PARADIGM FOR A COMPUTATIONAL CONSPICUITY MODEL

The previous conclusions lead us to search for a new paradigm for the computational estimation of the search conspicuity of traffic signs in real road environment, based on modeling priors on the objects of interest. Our approach is to perform this modeling by using recently developed statistical learning algorithms.

Figure 1. The original image on the top left is processed using the attentional conspicuity model on the top right, and modeling priors "no entry" signs on the bottom. From left to right, conspicuity model are computed using different kinds of feature vectors: pixel colors, color histograms on windows, small windows.

In the last decade, many new and efficient learning algorithms were proposed such as the ones derived using the "Kernel trick" [7]. The best known algorithm in this category is the so-called Support Vector Machine (SVM) algorithm which demonstrates reliable performances in learning object appearances in many pattern recognition applications [7]. The learning is performed from a set of positive and negative examples of feature vectors. Each positive feature vector corresponds to a sample of the aspect of the object of interest, when each negative feature vector is a sample of the aspect of the background. From this set, usually called the learning database, the SVM is able to infer the frontier that splits the feature space into parts associated to the aspect of the object of interest and parts associated to the background. Once the learning stage is performed the resulting classifier can be used for object recognition and detection in any new images. Moreover, the SVM also gives a positive confidence value for each recognition decision. When the confidence is between 0 and 1, the decision is not reliable. The proposed new paradigm is to rely on a learning algorithm for modeling the aspect of the object of interest and to define the search conspicuity of a given patch in an image as the confidence in recognizing this patch as the object.

The proposed paradigm allows us to design algorithms for the estimation of the search conspicuity of different kinds of objects in road and other environments. An example of the difference between results obtained with the attentional conspicuity model and our search conspicuity model for "no entry" traffic sign is shown in Figure 1. It is clear from the location selected as conspicuous that the proposed approach outperforms the saliency model.

4. ON THE CHOICE OF FEATURES

The question of how to combine the photometrical and geometrical aspects in the feature vector is not solved by the use of learning algorithms. To tackle this question, we ran experiments to compare conspicuity estimates obtained with different kinds of features. For each feature, the learning is performed on 29 “no entry” sign of various aspects. Three kinds of features with different complexities are selected: pixel colors, color histograms in a square window, pixel colors within a small window. Different learning algorithms are used depending of the feature type. For pixel color, we simply build a color histogram of the “no entry” signs and of the background. For color histogram feature, we use SVM learning. For pixel colors within a window, we use Parzen windowing approach [8]. 20 images of road scenes were processed. Examples of obtained maps with conspicuous patches are shown in figure 1 for one image.

Figure 2. On the left, the scanpath of one subject searching for "no entry" signs. Each circle represents a fixation. The duration is indicated in ms. The gaze starts at the image center. Note that the sign on the left is missed. The image on the right shows the predicted conspicuous locations using color histogram in windows of different sizes.

We ran a psychophysical experiment with three subjects on the same 20 images in order to have a reference. The subjects were asked to count the “no entry” signs in each image. The answers allowed to check if all signs had been seen or not. By using SensoMotoric Instruments’ remote eye tracker, gaze positions on the images were also recorded and processed to detect the positions and durations of the subjects’ fixations. An example of typical scan paths is shown in Figure 2. All four “no entry” signs are predicted as conspicuous but the subjects only look at three, whereas they count four. This can be explained by the fact that the subjects need not to focus on very conspicuous signs to detect them, they may rely on parafoveal vision. We observed however that the locations predicted as conspicuous using the different feature types are most of the time fixation locations. Among those that we tested, the color histogram in windows feature gives the best results in terms of number of false alarms and correct predictions. Subjects also focus on objects important to analyse in the road scene such as vehicles, which is a natural behaviour in a driving task.

The link between the predicted conspicuity and the mean duration of subjects’ gaze is displayed in Figure 3 for “no entry” signs. Conspicuity is predicted using SVM on color histogram windows. It is interesting to notice that the lower the conspicuity, the higher the duration. Moreover, a threshold around 280 ms appears which, according to behaviourists [9], corresponds to the minimum duration needed for a decision. The more conspicuous signs are analysed in less than 280 ms, whereas the others, less conspicuous according to the prediction, are examined for a longer duration. These results are preliminary due to the reduced number of subjects.

5. CONCLUSION AND FUTURE WORK

From the analysis of the limits of the available computational models of conspicuity, we propose a new paradigm to define search conspicuity. From our preliminary experiments with subjects, the new model seems to outperform the previous attentional model. In future work, we will continue to test our model for other road signs and other objects, to refine our conclusions on the optimal choice of the type of features. Particularly, it seems necessary to combine shape information with color histograms in windows, in order to better account for human capabilities.

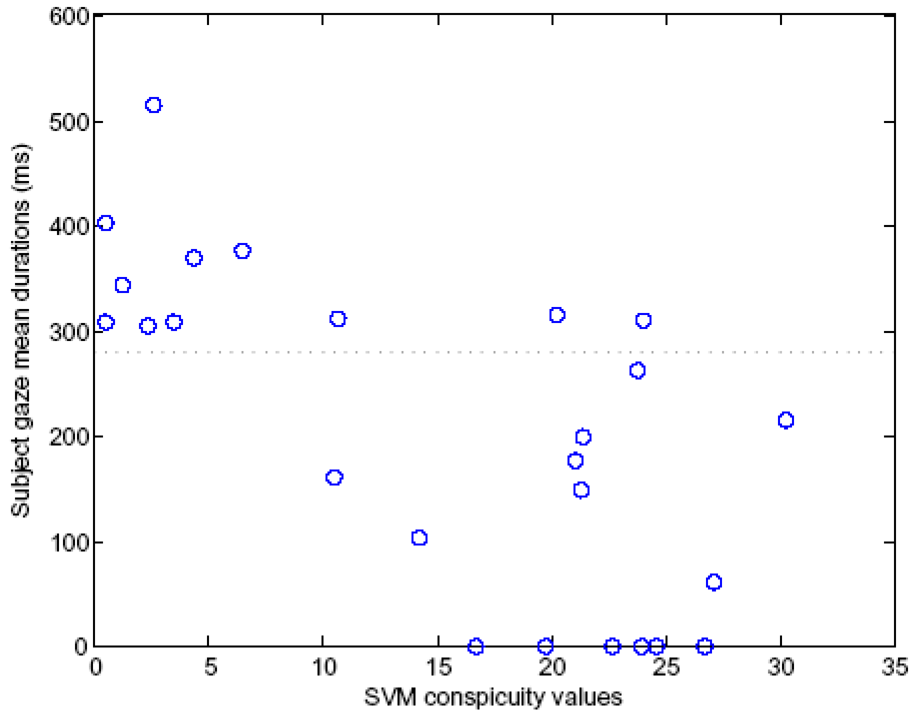

Figure 3. Mean duration of the subjects' gaze versus the predicted "no entry" sign conspicuity.

REFERENCES

[1] CIE 137 - 2000. The conspicuity of traffic signs in complex backgrounds.

[2] ITTI, L, KOCH, C, NIEBUR, E. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. on Pattern Analysis and Intelligence,* 1998, **20** (11), 1254-1259.

[3] UNDERWOOD, G, FOULSHAM, T, VAN LOON, E, HUMPHREYS, L, BLOYCE, J. Eye movements during scene inspection: A test of the saliency map hypothesis. *European Journal of Cognitive Psychology,* 2006, **18**(3), 321-343.

[4] BREMOND, R, TAREL, J-P, CHOUKOUR, H, DEUGNIER, M. La saillance visuelle des objets routiers, un indicateur de la visibilité routière. Proc. *Journées des Sciences de l'Ingénieur du LCPC*, 2006.

[5] NAVALPAKKAM, V and ITTI, L. Modeling the influence of task attention. *Vision Research,* 2005, **45**, 205-231.

[6] SUNDSTEDT, V, DEBATTISTA, K, LONGHURST, P, CHALMERS, A. Visual Attention for Efficient High-Fidelity Graphics. Proc. *Spring Conference on Computer Graphics,* 2005, 162-168.

[7] SCHOLKOPF, B. *Learning with Kernels,* MIT press, 2002.

[8] TAREL, J-P, BOUGHORBEL, S. Object predetection based on kernel parametric distribution fitting. Proc. *International Conference on Pattern Recognition,* 2006.

[9] JACOB, R J K, KARN, K S. Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises, in HYONA, J., *The Mind's Eyes: Cognitive and Applied Aspects of Eye Movements,* ed.: RADACH, R & DEUBEL, H. Elsevier Science, 2003.

Contact author

Name	Ludovic Simon - Jean-Philippe Tarel - Roland Brémond
Affiliation	Laboratoire Central des Ponts et Chaussées
Address	58, boulevard Lefebvre - 75732 Paris cedex 15 - France
Phone	+33(0)1 4043 6556 Fax +33(0)1 4043 5499
E-mail	ludovic.simon@lcpc.fr - jean-philippe.tarel@lcpc.fr - roland.bremond@lcpc.fr