ROAD SIGN RECOGNITION BY MEANS OF THE BEHAVIORAL MODEL OF VISION

Abstract

Algorithms and procedures to solve the task of road sign detection and recognition invariant of viewing conditions and results of testing during computer simulation with British and Russian signs are presented. After preliminary colour segmentation of initial real world images and classification according to road sign colours and external forms, biologically plausible Behavioral Model of Vision (BMV) [1,2] which was modified under task, identified correctly about 80% potential traffic sign images for various weather conditions, shading, and other transformations. Possible ways to increase the model performance are considered.

Introduction. An important task in developing intelligent systems of driver support and traffic safety is road sign detection and recognition [3-5]. At present, the problem of traffic signs recognition invariantly to their possible transformations in real world conditions has no effective solution in the frameworks of standard computer vision approaches. Evidently, some aspects of this problem may be solved by means of biologically plausible approach in image recognition [1,2,4]. In the present paper, modified algorithms of the BMV [1,2], based on imitation of foveal vision properties in the real visual system, and results of their testing for traffic sign recognition under various viewing and weather conditions are presented.

1. Algorithms and procedures. Colour Segmentation. Images from the standard databases are firstly utilised to find the range of colour vectors for the colours used in the signs, mainly red, blue, black and white. The ranges for each colour vector, e.g., (red, lightness, chroma) and (blue, lightness, chroma), are found by calculating the values by use of the CIECAM97 model [6] and are
plotted on a u’v’ chromaticity diagram after conversion. During the study, images of British road signs have been taken using a digital camera, Olympus Digital Camera C-3030, Russian real-world images were got by Fuji photo camera and then scanned (450 dpi). These images are then classified visually according to the viewing and weather conditions, such as cloudy, sunny, etc. Based on the images in each group, the parameters for each viewing condition are found from [7] (e.g., direct sun light with colour temperature 5335K and light from overcast sky with colour temperature 6500K) for application of the colour appearance model. Images taken under real viewing conditions are then transformed from RGB space to CIE XYZ values and then to LCH (Lightness, Chroma, Hue) by use of the model of CIECAM97. Based on the range of sign colours, traffic-signs-to-be are segmented from the rest of scenes for further identification.

Classification of traffic signs according to their external form. For all signs, both from standard databases and from real world images, preliminary classification according to the colour, their external form (circle, rectangle, or triangle) can be determined by means of histograms of orientations detected at resolution level 3 (RL 3). RL 3 is emulated by Gaussian convolution (kernel size is equal 9). Each sign with a certain external form (in spite of its inner content) has characteristic relationship of horizontally, vertically, and obliquely oriented segments at RL 3. In particular, all oriented elements had nearly equal representation for circle signs contrary to rectangle signs (Fig. 1) that had preferable horizontal and vertical orientations (in sum, more than 50% of all oriented segments). For each external form, quantitative estimations were obtained for classification into particular groups of signs from British and Russian standard databases and from real world images.
Recognition algorithms. Traffic sign recognition task was solved by the BMV developed earlier for invariant memorising and recognition of complex grey-level images such as human faces [1,2]. The basic properties of the BMV consist in: (i) space-variant representation of image fragment in each fixation (a change of resolution level from the centre to the periphery of the Attention Window (AW)); (ii) description of image oriented elements in each fixation by 49-dimensional vector; (iii) transition description from a fixation point to the next one, and (iv) image viewing trajectory formation. The basic BMV demonstrated invariance to scale, rotation, noise, shift, and in part to point of view. The peculiarities of the traffic sign recognition task demanded some modifications of the model. First, these modifications are determined by conventional format and relative simplicity of traffic sign images. Second, the basic model invariance to plain rotations should be eliminated because the same geometrical objects in traffic signs rotated in plain (such as arrows) have different meaning for driver. The modifications were performed in the (iii) – (iv) algorithms.

Similar to the basic version of the model, an image fragment in each fixation point is presented by 49-dimensional vector of oriented elements in relative coordinate system (Fig.2, a); orientation of segments near each of 49 AW point is determined by means of the Gaussians with spatially shifted centres. Such procedures allow to receive sign fragment description that is relatively stable to any local image disturbances in some range. Contrary to the basic model version, transition description from a fixation point to the next one was realised in absolute coordinate system to exclude rotation invariance mentioned above. Example of oriented elements detected in a sign fragment is shown in Fig. 2, b.
Formation of trajectory for traffic sign viewing. Some modifications were also developed in algorithms of image viewing trajectory formation for preferable processing of inner informative parts of signs. In particular, initial fixation point of the AW was set near geometric centre of oriented elements detected in a sign (Fig. 2, b). Three different algorithms for formation of image viewing trajectories were developed, namely standard method as in [2], estimation of filling-in of 49-dimensional vector, and calculation of external contour location. Comparison of these algorithms shown that maximum attraction of AW to the inner sign parts during image viewing was provided by an algorithm based on preliminary estimation of location of external image contour elements (Fig. 3). This algorithm was used for description of signs from standard databases at memorising stage. At recognition stage, model-specific description of real world signs were compared with those from standard databases by means of algorithms similar to [2].
Fig.3. Examples of sign viewing trajectories formed by different algorithms: successive fixation points are chosen (a) by standard method as in [2]; (b) by estimation of filling-in of 49-dimensional vector; (c) by calculation of external contour location.

2. Computer simulation. Two standard databases have been established in our study. One is based on British traffic signs (n= 142) scanned from the book of HighWay code. The other is Russian traffic signs (n= 158) obtained from the web site http://www.domkrat.ru/Laws/rules/znak1.shtml. According to conventional standards in traffic sign recognition area, size of each sign both from standard databases and from real world images is normalised to 40x40 pixels.

Memorising. Each traffic sign database was preliminary classified into groups according to colour and external form. Then for each group, each database traffic sign was transformed into a model-specific form. As the first step, all coloured images from a database were converted into grey-level representation. Then, for each database image a specific description was obtained based on trajectory of its viewing according to the inner informative regions of the sign. These descriptions have been stored and form a model-specific database of traffic sign images. The model-specific databases (separately for British and Russian signs) for traffic sign images needs to be built only once. The descriptions or features for each image are then utilised in all further computer experiments on recognition of real world images of traffic signs.

Recognition. After a coloured real world image is segmented using the CIECAM97 colour appearance, it is first converted into a grey-level representation. The BMV model then starts to find representative features from the image-to-be-identified and to search for a hypothesis to be generated about the image in accordance with the model-specific database. During this search the representative description of the query image is compared to the model-
specific description of the database traffic signs. If a successful match occurs the presented image is recognised, and the matched sign image is retrieved.

Experimental results show that the majority (more than 90%) of signs can be segmented correctly by using CIECAM97 colour vision model. After segmentation and classification according to colour and external form, the modified BMV, correctly identified 28 out of 35 potential traffic sign images for Russian signs and 90 out of 109 for British signs, which gives 80% and 82% success rates respectively. Similar results were obtained at different viewing and weather conditions. The non-identified or falsely identified signs are either of low resolution (taken from very far distance, more than 60 meters) or have a very complex information content, for example, the sign “GIVE WAY” with blurry letters, or a complex disturbing background (Fig. 4). The results of testing were identical for 15 and 10 fixation points of AW during image viewing, but they sharply depended from AW positioning on a sign. Recognition time varied from 0.35 seconds up to 0.6 seconds per image on a standard Pentium.

![a) Standard database signs](image1)

![b) Real world signs](image2)

![c) Standard database signs](image3)

Fig. 4. Three examples of recognised (a) and non-recognised (b, c) real world signs.

3. Conclusion. Overall the models-based approach can give accurate identification for traffic signs located at a moderate distance for still images in various weather conditions and shows a good performance for a wide variety of traffic signs of different colours, forms, and informative content. The results of computer experiments with the BMV demonstrate the importance of AW fixation points chosen while viewing trajectory formation to find the most informative regions in traffic sign images. In particular, the number of fixation points of AW during image viewing may be reduced significantly, up to one
fixation for simple signs. Evidently, simple or symmetric signs could be recognised by one fixation model version, complex or asymmetric signs will be recognised by multiple fixation points model version.

Modification of the BMV in accordance with the results of these experiments and the use of special acceleration boards can lead to improvement in its performance and therefore increase its importance for practical applications.

This work is supported by The Royal Society, UK, under the International Scientific Exchange Scheme.

References