Road Sign Recognition by Single Positioning of Space-Variant Sensor Window

Dmitry G. Shaposhnikov, Lubov N. Podladchikova, Alexander V. Golovan, Natalia A. Shevtsova
A.B.Kogan Research Institute for Neurocybernetics, Rostov State University, Russia
nisms@krinc.ru

Kunbin Hong, Xiaohong Gao
School of Computing Science, Middlesex University, London, United Kingdom
K.Hong@mdx.ac.uk, X.Gao@mdx.ac.uk

Abstract

A biologically plausible model of traffic sign detection and recognition invariantly with respect to variable viewing conditions is presented. The model simulates several key mechanisms of biological vision, such as space-variant representation of information (reduction in resolution from the fovea to retinal periphery), orientation selectivity in the cortical neuron responses, and context encoding of information. The model was tested on British traffic signs and demonstrated the ability to recognize these signs from a single position of a space-variant sensor window. After performing colour segmentation and classification and finding the sign centre, 85% of the traffic signs tested were identified under various environmental conditions.

1. Introduction

Many methods for automatic traffic sign identification have been developed with some promising results [1, 5, 9, 13]. However, identification of traffic signs invariantly with respect to various natural viewing conditions still remains a challenging task. In particular, with the account of safety requirements on public roads, more robust and fast algorithms are required to provide the necessary accuracy in recognition of traffic signs.

A biologically plausible model of visual recognition BMV (Behavioural Model of Vision) [8] was previously implemented for solution of the task of traffic sign recognition [1]. That model version required visual processing in multiple positions of the space-variant sensor window (SW) and demonstrated the ability of recognition of traffic signs, invariantly with respect to viewing and environmental conditions, with recognition rate about 80%. Several means for the improvement of recognition efficiency (i.e., increasing recognition rate and reduction of computations) have been proposed. In particular, it was suggested that the signs could be quickly recognized from a single position of SW if this position is close to the sign centre, since for most traffic signs the geometrical centre is also the centre of information content.

Choice of informative image fragments for detailed processing is one of the most important problems in the field of image recognition [2, 7]. There are many algorithms for detection of most informative parts of images in the frameworks of both conventional and biologically plausible approaches. Most of them are image-dependent because each type of images has specific important fragments. For example, for face images the most informative fragments are eyes, nose, and mouth, [2, 7, 12]. As mentioned above, for most traffic signs the most informative fragments are concentrated around the sign centre. However, in traffic sign recognition only a few approaches are known that attempt to use a selection of most informative fragments of signs for detailed processing [1, 5, 9].

Colour is a dominant visual feature, which undoubtedly represents a key piece of information used by drivers. Therefore colour is widely used in traffic sign recognition systems [5, 13], especially for segmentation of traffic sign images from the rest of a scene. The colour segmentation can be also used for finding the sign center.

In this study, traffic signs were segmented from road scenes under various environmental conditions by colour contents using a standard colour appearance model CIECAM97 [4, 6, 11]. First, the colour segmentation and classification based on colours and shapes provided the detection of the sign centre. After that, the sign was recognized from a single position of a space-variant sensor window centered in the sign centre.

2. Algorithms and procedures

British traffic sign images (n=105) for standard database were scanned from the book of Highway Code. These images were used for preliminary testing the developed algorithms and procedures. Besides, they served as prototypes for recognition of real world images. Traffic signs (n=97) have been taken in London under various environmental conditions using a digital camera (Olympus Digital Camera C-3030). According to conventional standards, the size of each sign in both the standard database and in real world images was normalized to 40x40 pixels.

2.1. Colour Segmentation

Images taken from real word under different viewing conditions were preprocessed to find the range of colour vectors for the colours usually used in the signs, namely red, blue, black, and white. This preprocessing was performed using model of CIECAM97 (11). CIECAM97 is a standard colour appearance model recommended by CIE in 1997 for measuring colour appearance under various viewing conditions. This model can estimate a colour appearance as accurate as an average observer. For human perception, the most common terms used for colour description or colour appearance are lightness, chroma, and hue. A representative set of traffic signs was classified visually according to the viewing and environmental conditions, such as cloudy, sunny, etc. Based on the images in each group, the parameters for each viewing condition were found from [4] (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. Test images taken under real viewing conditions were transformed from RGB space to CIE XYZ values and then to LCH (Lightness, Chroma, Hue) using the model of CIECAM97. The lightness was similar for red and blue signs and background. Therefore, only Hue and Chroma were used for segmentation.

Table 1: *The ranges of Hue and Chroma for red and blue signs.*

Colour	Hue	Chroma
Red	393 - 423	57 - 95
Blue	280 - 290	57- 95

Based on the range of sign colours, traffic-signs-to-be are segmented from the rest of scenes for further identification and classified according to the colour. Only blue and red signs were used in this study (Table 1). All sign images with size more than 10x10 pixels (pictures were taken within 100 meters distance) could be segmented correctly. Sometimes, some other contents, such as the rear red lights of cars were also segmented. However, these non-sign segments could be rejected during recognition stage using BMV.

Information about the colour of the sign was also used for the localization of sign centre based on estimation of the location of colour contour elements.

2.2. Classification of traffic signs according to their shapes

After preliminary colour classification, the signs were further classified by shapes (circle, rectangle, or triangle) which were detected using histograms of oriented segments detected on a sign image. Each sign with a certain shape has its own characteristic pattern of oriented segments. In particular, elements of various orientation have nearly equal representation for circle signs in contrary to rectangle signs (Fig. 1, a), which in turn have preferably horizontal and vertical oriented elements (in sum, more than 50% of all oriented elements). This simple method provides classification of signs in frameworks of each colour group into two subgroups according to their shapes: blue rectangle and circular signs and red triangle and circular signs. Quantitative parameters for this classification were obtained for particular groups of signs in the standard database. The same parameters were used for classification of real world images.

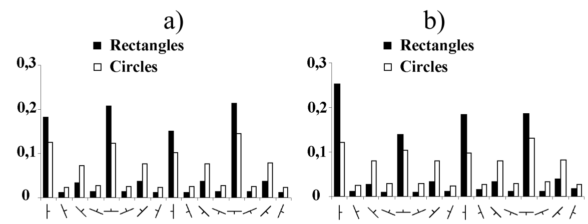

Figure 1: *Averaged histograms of oriented element representation for blue traffic signs in (a) standard database (n=76) and (b) real world images (n=56).*

2.3. Determination of sign centre for positioning of sensor window

The recognition method of the BMV model [8] is based on the encoding of the image according to the path of image viewing and the description of image fragments at each position of SW by a set of primary features (oriented elements). For traffic sign images, we have found that an effective recognition may be performed from a single position of SW if the latter is placed in about the sign centre. Such location of the SW appears to provide the most specific sign description.

The algorithm for determination of colour contour geometric centre has been developed to find the centre of the internal informative part of signs. In the algorithm, spatial location of colour contour elements in the real world signs was determined on the basis of quantitative estimations of RGB composition for a signs of a given colour in the standard database. This algorithm provided for the exact normalization of sign size to 40x40 pixels, and extraction of a "pure" (without background) real world sign (Fig. 2). This algorithm provided for determination of the geometric centre of a sign with necessary accuracy (up to 6 pixels).

a) b)

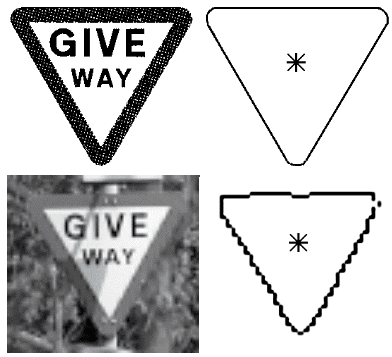

Figure 2: *Examples of determination of the sign centre. Images from the standard database (upper row) and a real world picture (lower row) are shown in (a); * indicates location of the sign centre determined by colour contour (b).*

2.4. Feature description of traffic sign

The description of each sign at the memorising stage (for standard database images) and recognition stage (for real world images) was provided by the specific structure of the SW, which imitates some features of the real visual system such as space-variant representation of information from the centre (fovea) to the periphery of the retina [10, 12], neuronal orientation selectivity [3], and context encoding of the foveal information [12].

Basic algorithms of sign processing by space-variant SW were similar to [8], namely: (i) an image was represented by 49-dimensional vector of oriented elements extracted in vicinity of each of 49 nodes of SW; (ii) the SW nodes were located at the intersections of sixteen radiating lines and three concentric circles with increasing radii (Fig.3, a); (iii) orientation of segments in the vicinity of each SW node is determined by means of calculation of the difference between two oriented Gaussians with spatially shifted centres; (iv) space-variant image representation is emulated by Gaussian convolution with different kernels depending on distance from the SW centre. The SW size increased to 36 pixels (instead of 16 pixels in the basic BMV [8]), and kernel sizes were changed to process a sign from a single position of the SW, i.e., they were equal to 5x5 for the central (foveal) part of the SW, 7x7 for the immediate (parafoveal), and 9x9 for the peripheral part. Besides, estimation of oriented elements in the context area of 48 SW nodes (except for the central node) was used to receive a detailed feature description: it was equal to 3x3 pixels for16 nodes located in the central circle of the SW, 5x5 for the immediate, and 7x7 for the peripheral circle. Each SW node was described by two values, namely, an orientation node was described by two values, namely, an orientation

a) b)

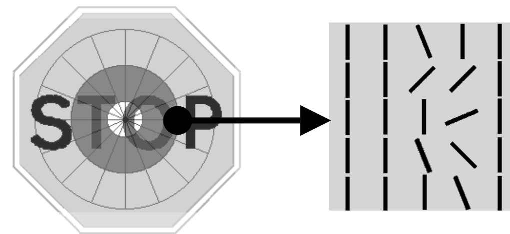

Figure 3: *(a) Schematic of the SW located in the centre of the informative part of a sign. Circles with different gray levels represent different resolutions in the SW structure. (b) Oriented elements detected in the context area of a SW node (indicated by a small black circle in (a)).*

dominated in the context area and a density of oriented elements detected in this area (Fig. 3, b). Such structure of the SW and its location in the sign centre provided maximal representation of oriented elements (up to 90%) in the informative region of a sign at the first and second resolution levels. An example of detection of oriented elements in the context area of the indicated node of a sign is shown in Fig. 3, b.

2.5. Recognition algorithms

The 49-dimensional vector for an incoming traffic-sign-to-be image preliminary classified by colour and shape was compared with descriptions of database images of the corresponding subgroup by the formula:

$$K^b = \sum_{i=0}^{i<49} \left[\mathrm{sgn}\left(Or_i^b - Or_i^{rw}\right) \cdot \left(1 - abs\left(\rho_i^b - \rho_i^{rw}\right)\right)\right]$$

$$\mathrm{sgn}(x) = \begin{cases} 1, & if \;\; x = 0; \\ 0, & overwise; \end{cases}$$

where *Or* is dominating segment orientation in the context area of a given SW node (orientations are determined by the step 22.5° and indicated as 1, 2, ..., 16); superscript *b* stands for prototype database images, *rw* - for the incoming image; ρ is the density of the dominating oriented segment in the context area (see Section 2.4.) of the given SW node. A prototype image from the database with maximal K^b was considered as the result of recognition.

3. Computer simulation

During the stage of memorising, each traffic sign from the database was preliminary classified according to colour and shape. Thus, the database was divided into 4 subgroups. Then for each subgroup, each sign was transformed into a model-specific form (i.e., presented by a 49-dimensional vector of oriented elements extracted in vicinity of each of 49 nodes of SW) and a specific description for each image in each subgroup was obtained. These descriptions were stored in a model-specific prototype database. The model-specific database for traffic sign images needs to be built only once. The descriptions or features for each database image are then utilised in all further computer experiments on recognition of traffic signs.

During recognition, first a test colour image is extracted from the scene using the CIECAM97 colour appearance model. Then the image is classified according to its colour and shape. After that, the sign centre is determined. Then the recognition algorithm searches for a prototype image in the model-specific database. During this search the representative description of the incoming image is compared to the model-specific description of the database traffic signs. If a successful match occurs the presented image is recognized, and the matched database sign image is retrieved.

Our experimental studies have shown that the majority of signs (more than 90%) can be segmented correctly by using CIECAM97 colour vision model. After segmentation and classification according to colour and shape, the model identified 83 out of 97 potential traffic sign images, which gives 85% success rate. Similar results were obtained for different viewing and environmental conditions (87% and 84% for sunny and cloudy weather respectively). The non-identified or falsely identified signs were either of low resolution (taken from very far distance, more than 60 meters) or have information content similar to other signs (for example, the triangle sign “Roundabout”) (Fig. 4), or a complex disturbing background. Recognition time varied from 0.25 seconds up to 0.4 seconds per image on a standard Pentium 233.

a) b) c)

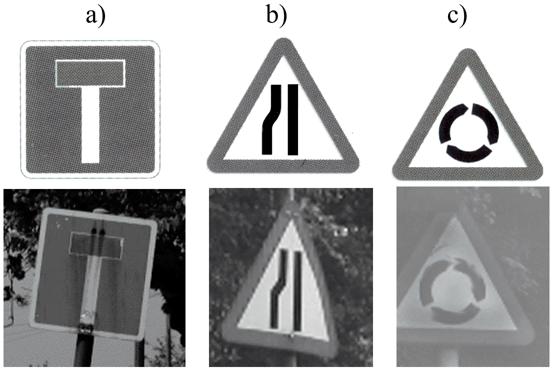

Figure 4: *The examples of recognised (a, b) and non-recognised (c) signs (lower row). The prototype images are shown in the upper row.*

4. Conclusion

In this work we tested our suggestion that the BMV model would be able to recognize traffic signs from a single position of the space-variant sensor window [1]. The presented results confirm this suggestion and even demonstrate an increase of recognition efficiency from a single SW position as compared to the version of the model based on multiple positions of the SW while processing the same signs. Task-oriented modifications of the SW, including increase of the SW size, determination of context in the vicinity of each SW node, setting the SW in the sign centre, etc. – all together allow a detailed feature description of the informative part of a sign. This description is sufficiently compressed while quite enough for effective recognition. In addition, this description is stable to local image disturbances in a certain range. Overall, the described model-based approach provides an accurate identification of the traffic signs located at a moderate distance under various environmental conditions. The resultant recognition system shows a good performance for a wide variety of traffic signs of different colours, forms, and informative content.

Acknowledgement

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

References

[1] X.W. Gao, A.V. Golovan, K. Hong, L.N. Podladchikova, D.G. Shaposhnikov, N.A. Shevtsova. Road Sign Recognition by Means of the Behavioral Model of Vision. *Proc. of Conf. on Neuroinformatics*, January 2002, Moscow, 63-69.

[2] A.V. Golovan, N.A. Shevtsova, L.N. Podladchikova, S.N.Markin, D.G.Shaposhnikov. Image Preprocessing for Identifying the Most Informative Regions in Facial Images. *Pattern Recognition a. Image Analysis*, 2001, 11, 2, 313-316.

[3] D.H. Hubel, T.N. Wiesel. Sequence, Regularity and Geometry of Orientation Columns in the Monkey Striate Cortex. *J. Comparative Neurology,* 1974, 158, 267-293.

[4] D. Judd, D. MacAdam, and G. Wyszecki, Spectral Distribution of Typical Daylight as a Function of Correlated Colour Temperature, *J. Opt. Soc. Am*., 1964, 54(8), 1031-1040.

[5] M. Lalonde, Y. Li. Road Sign Recognition - Survey of the State of Art. *Tech. rep for Sub-Project 2.4*, CRIM-IIT-95/09-35, http://nannetta.ce.unipr.it/argo/theysay/rs2/#publications.

[6] M.R. Luo, R.W.G. Hunt, The Structure of the CIE 1997 Colour Appearance Model (CIECAM97s), *Color Res. Appl.*, 1998, 23, 138-146.

[7] W. Osberger and Maeder A.J. Automatic Identification of Perceptually Important Regions in an Image. *Proc. 14th Int. Conf. on Pat. Rec.*, Brisbane, Australia, 1998, 701-704.

[8] A. Rybak, V.I. Gusakova, A.V. Golovan, L.N. Podladchikova, N.A. Shevtsova I. A Model of Attention-Guided Visual Perception and Recognition. *Vision Research*, 1998, 38, 2387-2400.

[9] G. Salgian, D. Ballard. Visual Routines for Vehicle Control. In D. Kriegman, G. Hager, and S. Morse, editors, *The Confluence of Vision and Control*, Springer Verlag, 1998.

[10] E.L. Schwartz. Computational Anatomy and Functional Architecture of Striate Cortex: A Spatial Mapping Approach to Perceptual Coding. *Brain Research*, 1980, 20, 645-669.

[11] CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), *CIECAM97s*, CIE TC1-34, April 1998.

[12] A.L.Yarbus. *Eye Movements and Vision*, Plenum, New York, 1967.

[13] M. M. Zadeh, T. Kasvand, C. Y. Suen. Localization and Recognition of Traffic Signs for Automated Vehicle Control Systems. In *Conf. on Intelligent Transportation Systems, part of SPIE's Intelligent Systems & Automated Manufacturing,* Pittsburgh, PA, October 1997.