Vision Models Based Identification of Traffic Signs

X. Gao, N. Shevtsova*, K. Hong, S. Batty, L. Podladchikova*, A. Golovan*, D. Shaposhnikov*, and V. Gusakova*
School of Computing Science, Middlesex University, Bounds Green London, United Kingdom
**** Laboratory of Neuroinformatics of Sensory and Motor Systems, A.B. Kogan Research Institute for Neurocybernetics, Rostove State University, Rostov-on-Don, Russia***

Abstract

During the last 10 years, computer hardware technology has been improved rapidly. Large memory, storage is no longer a problem. Therefore some trade-off (dirty and quick algorithms) for traffic sign recognition between accuracy and speed should be improved. In this study, a new approach has been developed for accurate and fast recognition of traffic signs based on human vision models. It applies colour appearance model CIECAM97s to segment traffic signs from the rest of scenes. A Behavioural Model of Vision (BMV) is then utilised to identify the signs after segmented images are converted into grey-level representation. Two standard traffic sign databases are established. One is British traffic signs and the other is Russian traffic signs. Preliminary results show that around 90% signs taken from the British road with various viewing conditions have been correctly identified.

1. Introduction

Identification of traffic signs correctly at the right time and the right place is very important for car drivers to insure themselves and their passengers' safe journey. However, sometimes, due to the change of weather conditions or viewing angles, traffic signs are difficult to be seen until it is too late. Development of an automatic system to be implemented inside cars for recognition of traffic signs will certainly improve driving safety a great deal.

An automatic real time system requires the identification of traffic signs invariant with respect to various transformations of signs and viewing environment. In particular, in view of the extremely stringent safety requirements for routine use of approaches on public roads, more computing power than was available a few years ago and more robust algorithms will be required in order to provide the necessary accuracy in recognition of traffic signs.

There are broadly 3 major methods applied for traffic sign recognition. They are colour-based, shape-based, and neural network-based recognition. Colour is a dominant visual feature and undoubtedly represents a piece of key information for drivers to handle. Colour is regulated not only for the traffic sign category (red = stop, yellow = danger, etc.) but also for the tint of the paint that covers the sign, which should correspond, with a tolerance, to a specific wavelength in the visible spectrum.[1] Therefore it is widely used in the systems for traffic sign recognition,[1] especially for segmentation of traffic sign images from the rest of a scene. The eight colours, red, orange, yellow, green, blue, violet, brown and achromatic, are the most discriminating colours for traffic signs.

Most colour-based techniques in computer vision run into problems if the illumination source varies not only in intensity but also in colour as well. This is the main reason why many researchers have tried to come up with algorithms for separating the incident illumination from the colour signal perceived by the sensors. So that after this kind of separation, a sign becomes illumination-invariant and is full of characteristics of the surface that reflects the light. As the spectral composition, and therefore the colour, of daylight is known to change depending on weather conditions, e.g., sky with/without clouds, time of day, and night when all sorts of artificial lights are surrounded,[2] no method has been widely accepted yet.

In this study, traffic signs are segmented from real world road scenes based on colour contents using a standard colour appearance model CIECAM97 recommended by the CIE (International Committee on Illumination)[3,4] and are identified after segmentation by the application of a Behavioural Model of Vision (BMV).[5,6]

2. Methodology

In this study, two standard databases have been established. One is based on British traffic signs (n=142) scanned in 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.

2.1 Segmentation

The first step to process the images (still images for the time being) taken by a video camera is to segment the sub-images of traffic-sign-to-be from the rest of scenes. To achieve this, 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 corresponding values using the CIECAM97 model.

When an image is downloaded to a computer, it is presented in a RGB space. To convert RGB space to CIE standard XYZ space, the following equations are applied as shown in Eq (1) for average daylight with CIE standard illuminant D65 as reference white, i.e., $[X_w, Y_w, Z_w] = [0.95045\ 1.0\ 1.088754]$.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

Then from CIE XYZ space, the hue, chroma, and lightness are obtained using CIECAM97 model. The reason to apply this model is that it takes viewing condition into account and can predict colours as accurate as an average observer.

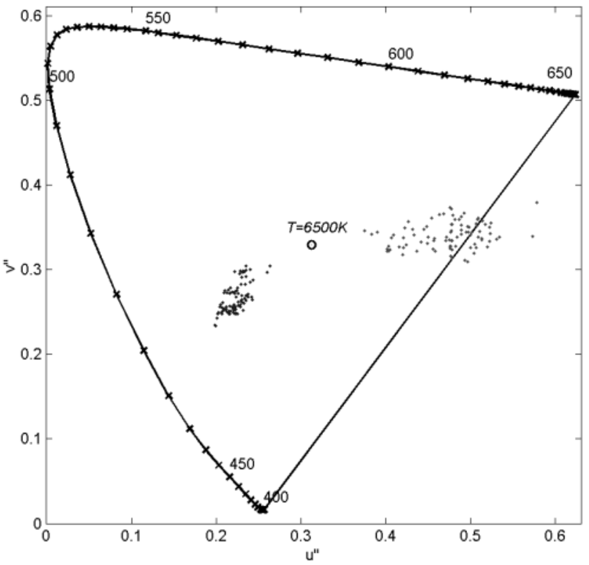


Figure. 1 The ranges of blue and red signs are plotted on a u'v' diagram.

During the study, 83 images of British road signs have been taken with variety of viewing and weather conditions using a digital camera, Olympus Digital Camera C-3030. These images are then classified visually according to the viewing conditions, such as cloudy, sunny, etc. Based on the images in each group, the parameters for each viewing condition are found out from Ref. 3 (e.g., direct sun light with colour temperature 5335K and light from overcast sky with colour temperature 6500K) for the application of the colour appearance model. Images taken under real viewing conditions are then transformed from RGB space (the format used in computers to represent an image) to CIE XYZ values and then to LCH (Lightness, Chroma, Hue) using the model of CIECAM97. The range for red sign is between 393-423 that is calculated using mean ± standard deviation, and is between 280 to 290 for blue hue respectively. While for chroma, the range is between 57 to 95. The background also has chroma value ranged between 7 to 50. Figure 1 illustrates the range of blue and red signs plotted on a u'v' chromaticity diagram.

2.2 Recognition

The second step to recognise a sign is to process segmented sub-images. The identification of traffic signs in this study is carried out by the application of the BMV.[5,6] This model is developed on the base of biologically plausible algorithms and has demonstrated the ability to recognise complex grey-level images invariantly with respect to shift, plain rotation, and in a certain extent to scale.

To apply the model to the traffic sign identification task, the traffic sign database was transformed into a model-specific form. Firstly, all coloured images from the database were converted into grey-level representation. Then, for each image in the database a specific description was obtained based on trajectory of its viewing according to the most informative regions of the sign.[5] The model provides a compressed and invariant representation of each image fragment along the trajectory of viewing by space-variant features extracted in the fragment by Attention Window (AW). These descriptions have been stored with the images and form a model-specific database of traffic sign images. The model-specific database for traffic signs 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.

To extract features of signs, or to describe a sign at both memorising (for database images) and recognition (for real world images) stages, a specific structure of the AW is used. The AW simulates some mechanisms of the biological visual systems, such as space-variant representation of information from the centre (the fovea) to the periphery of the retina,[7,8] neuronal orientation selectivity,[9] and context encoding of the foveal information.[8]

Similar to Ref. 6, the following steps are applied to obtain a model-specific description of an traffic sign image.

(i) An image fragment in each viewing trajectory point is presented by 49-dimensional vector of orientations extracted in vicinity of each of 49 nodes of AW (Fig. 2,a) the centre of which is located in a given point of the viewing trajectory;
(ii) the AW nodes are located at the intersections of sixteen radiating lines and three concentric circles, each with a different radius;
(iii) orientation of segments in the vicinity of each AW node is determined by means of Gaussian with spatially shifted centres with the step of 22.5°;
(iv) space-variant representation of image features is emulated by Gaussian convolution with different kernels increasing with the distance from the AW centre.

a) b)

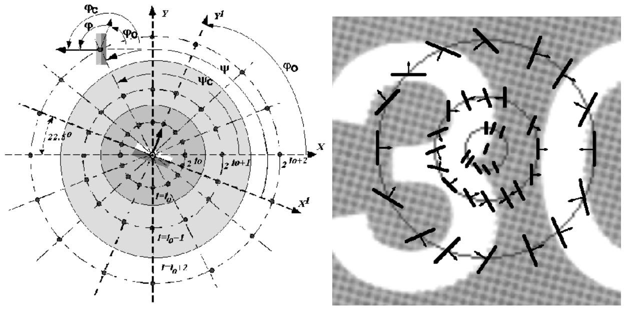


Figure 2. (a) Schematic of the attention window (AW); the areas of different resolution are separated by shadings. The AW context nodes are located at the intersections of sixteen radiating lines and three concentric circles, each in a different resolution area. X0Y is the absolute coordinate system. The relative coordinate system X^1OY^1 is attached to the edge in the centre of the AW; (b) example of oriented elements detected by the AW in initial point of the viewing trajectory of a traffic sign image.

a) b) c)

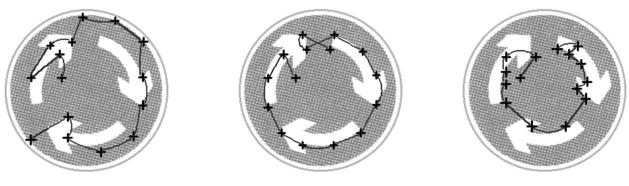


Figure 3. Examples of sign viewing trajectories formed by different algorithms successive fixation points are chosen (a) by the standard method as in [6]; (b) by estimation of filling-in of 49-dimentional vector; (c) by determination of external contour location.

Such procedure allows to obtain image fragment description in each fixation of the AW that may be relatively stable to any local image disturbances in some range. Example of oriented elements detected in a sign fragment is shown in Fig. 2, b. The whole image is presented by a set of fragment descriptions along a trajectory of image viewing. Each next AW fixation point is chosen among the context nodes of the previous fixation of the AW. A transition from a given fixation point to the next one forms a trajectory of viewing and altogether with fragment descriptions provides a model-specific description of an image.

The pecularities of the traffic sign recognition task demand some modifications of the model. In particular, contrary to the original model version where transition description from a fixation point to the next one was realised in a relative coordinate system (see Fig. 2a) to provide rotation invariance, the current model version uses the absolute coordinate system. This was made to exclude invariance to rotation because the same objects in signs just rotated in plain have different meaning for driver (such as arrows). Besides, some modifications were developed in algorithms of image viewing trajectory formation for preferable representation of inner informative parts of signs, which is believed to be the most attractive for driver's attention. Three algorithms for formation of image viewing trajectories were developed and compared to each other while sign processing. Fig. 3 illustrates that maximal attraction of the AW to the inner sign parts during image viewing is provided by algorithm based on preliminary estimation of location of external image contour elements. This algorithm was used for description of signs from standard databases at memorizing stage.

At recognition stage, a model-specific description of a real world sign was compared with those from the standard databases similar to Ref. 6.

3. Results

After a coloured image is segmented using the CIECAM97 colour appearance model, it is firstly 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 of traffic signs. If a successful match occurs the presented image is recognised, and the matched sign image is retrieved.

Initial experimental results show that the majority signs can be segmented correctly by using CIECAM97 colour vision model, up to 90% for sunny days.

After segmentation, the BMV model, correctly identified 37 out of 41 potential traffic sign images for sunny weather conditions and 37 out of 42 for cloudy weather conditions, which gives 90% and 88% success rates respectively. 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. Recognition time is varied from 0.35 seconds up to 0.6 seconds per image on a standard Pentium III, 400Hz.

Figure 4 illustrates an example of segmentation. The third segmented image (with the rear light from the left-hand-side car) is not traffic sign but it has the colour range characteristic for them. Although it is segmented at this stage, it will be recognised as a non-sign image during the recognition stage.

Figure 5 shows some recognition results using the BMV. It can be seen that this model works very well for some images with various disturbances (as shown in (a) and (b)).

4. Discussion

Overall the models-based approach can give accurate identification for traffic signs located in 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 use of the CIECAM97 colour vision model allows the segmentation of the majority of traffic signs from the rest of the scenes.

Computer experiments with the BMV indicate that a preliminary separation of traffic signs by shape for each colour (for example, rectangle versus circle for blue traffic sings or triangle versus ring/circle for red ones) can accelerate sign identification which will be implemented in future model modification.

In addition, experimental results demonstrate the importance of AW fixation points chosen while viewing trajectory formation. Also the adequate template image encoding indicates that it is necessary that psychophysical experiments should be conducted to achieve better understanding what attracts a driver's visual attention while driving along the road in order to find the most informative regions in traffic sign images. 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.

Figure 4. An example of image segmentation using CIECAM97 model. The third segmented image is not a traffic sign that will be recognised using the BMV.

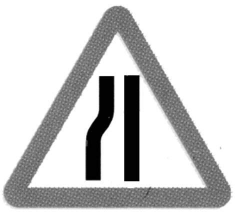
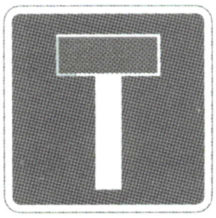
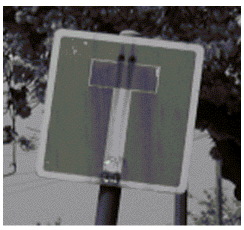
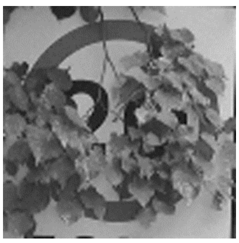
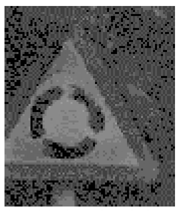
a)	**b)**	**c)**	**d)**	
		Non identified image	Non identified image	Standard signs stored in the database
				Sign images taken from real road scene

Figure 5. Some retrieval results of recognition using BMV model. In (a) and (b)

Acknowledgement

This project is founded by The Royal Society, UK, under the International Scientific Exchange Scheme. Their support is gratefully acknowledged.

References

1. M. Lalonde and Y. Li, Road Sign Recognition -- Survey of the State of Art, technique report for Sub-Project 2.4, CRIM-IIT-95/09-35, http://nannetta.ce.unipr.it/argo/theysay/rs2/#publications.
2. D. Judd, D. MacAdam, and G.. Wyszecki, Spectral distribution of typical daylight as a function of correlated color temperature, J. Opt. Soc. Am., 54(8):1031-1040 (1964).
3. M.R. Luo and R.W.G. Hunt, The structure of the CIE 1997 colour appearance model (CIECAM97s), Color Res. Appl., 23:138-146 (1998).
4. CIE, The CIE 1997 Interim Colour Appearance Model (Simple Version), CIECAM97s, CIE TC1-34, April, (1998).
5. L.N. Podladchikova, V.I. Gusakova, D.G. Shaposhnikov, A. Faure, A.V. Golovan, and N.A. Shevtsova, MARR: Active vision model, Proceedings SPIE'97, 3208:418-425 (1997).
6. I.A., Rybak, V.I., Gusakova, A.V., Golovan, L.N., Podladchikova, N.A., Shevtsova, A Model of Attention-guided Visual Perception and Recognition, Vision Research, 38:2387-2400 (1998).
7. E.L. Schwartz. Computational anatomy and functional architecture of striate cortex: A spatial mapping approach to perceptual coding. Brain Research, 20, 645-669 (1980).
8. A.L.Yarbus. Eye Movements and Vision. 1967, New York: Plenum.
9. D.H. Hubel. & T.N. Wiesel. Sequence, regularity and geometry of orientation columns in the monkey striate cortex. J. Comparative Neurology, 158, 267-293 (1974).

Biography

Xiaohong Gao received her B.S. degree in Applied Mathematics from Liaoning University in China in 1984 and a PhD in Modelling of Colour Appearance from Loughborough University at U.K. in 1994. Her current research interests include colour imaging, image indexing and retrieval, and medical imaging.