14

Traffic Sign Detection and Recognition

Hasan Fleyeh

Dalarna University, Falun, Sweden

14.1 Introduction

Traffic signs define a visual language that can be interpreted by drivers. They represent the current traffic situation on the road, show danger and difficulties around the drivers, give them warnings and help them with their navigation by providing useful information that makes driving safe and convenient [1, 2].

The human visual perception abilities depend on the individual's physical and mental conditions. In certain circumstances, these abilities can be affected by many factors such as fatigue and observatory skills. Giving this information in a good time to drivers can prevent accidents, save lives, increase driving performance and reduce the pollution caused by vehicles [3–5].

Traffic sign recognition is a field which is concerned with the detection and recognition of traffic signs in traffic scenes acquired by a camera. It is a technique which uses computer vision and artificial intelligence to extract the traffic signs from outdoor images taken in uncontrolled lighting conditions where these signs may be occluded by other objects, and may suffer from different problems such as colour fading, disorientation and variations in shape and size. It is the field of study that can be used to aid the development of an inventory system (for which real-time recognition is not required) or to aid the development of an in-car advisory system (when real-time recognition is necessary). Both road sign inventory and road sign recognition are concerned with traffic signs, face similar challenges and use automatic detection and recognition.

A traffic sign recognition system could in principle be developed as part of intelligent transport systems (ITS) that continuously monitors the driver, the vehicle and the road, for example, to inform the driver in time about upcoming decision points regarding navigation and potentially risky traffic situations. Detection of these signs in outdoor images from a moving vehicle will help the driver to take the right decision at the right time, which means fewer accidents, less pollution and better safety.

Computer Vision and Imaging in Intelligent Transportation Systems, First Edition.
Edited by Robert P. Loce, Raja Bala and Mohan Trivedi.

Companion website: www.wiley.com/go/loce/ComputerVisionandImaginginITS

14.2 Traffic Signs

Traffic signs are those that use a visual/symbolic language about the road(s) ahead that can be interpreted by drivers. They provide the driver with pieces of information that make driving safe and convenient. Road signs are designed, manufactured and installed according to strict regulations [6]. However, they can appear in different conditions, including partly occluded, distorted, damaged and clustered in a group of more than one sign [7, 8].

Traffic signs are characterised by a number of features which ideally make them recognisable with respect to the environment, but certain factors can also affect a driver's perception. They are, but not limited to, as follows:

- They are designed in fixed two-dimensional (2D) shapes such as triangles, circles, octagons or rectangles [7, 9].
- The colours of the signs are chosen to contrast with the surroundings, which make them easily recognisable by drivers [10].
- The colours are regulated by the sign category [11].
- The information on the sign has one colour and the rest of the sign has another colour.
- The tint of the paint which covers the sign should correspond to a specific wavelength in the visible spectrum [6, 8].
- The signs are located in well-defined locations with respect to the road, so that the driver can, more or less, anticipate the location of these signs [11].
- They may contain a pictogram, a string of characters or both [8].
- In every country the traffic signs are characterised by using fixed text fonts and character heights.

14.2.1 The European Road and Traffic Signs

According to 1968 Vienna Convention on Road Signs and Signals, the European traffic signs are categorised into four groups:

- Warning signs: This group of traffic signs indicates a hazard ahead on the road. It is characterised by an equilateral triangle with a thick red rim and a white or yellow interior. A pictogram is used to specify different warnings. Other signs such as the Yield sign, the distance to level crossing signs and track level crossing also belong to this class.
- Prohibitory signs: They are used to prohibit certain types of manoeuvres for some types of traffic. The no entry, no parking and speed limit signs belong to this category. Normally, they are designed in a circular shape with a thick red rim and a white or yellow interior. There are few exceptions; the Stop sign is an octagon with a red background and white rim, the No Parking and No Standing signs have a blue background. The end-of-restriction signs are marked with black bars.
- Mandatory signs: They are characterised by a complete blue circle and a white arrow or pictogram. They control the actions of drivers and road users. Signs ending obligation have a diagonal red slash.
- Indicatory and supplementary signs: These signs are characterised by using rectangles with different background colours such as yellow, green or blue. The pictograms are either white or black. This category includes the diamond-shaped rectangle and the signs which give information about road priority.

The colours used on road signs have specific wavelengths in the visible spectrum. They are selected to be distinguishable from the natural and man-made surroundings so that they can be easily recognisable by road users. The Swedish traffic signs are illustrated in Figure 14.1.

(a)

(b)

Figure 14.1 The Swedish traffic signs. (a) Warning signs, (b) prohibitory signs, (c) mandatory signs and (d–f) indicatory and supplementary signs.

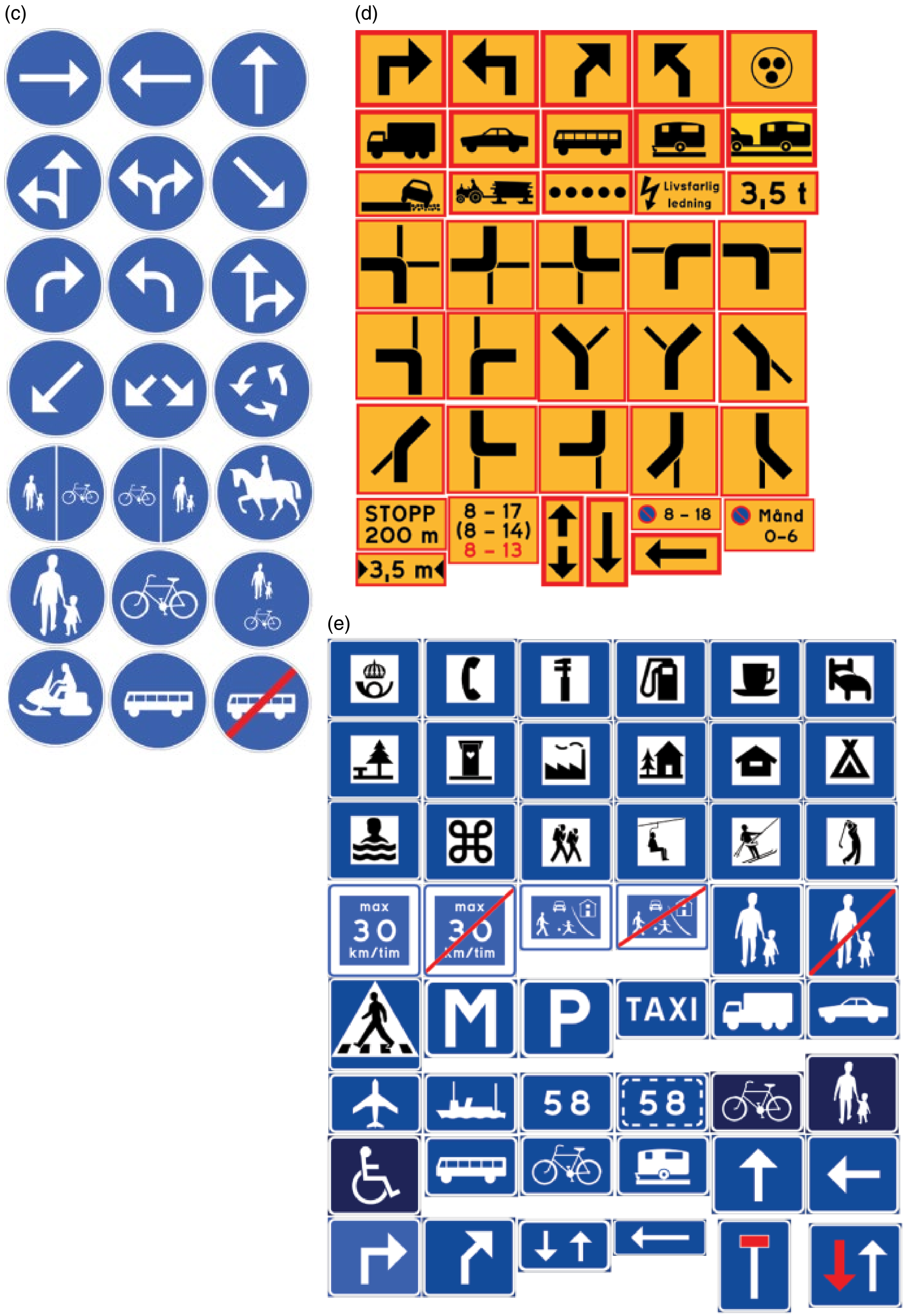


Figure 14.1 (Continued)

(f)

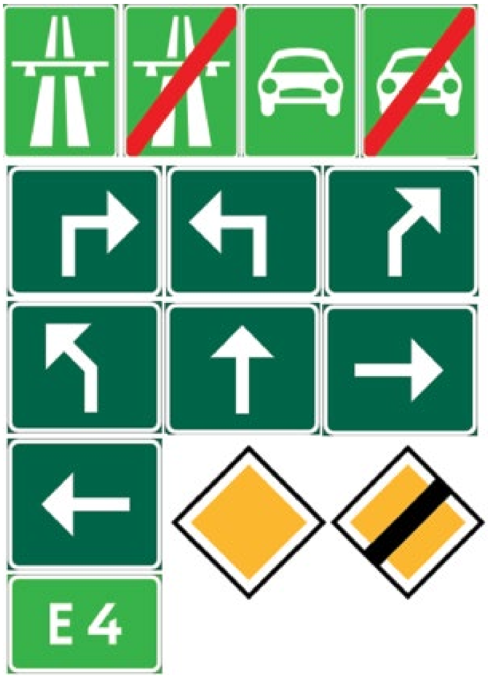


Figure 14.1 (Continued)

14.2.2 The American Road and Traffic Signs

American traffic signs follow the *Manual on Uniform Traffic Control Devices* and its companion 'Standard Highway Signs'. They did not follow Vienna Convention on Traffic Signs and Signals, and they are divided into eight categories: Regulator, Schools, Warning, Guide, Toll Road Signs, Hospital, Non-Compliant to MUTCD Signs and References. Regulatory and Warning signs are further divided into several subgroups called Series. These groups are completely different than that of European traffic signs. Yield and Stop signs are grouped in R1 Series, for instance, and speed limit signs are characterised by white background and black text [12].

14.3 Traffic Sign Recognition

The first paper on the subject was published in Japan in 1984 [13]. The aim was to try various computer vision methods for the detection of road signs in outdoor scenes. Since that time many research groups and companies have shown interest, conducted research in the field and generated an enormous amount of work. Different techniques have been used to cover different application areas, and vast improvements have been achieved during the past decade.

The identification of the traffic signs is achieved through two main stages: detection and recognition. In the detection phase, the image is preprocessed, enhanced and segmented according to the sign properties such as colour or shape or both. The output is a segmented image containing potential regions which could be recognised as possible road signs. The efficiency and speed of the detection are important factors because they reduce the search space and indicate only potential regions. In the recognition stage, each of the candidates is tested against a certain set of features (patterns) to decide whether it is in the group of traffic signs or not, and then according to these features they are classified into different groups. These features are chosen so as to emphasise the differences among the classes. The shape of the sign plays a central role in this stage and the signs are classified into different classes such as triangles, circles and octagons. Pictogram analysis allows a further stage of classification. By analysing pictogram shapes together with the text available in the interior of the sign, it is easy to decide the individual class of the sign under consideration. The system can be implemented by either colour information or shape information or both. Combining colour and shape may give better results if the two features are available, but many studies have shown that detection and recognition can be achieved even if one component, either colour or shape, is missing.

14.4 Traffic Sign Recognition Applications

Techniques for traffic sign detection and recognition have been developed in a range of application areas. These include the following:

- Driver support system (DSS): This system can detect and recognise traffic signs in real time. The system can also help to improve traffic flow and safety [14, 15], and avoid hazardous driving conditions, such as collisions. Until the past decade, there were very small number of studies about traffic sign detection and classification. Research groups focused on other aspects of sign detection, more related to the development of an automatic pilot, such as the detection of the road borders and/or the recognition of obstacles in the vehicle's path, for example, other vehicles or pedestrians. Intelligent vehicles such as Volvo S, V and XC series introduced systems which could take decisions about their speed, trajectory, etc., depending on the signs detected [16, 17]. Although, in the future, it can be part of a fully automated vehicle, now it can be a support to automatically limit the speed of the vehicle, send a warning signal indicating excess speed, warn or limit illegal manoeuvres or indicate earlier the presence of a sign to the driver. The general idea is to support the driver in some tasks, allowing him or her to concentrate on driving.
- Highway maintenance: This is used to check the presence and condition of the signs. Instead of an operator watching a video, which is a tedious work because the signs appear from time to time and the operator should pay a great attention to find the damaged ones, the traffic sign detection and recognition system can do this job automatically for the signs with good conditions and alerts the operator when the sign is located but not classified which can be interpreted as a sign in poor condition.
- Sign inventory: The many millions of roadway signs necessary to keep roadways safe and traffic flowing present a particular logistical challenge for those responsible for the installation and maintenance of those signs. Traffic signs must be properly installed in the necessary locations and an inventory of those signs must be maintained for future reference.
- Mobile robots: Road and traffic signs can be used to automatically mobilise robots (unmanned vehicles) depending on the detection and recognition of these landmarks by the robot [14]. Mei et al. developed an unmanned vehicle which has the ability to navigate in urban environments without GPS or any other satellite navigation system. The vehicle navigates solely by perceiving traffic signs [18].

14.5 Potential Challenges

In addition to the complex environment of the roads and the scenes around them, traffic signs can be found in different conditions such as aged, damaged and disoriented as depicted in Figure 14.2. Hence, the detection and recognition of these signs may face one or more of the following difficulties:

- The colour of the sign fades with time as a result of long exposure to sun light, and the reaction of the paint with the air [2, 5].
- Visibility is affected by the weather conditions such as the fog, rain, clouds and snow [2].
- Visibility can be affected by local light variations such as the direction of the light, the strength of the light depending on the time of the day and the season and the shadows generated by other objects [8, 19, 20].
- Colour information is very sensitive to the variations of the light conditions such as shadows, clouds and the sun [2, 5, 21]. It can be affected by illuminant colour (daylight), illumination geometry and viewing geometry [22].
- The presence of obstacles in the scene, such trees, buildings, vehicles and pedestrians or even signs which occlude other signs [19, 21].
- The presence of objects similar in colour and/or shape to the road signs in the scene under consideration, such as buildings or vehicles [5, 19]. They could be similar to the road sign in colour or shape or both.
- Signs may be found disoriented, damaged or occluded by any kind of obstacles, even by some other signs.
- The size of the sign depends on the distance between the camera and the sign itself. Traffic signs may appear rotated due to the imaging orientation [23].
- The acquired image often suffers from motion blur and car vibration [24]. This motion blur cannot be predicted above a certain level, because the car movements are not known to the recognition process. It is possible to make an assertion about the movements of objects in the future if the motion is continuous and unchanged.
- Sign boards can appear to have bright white or near-white spots due to first surface reflection from the light sources. In first surface reflection the light is reflected prior to penetrating to a depth where certain wavelengths are absorbed, thereby imparting a colour associated with the sign. This is called 'highlight'.
- Vandalism of sign boards by people who put stickers or write on them or damage the pictograms by changing the pictogram shape.
- Different countries use different colours (Figure 14.3) and different pictograms (Figure 14.4).
- The absence of a standard database for evaluation of the existent classification methods [25].

It is extremely important for the algorithms to be developed for the detection and recognition of road and traffic signs to have high robustness of colour segmentation, high insensitivity to noise and brightness variations, and should be invariant to geometrical effects such as translation, in-plane and out-plane rotations and scaling changes in the image [26, 27].

14.6 Traffic Sign Recognition System Design

A system to detect and recognise traffic signs should be able to work in two modes: the training mode in which a database can be built by collecting a set of traffic signs for training and validation, and the prediction mode in which the system can recognise a traffic sign which has not been seen before. Figure 14.5 illustrates the main stages to recognise a traffic sign.

Figure 14.2 Potential challenges when working with traffic signs. (a) Faded sign, (b) bad weather condition, (c) bad lighting geometry, (d) obstacles in the scene, (e) similar background colour, (f) damaged sign, (g) distance related size, (h) motion blur, (i) reflection from sign board and (j) stickers.

Figure 14.2 (Continued)

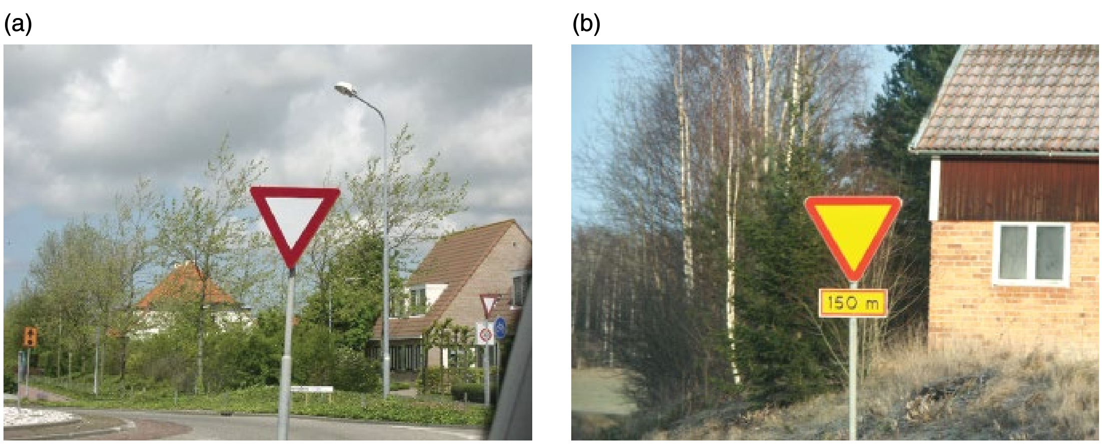


Figure 14.3 Different countries have different colour standard. (a) The Netherlands and (b) Sweden.

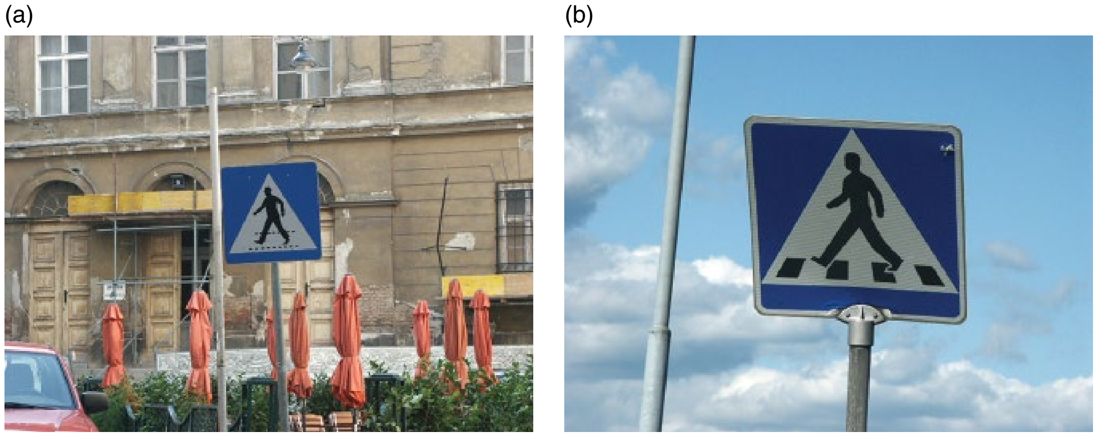


Figure 14.4 Traffic sign pictograms are different in different countries. (a) Austria and (b) Sweden.

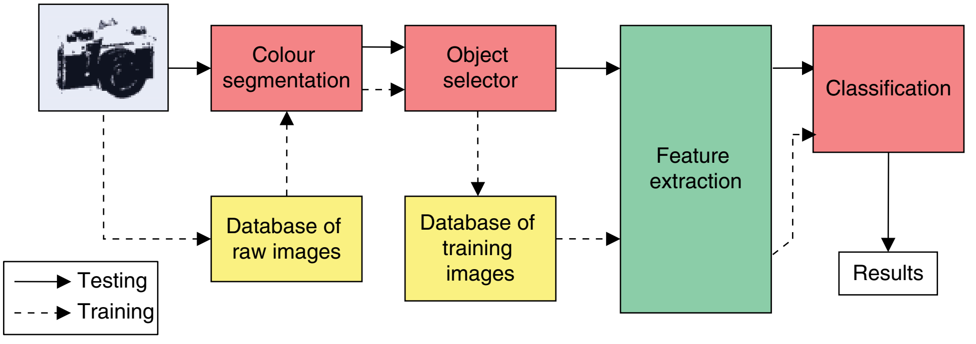


Figure 14.5 Block diagram of traffic sign recognition system.

Recognition and classification of traffic signs can be achieved by combining the two main traffic sign features: colour and shape. This method helps the recognition algorithm to perform in a better way and to reduce the number of false alarms generated by this algorithm. Therefore, the detection and recognition of different signs requires testing the presence of different colour combinations in the image together with the presence of the specific shape. Therefore, recognition and classification is carried out by two stages. In the first stage colour segmentation is applied. Two rim colours exist for traffic signs in Sweden: red and blue. A traffic sign shape tree is built according to these two colours as depicted in Figure 14.6. Based on the colour of the traffic sign's rim, the type of the traffic sign is specified by combing the shape of the rim with its pictogram. This requires two different shape analysis stages: the shape of the rim and the analysis of the pictogram. Very often, this shape analysis is achieved by training two classifiers to classify the rim and the pictogram.

14.6.1 Traffic Signs Datasets

There are three main public traffic sign datasets in Europe; among them are two Swedish datasets and one German. The datasets were collected for research purposes and create a standard benchmark for developing and testing algorithms for traffic sign recognition.

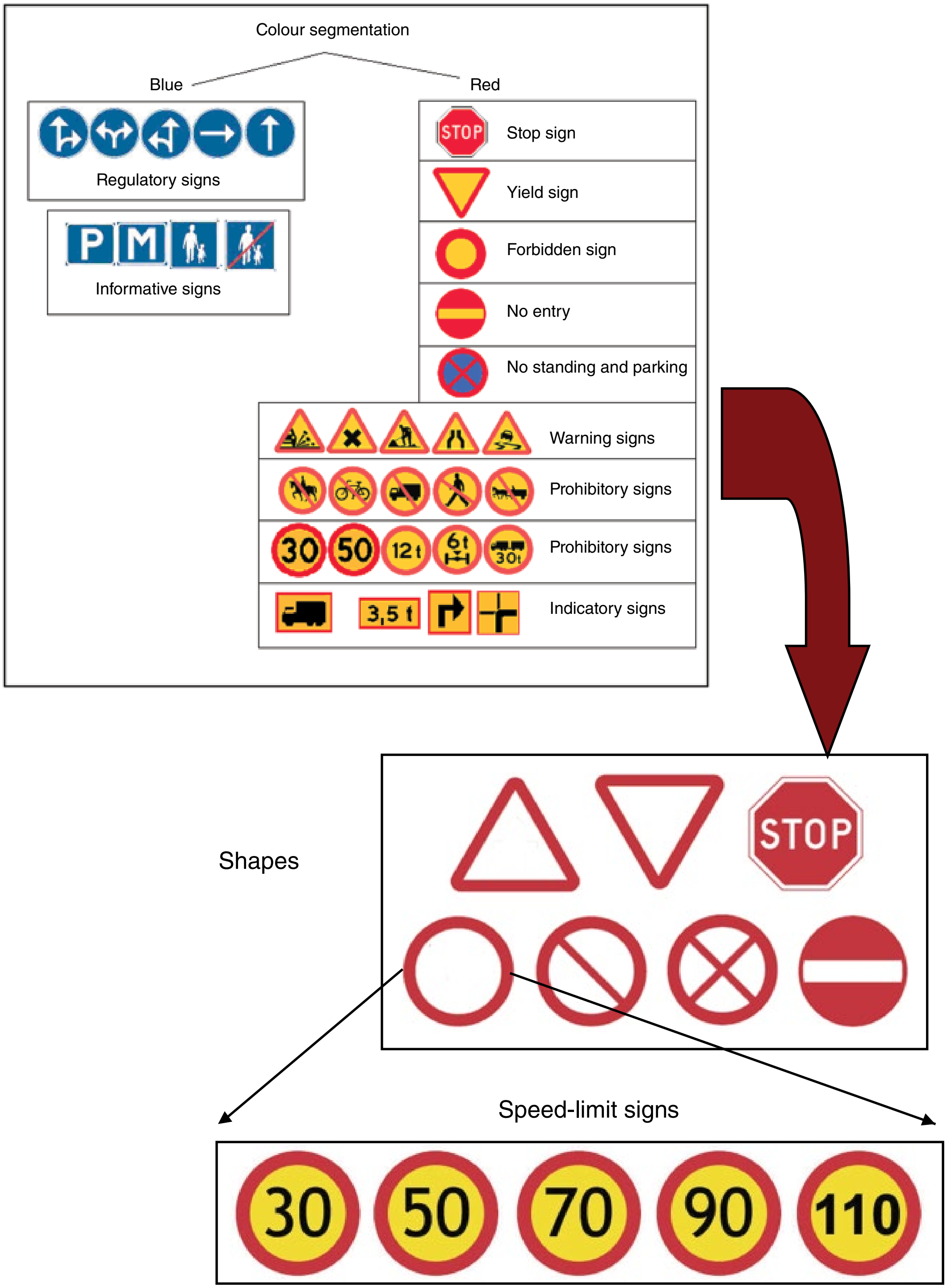


Figure 14.6 Traffic sign tree based on colour and shape information.

Dalarna University in Sweden released a traffic sign dataset which consists of 4338 image collected in Sweden and 330 images collected in other counties [28]. All still images were taken manually when traffic signs were seen by the camera operator. They were collected in different light conditions, in different weather conditions and in different road conditions including different speeds. For all images and without any exception, the camera was set to 640 × 480 pixels. Images in this database are classified into 30 categories depending on weather conditions, type of the sign, sign condition, image condition and light geometry.

Ruhr-Universität Bochum in Germany released the German Traffic Sign Benchmark dataset which is a multiclass, single-image in conjunction with the International Joint Conference on Neural Networks (IJCNN) 2011. The benchmark traffic sign dataset contains 40 traffic sign classes. It consists of more than 50,000 images [29].

Linköping University in Sweden released a public traffic signs dataset in conjunction with Scandinavian Conference on Image Analysis (SCIA) 2011 conference. The dataset contains 20,000 images with 20% of the images are labelled. It contains 3488 traffic signs which were collected from highways and cities from more than 350 km of Swedish roads [30].

14.6.2 Colour Segmentation

Colours represent an important part of the information provided to the driver to ensure the objectives of the traffic sign. Therefore, traffic signs and their colours are selected to be different from the nature or from the surrounding in order to be distinguishable. One of the important steps in Traffic Sign Recognition (TSR) is segmentation. It is the process by which candidate objects are specified for further analysis according to certain properties such as colour or shape. Colour segmentation has always been considered a strong tool for image segmentation because it is computationally inexpensive. Due to the fact that colour gives more information than grey it is used in segmentation algorithms instead of edge-based and luminance histogram-based techniques.

In order to develop a robust colour segmentation algorithm, it is necessary to understand the nature of the colour and the circumstances under which the traffic scene image is collected. The first problem for any colour segmentation algorithm is that the apparent colour of the object varies because of the chromatic variation of daylight. The second problem is that the colours of traffic sign boards change with age, which means newly installed traffic signs have different colours than older traffic signs. In addition to this, different countries use different standard colours for traffic signs. Finally, traffic sign images may suffer from the effect of shadows and highlights. The effect of shadows occurs when different parts of the object are exposed to different illumination levels, while in the case of highlight the object reflects some of the light of the illuminant directly to the viewer. These problems are discussed in the following text in detail.

1) *Colour variations in outdoor images*

   One of the most difficult problems in using colours in outdoor images is the chromatic variation of daylight which causes the apparent colour of the object to vary as daylight changes. The irradiance of any object in a colour image depends on the following three parameters:

   The colour of the incident light: Daylight's colour varies along the CIE curve. It is given by

$$y = 2.87x - 3.0x^2 - 0.275 \text{ for } 0.25 \le x \le 0.38 \tag{14.1}$$

   The variation of daylight's colour is a single variable which is independent of the intensity.

   The reflectance properties of the object: The reflectance of an object $s(\lambda)$ is a function of the wavelength λ of the incident light. It is given by

$$s(\lambda) = e(\lambda)\phi(\lambda) \tag{14.2}$$

where $e(\lambda)$ is the intensity of the light at wavelength λ and $\phi(\lambda)$ is the object's albedo at each wavelength.

The camera properties: The observed intensities depend on the lens diameter d, its focal length f and the image position of the object measured as angle a off the optical axis. This is given by

$$E(\lambda) = L(\lambda) \cdot \left(\frac{\pi}{4}\right)\left(\frac{d}{f}\right)^2 \cos(4a) \tag{14.3}$$

According to Equation 14.3, the radiance $L(\lambda)$ is multiplied by a constant which will not affect an object's observed colour. By cancelling the camera's lens chromatic aberration, only the density of the observed light will be affected.

As a result, the colour of the light reflected by an object located outdoors is a function of the temperature of the daylight and the object's albedo [22, 31].

2) *Aging of traffic signs*

Traffic signs may be mounted on a pole for a long time without any kind of maintenance or replacement. Over time the properties of the material used to reflect the light in these traffic signs changes its properties because of environmental reactions and so its colour fades. Figure 14.7 depicts two traffic signs; the sign on the left is a new one, while that on the right is an old one.

To investigate the effect of aging, a set of equal number of old and new traffic signs was selected. Colour located in the red part of each traffic sign was extracted and converted from RGB (red, green, blue) into HSV (hue, saturation, value). HSV was selected because it is invariant to the variations in light conditions as it is multiplicative/scale invariant, additive/shift invariant, and it is invariant under saturation changes. In addition, it has been proven by Gevers and Smeulders [32] that Hue is invariant against shadow and highlights.

Vitabile et al. [8] defined three different areas in the HSV colour space as follows:

1) The *achromatic* area, characterised by $s \le 0.25$ or $v \le 0.2$ or $v \ge 0.9$.
2) The *unstable chromatic* area, characterised by $0.25 \le s \le 0.5$ and $0.2 \le v \le 0.9$.
3) The *chromatic* area, characterised by $s \ge 0.5$ and $0.2 \le v \le 0.9$.

(a) (b)

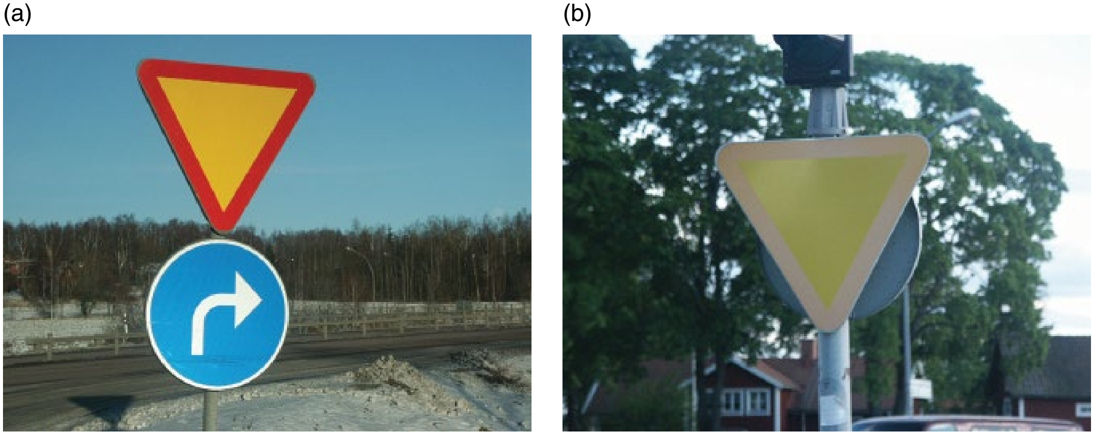


Figure 14.7 Effect of aging. (a) A new traffic sign. (b) An old traffic sign.

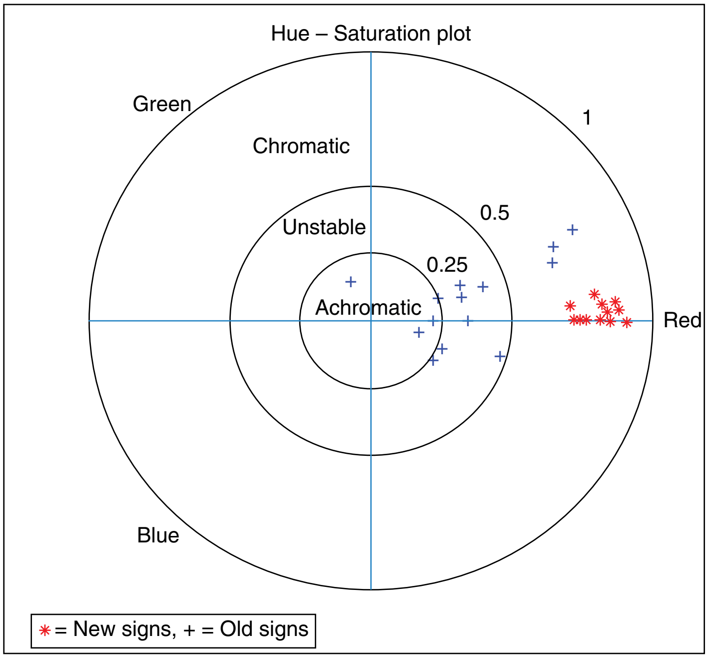


Figure 14.8 Hue–Saturation plot of new and old traffic signs.

Plotting the Hue and Saturation of each traffic sign in the Hue–Saturation plot is shown in Figure 14.8. The new traffic signs were grouped in the chromatic region, while old traffic signs shifted either towards the yellow part of the Hue or towards the unstable area or even towards the achromatic area of this plot. Old traffic signs which moved toward the yellow area but still in the chromatic area can be colour segmented, while the colour of the remaining traffic signs cannot be segmented because of the loss of Hue.

3) *Traffic signs from different countries*
Different countries use different colours for their traffic signs. Figure 14.3 shows two yield traffic signs from the Netherlands (left) and Sweden (right). To understand the distribution of traffic signs from different countries on the Hue–Saturation plot, a number of traffic signs were collected from a number of countries in Europe, the United States and Japan, Figure 14.9. Plotting the red colour on the Hue–Saturation plot clearly indicates that the red colour of these traffic signs is not similar and those countries do not follow any standard colour.

4) *Robustness of colour spaces against shadows and highlights*
Shadows and highlights represent a big challenge to computer vision researchers. In the first case different parts of the object are exposed to different illumination levels, and in the second case the object reflects some of the light of the illuminant directly to the viewer.

Consider an image of a certain surface patch which is illuminated by an incident light with a certain spectral power density (SPD) denoted $e(\lambda)$.This image is taken by a camera with RGB sensors characterised by their spectral sensitivities $f_C(\lambda)$ for $C = \{R,G,B\}$. The Cth sensor response of the camera is given by

$$C = m_b\left(\mathbf{n},\mathbf{s}\right)\int_\lambda f_C(\lambda)e(\lambda)c_b(\lambda)d\lambda + m_s\left(\mathbf{n},\mathbf{s},\mathbf{v}\right)\int_\lambda f_C(\lambda)e(\lambda)c_s(\lambda)d\lambda \qquad (14.4)$$

Figure 14.9 Hue–Saturation of traffic sign colours of different countries.

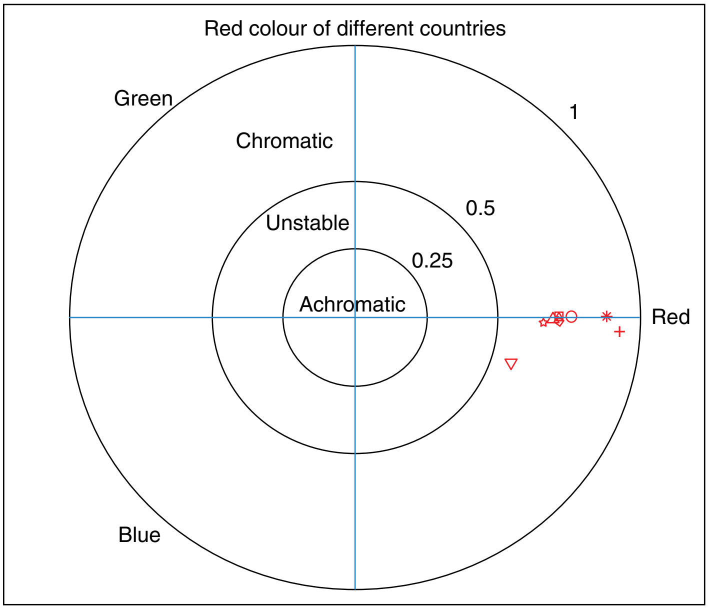
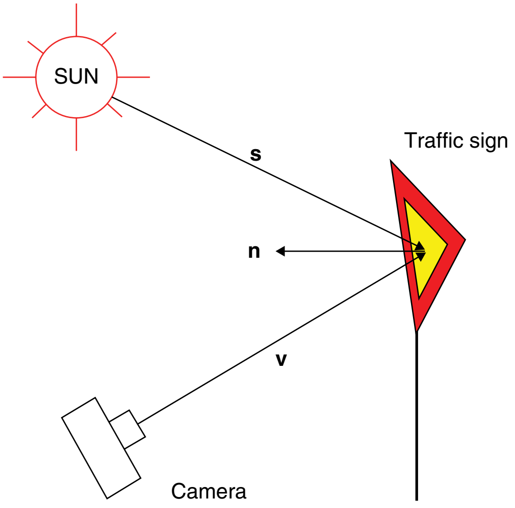


Figure 14.10 Traffic sign reflection model.



for $C = \{R, G, B\}$ where $c_b(\lambda)$ and $c_s(\lambda)$ are the body and surface albedo, respectively; λ is the wavelength at which the sensor responds; and $\mathbf{n}, \mathbf{s}, \mathbf{v}$ are unit vectors that represent the direction of the normal vector to the surface patch, direction of the source of illumination and direction of the viewer, respectively, Figure 14.10.

Furthermore, the terms m_b and m_s denote the geometric dependencies on the body and surface reflection component, respectively [32].

Assuming that surface albedo $c_s(\lambda)$ is constant and independent of the wavelength, and white illumination is used (white illumination means equal energy for all wavelengths within the visible spectrum), then $e(\lambda) = e$ and $c_s(\lambda) = c_s$ which are constants. The sensors responses can be modified as follows:

$$C_w = em_b(\mathbf{n},\mathbf{s})k_C + em_s(\mathbf{n},\mathbf{s},\mathbf{v})c_s\int_\lambda f_C(\lambda)d\lambda \text{ for } C_w = \{R_w, G_w, B_w\} \tag{14.5}$$

In Equation 14.5, C_w is the response of the RGB sensors under the assumption of white light source, and k_C is given by

$$k_C = \int_\lambda f_C(\lambda)c_b\, d\lambda \tag{14.6}$$

where k_C is the compact formulation depending on the sensors and the surface albedo only.
If the assumption of white illumination holds, then

$$\int_\lambda f_R(\lambda)d\lambda = \int_\lambda f_G(\lambda)d\lambda = \int_\lambda f_B(\lambda)d\lambda = f \tag{14.7}$$

and the reflection of the surface can be given by

$$C_w = em_b(\mathbf{n},\mathbf{s})k_C + em_s(\mathbf{n},\mathbf{s},\mathbf{v})c_s f \tag{14.8}$$

The first term of Equation 14.6 represents the effect of shadows on colour invariance, while the second term is effect of highlights on colour invariance. Fleyeh [33] has tested different colour spaces and showed that Hue is the only colour feature which is invariant against viewing direction, surface orientation, highlight, illumination direction and illumination intensity, as depicted in Table 14.1.

Colour segmentation can be designed and implemented in different ways. Simple rule-based colour segmentation algorithms were developed in the literature [34]. However, these algorithms cannot deal with all of the aforementioned problems. Therefore, more advanced and adaptive algorithm which can tackle these problems is needed. Fleyeh and Mumtaz [35] suggested a colour segmentation algorithm for traffic signs based on self-organising maps (SOM). SOM is a powerful unsupervised clustering technique which can be utilised to solve the problem of colour segmentation. Colour segmentation is based on reducing the number of colours in an image. It is achieved by selecting the most representative colours from all colours present in the original image, making a colour palette, and then mapping each colour in the image to the nearest colour in the palette. Since each colour in the segmented image represents an object, only a few colours are desired in the colour palette.

Table 14.1 Invariance of colour models to imaging conditions.

Colour feature	Viewing direction	Surface orientation	Highlight	Illumination direction	Illumination intensity
I	N	N	N	N	N
RGB	N	N	N	N	N
Nrgb	Y	Y	N	N	N
H	Y	Y	Y	Y	Y
S	Y	Y	N	Y	Y

'Y' denotes invariance and 'N' denotes sensitivity of colour models to imaging conditions.

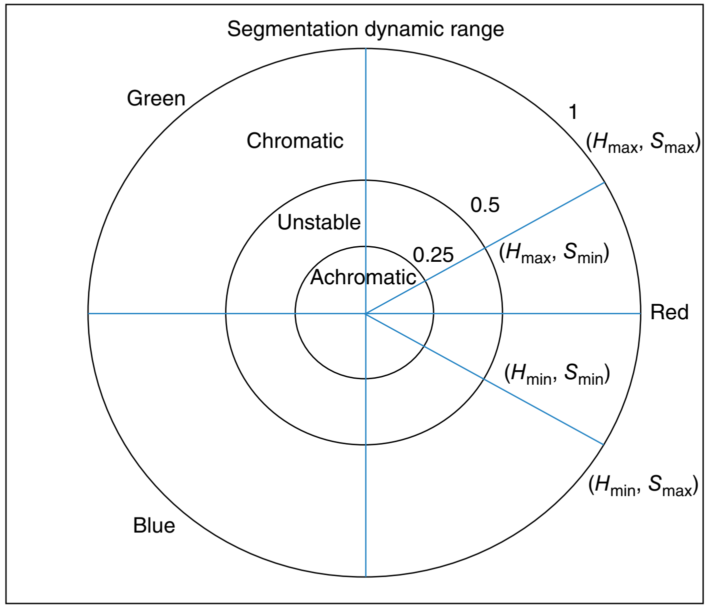


Figure 14.11 Dynamic search boundary by the proposed algorithm for the red colour.

An important feature on which the quality of segmentation depends is the search space boundaries. RGB images were converted into HSV colour space and the H and S values of the traffic signs were observed to assign a reasonable search boundary in the H and S space. Since the values of H and S differ from one traffic sign to another, it is not wise to choose static boundaries. Instead, boundaries of a search subspace in the H–S plot are specified dynamically by the SOM which is trained to find the upper and lower limits of H and S through the best matching neurons. It is, therefore, believed that the algorithm is adaptive and can be used in a wide range of environmental conditions and in different countries.

The values H_{min}, H_{max}, S_{min} and S_{max} in Figure 14.11 are specified dynamically by the SOM. This means that the position and the size of the search space dynamically depends on the desired colour to be segmented.

The training process divides the grid of nodes into most dominant colours depending on the size of the grid. For example if the grid size is 3, a maximum of nine dominant colours present in the test image are obtained. These nine colours can then be used to assign the boundary values for the colour under consideration through its neurons obtained by the proposed approach. The best matching neurons have their own associated weight vectors which can be used to segment the image. Therefore the larger the size of the grid the greater the number of representative colours which will be considered and thus better segmentation can be achieved. Results of colour segmentation are depicted in Figures 14.12 and 14.13, respectively.

14.6.3 Traffic Sign's Rim Analysis

The shape of the traffic sign's rim can be determined by different methods. One method is by exploiting Hough transform. Hough transform in its original form is not suitable for complex object detection such as triangles, circles and rectangles due to noise and shape imperfection. Circular Hough transform can instead be invoked to detect these shapes. It can be described as a transformation of centre point of a circle in x–y plane to the parameter space. The equation of a circle in x–y plane is given by

Figure 14.12 Colour segmentation results of traffic sign images in different light conditions.

$$(x-a)^2+(y-b)^2=r^2 \tag{14.9}$$

where a and b are the centre point of the circle in the x and y direction and r is the radius of the circle. The parametric representation of the circle is given by

$$x=a+r\cos(\theta) \tag{14.10}$$

$$y=b+r\sin(\theta) \tag{14.11}$$

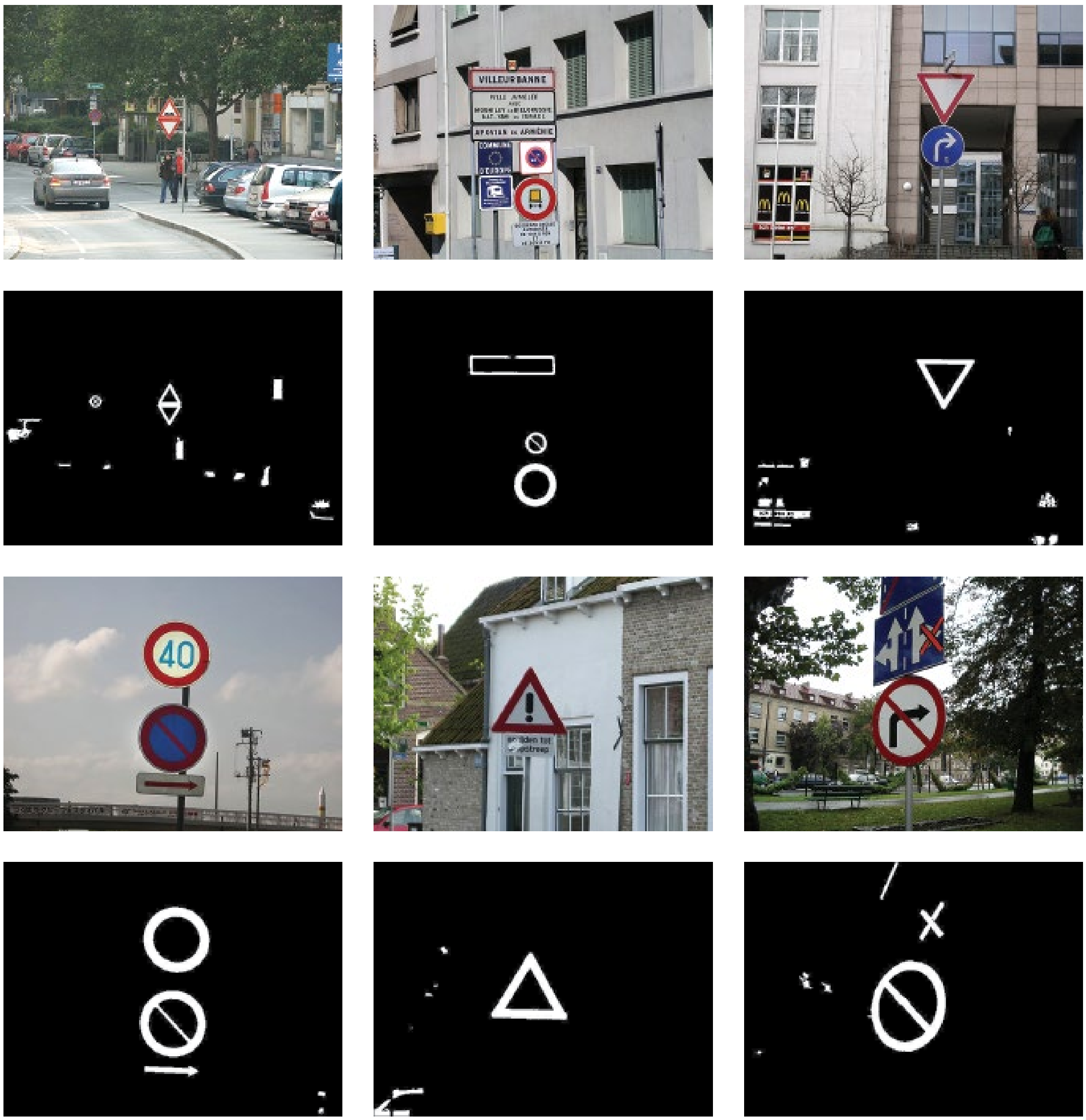


Figure 14.13 Traffic sign images from different countries. (Top left to bottom right) Austria, France, Germany, Japan, The Netherlands, and Poland.

To determine the presence of any traffic sign, it is necessary to accumulate votes in the 3D parameter space (a, b, r). The objective is to find the coordinates (a, b) of the object's centroid and the locus of the objects circumference. A voting mechanism which aims to find the distribution of the votes in the Hough space is illustrated in Figure 14.14.

This voting mechanism specifies the number of votes given to any object in the image and its location. In this mechanism, each white pixel of the object is considered as the centre of a set of concentric circles with different radii. For each radius of the set of circles, a vote is given to each intersection of

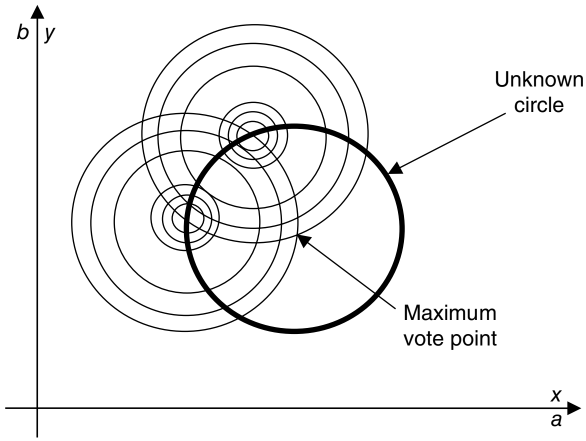


Figure 14.14 Circular Hough transform voting mechanism.

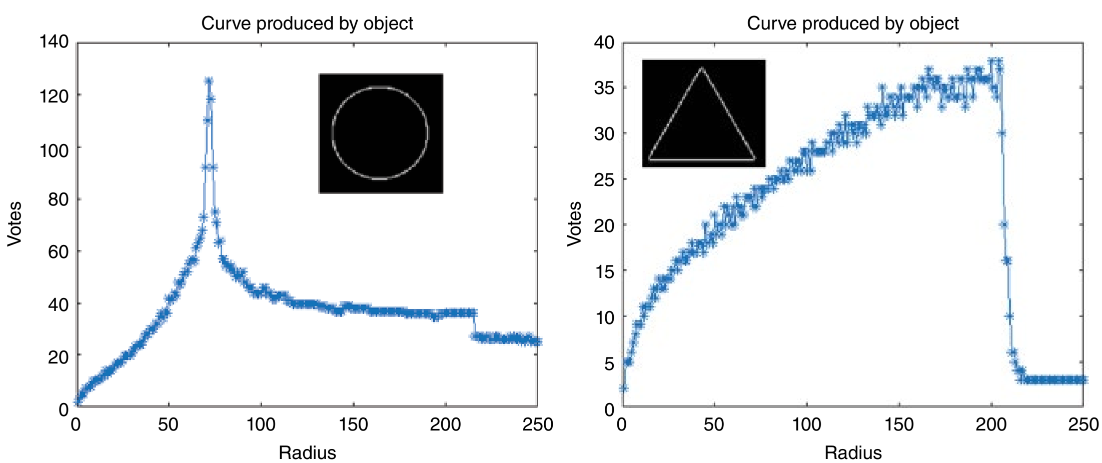


Figure 14.15 Property curves of a circle (a) and a triangle (b).

the circles, and the maximum number of votes is computed. This means that for each radius of the concentric circles there is one point in the Hough space (h, k) which represents the maximum voting.

A circle is detected by its peak generated by the voting mechanism, while a triangle is characterised by a smooth curve and the absence of the peak in the voting curve, as depicted in Figure 14.15. Property curves generated by the voting mechanism were smoothed by LOWESS regression [36] and normalised in both axes to [0, 1] as depicted in Figure 14.16. A training set of curves can be collected and an SVM classifier can be trained with these curves to recognise the different rims of traffic signs. In order to reduce the amount of computations in the recognition phase, edge detection algorithm such as Canny edge detector can be invoked to produce the edges of each of the candidate objects. These objects will sequentially be separated from the image and a voting curve will be created, smoothed and exposed to the SVM classifier for classification. Usually, Canny edge detector generates two edges which represent the outer and inner edges of the object under consideration, as depicted in Figure 14.17. To consider one edge only, the following set of rules were proposed [37, 38]:

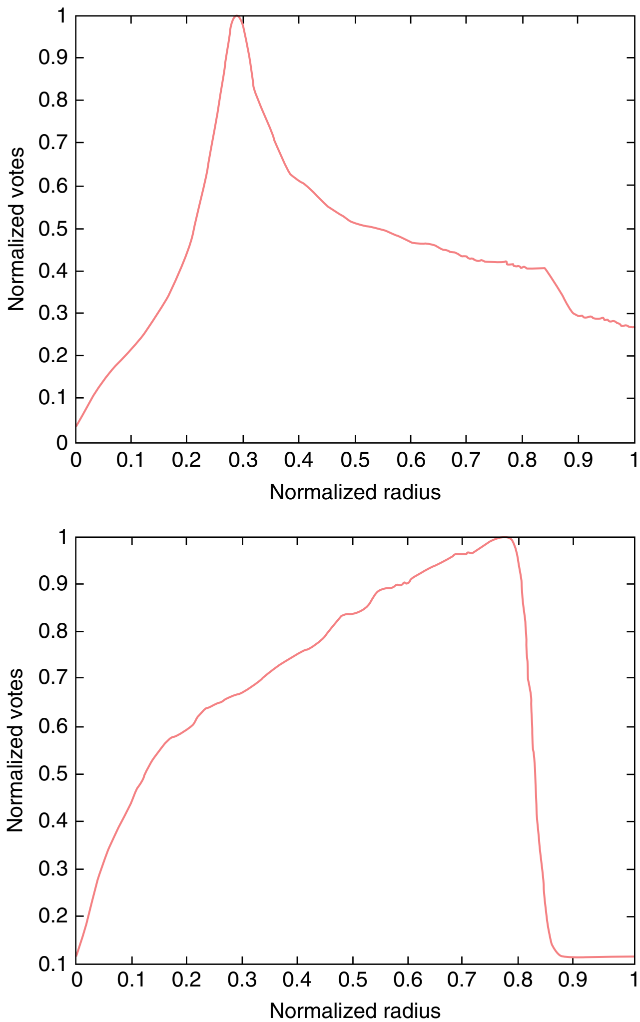


Figure 14.16 Smoothed property curves by LOWESS regression.

- If two concentrated edges are detected, the two edges of the traffic sign are healthy and one traffic sign is detected. The outer edge of the traffic sign is considered.
- If one edge, which is located inside another object, is detected, the outer edge of the traffic sign is destroyed, while the inner edge is healthy. The inner edge of the traffic sign is considered.
- If one edge, which is not located inside another object, is detected, the inner edge of the traffic sign is destroyed, while the outer edge is healthy. The outer edge of the traffic sign is considered.

Another simple method to detect traffic sign rim is by training a classifier by a set of features extracted from the segmented rim. Invariant image features based on integration over a transformation group which were introduced by Schulz-Mirbach [39] can be invoked for this purpose. In many cases during image retrieval, the exact position and orientation of objects in an image are only of secondary value. Thus, it is desirable to have features which are invariant to certain transformations, say translation and rotation.

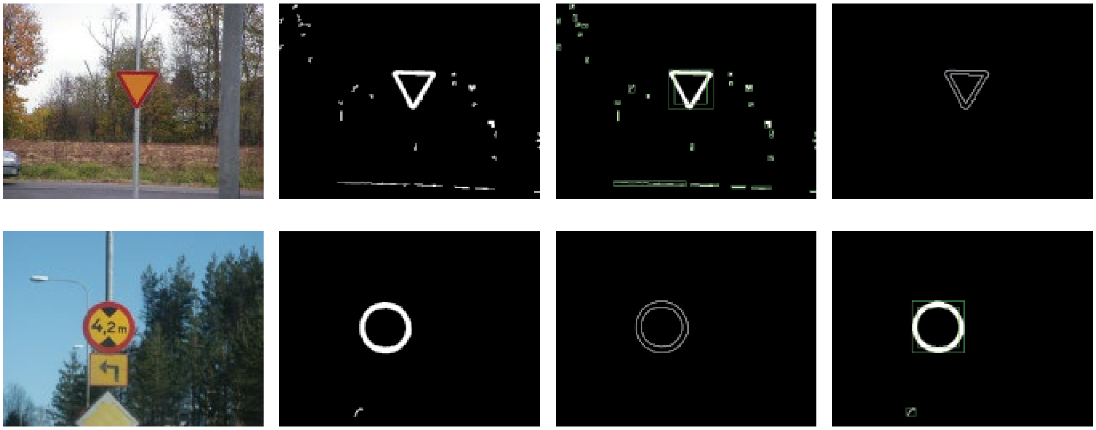


Figure 14.17 Results of rim detection.

Let $I = \{I(i,j)\},\ 0 \le i < N,\ 0 \le j < M$ be an image, with $I(i,j)$ representing the grey value at the pixel coordinate (i,j). Let G be the transformation group of translations and rotations with elements $g \in G$ acting on the image, such that the transformed image is gI. An invariant feature must satisfy $F(gI) = F(I),\ \forall g \in G$. Such invariant features can be constructed by integrating $f(gI)$ over the transformation group G.

$$F(I) = \frac{1}{|G|}\int_G f(gI)\,dg \tag{14.12}$$

For a segmented binary traffic sign image, transformations will be restricted to a certain group of translations. As a kernel function, binary operations among neighbour pixels are proposed. For example, a two-point kernel evaluated at a point (x, y) would be

$$k(x,y) = I(x,y)\,\mathrm{XOR}\,I(x+\Delta_x,\ y+\Delta_y) \tag{14.13}$$

where the pair (Δ_x, Δ_y) determines the local support.

The ith invariant feature is then given by

$$F_i = \frac{1}{MN}\sum_{x=0}^{N-1}\sum_{y=0}^{M-1}\mathrm{XOR}\Big(I(x,y),\ I\big(x+\Delta_x^i,\ y+\Delta_y^i\big)\Big) \tag{14.14}$$

where the pair (Δ_x^i, Δ_y^i) are the translation parameters for the ith kernel. The boundary pixels may optionally be discarded if the corresponding translated point falls outside the image.

Theoretically, the values of (Δ_x^i, Δ_y^i) are $0 \le (\Delta_x^i, \Delta_y^i) < \infty$; but practically, they should not exceed the size of the image. Figure 14.18 illustrates the features which are essentially discriminative for this classification task. The first column is the original image. The other columns depict the result of XORing of the original image with its translated version by the amount mentioned in the first row.

14.6.4 Pictogram Extraction

Once the rim of the traffic sign is specified, the next step is to extract its pictogram. All binary objects in the segmented image are labelled using connected components labelling, and all objects with red rims, yellow or white interiors, and appropriate dimensions will be selected as candidates for further

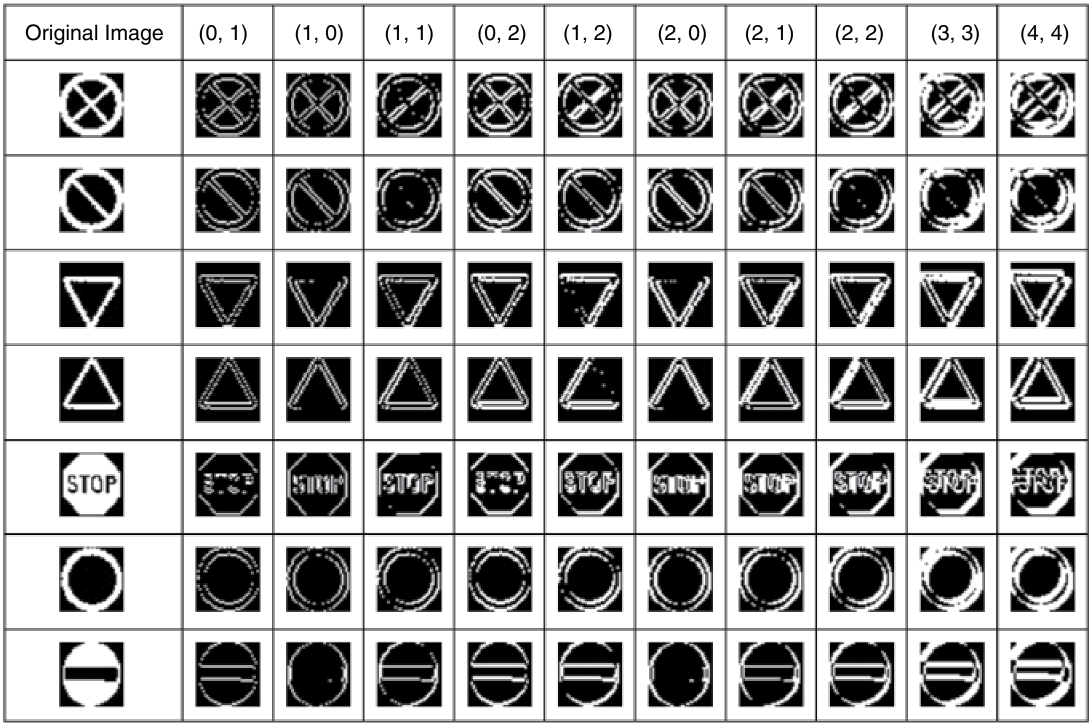


Figure 14.18 Binary Haar features of different traffic sign rims.

investigation. As illustrated in Figure 14.19, the process of extracting the pictogram from the image is described in the following steps [40]:

1) Fill the red region which represents a candidate sign (sub-image A) with white pixels in order to produce (sub-image B).
2) Extract the area corresponding to the pictogram of the sign by XOR operator between sub-image A and sub-image B, that is (sub-image C = sub-image A XOR sub-image B). The resulting area will correspond to the pictogram of the traffic sign.
3) Extract the corresponding pictogram and convert it into grey level (sub-image D).

14.6.5 Pictogram Classification Using Features

In order to classify the pictogram, a set of features which describe it is required. A classifier, which is trained by this set of features, is the tool by which this pictogram is specified. There are many features and descriptors in the literature, but histogram of oriented gradients (HOGs), which was developed by Dalal and Triggs [41] to detect pedestrian, has become a source of attention for many researchers.

To compute the HOG descriptors of any image containing the extracted pictogram, this image is divided into a number of cells and a number of orientation bins as depicted in Figure 14.20. For each cell a local 1D histogram of the gradient directions of edge orientations over the pixels of the cell is collected. For better invariance to illumination such as shadows, the local histogram is accumulated over a larger area called 'blocks'. To improve the contribution of the cells in the final image descriptor, overlapping between these cells is invoked.

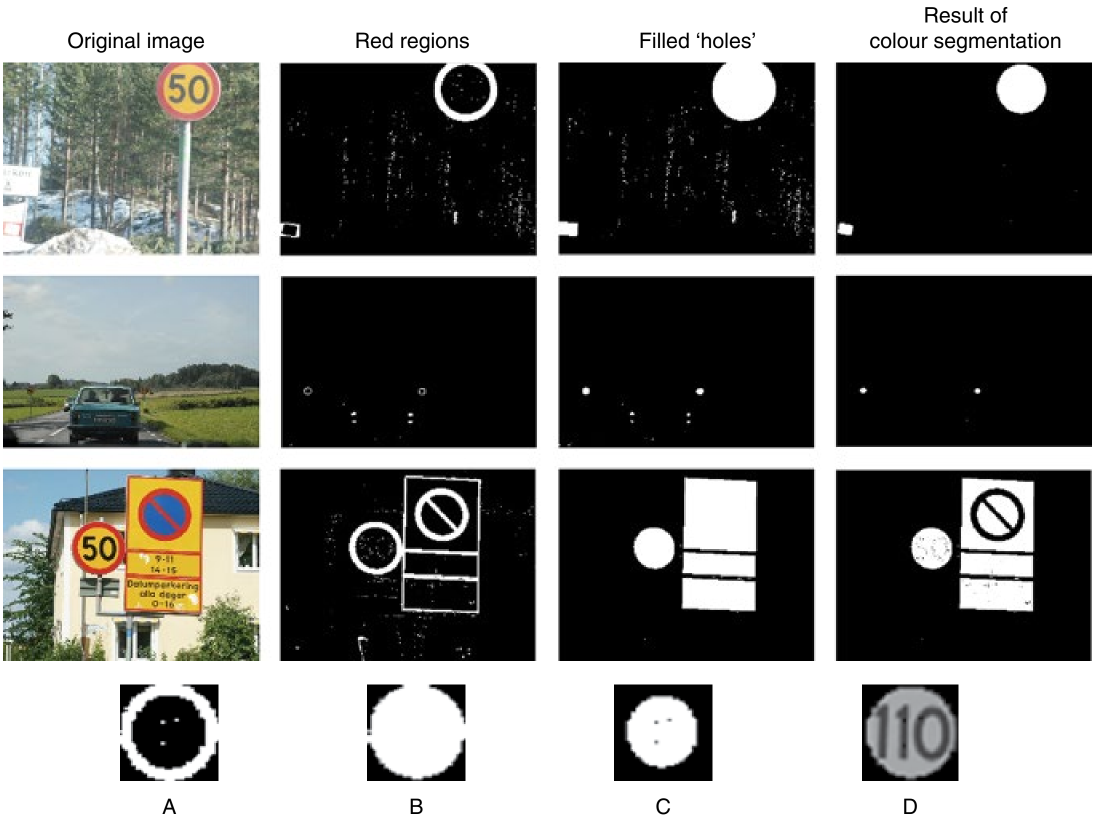


Figure 14.19 Steps of pictogram extraction.

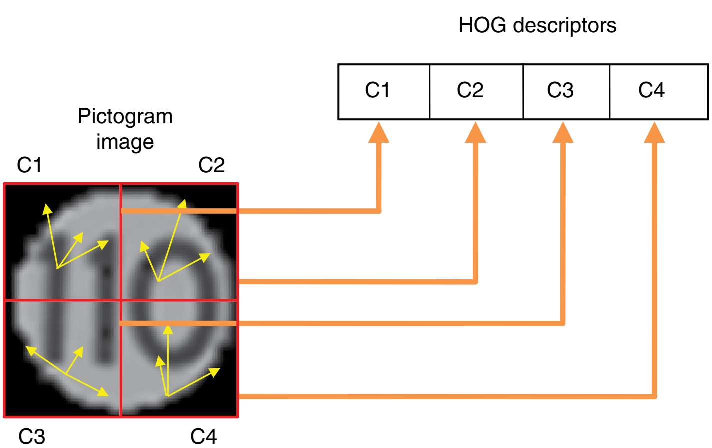


Figure 14.20 Computing the HOG descriptors of a pictogram. Source: Fleyeh and Roch [42]. Reproduced with permission of International Journal for Traffic and Transport Engineering.

Edge orientations are divided into a number of bins. These bins are equally spaced over the interval 0–180° for unsigned gradient and 0–360° for signed gradients. Edge orientations should fit into one of these bins. The histograms collected for the different cells in the bins in the same

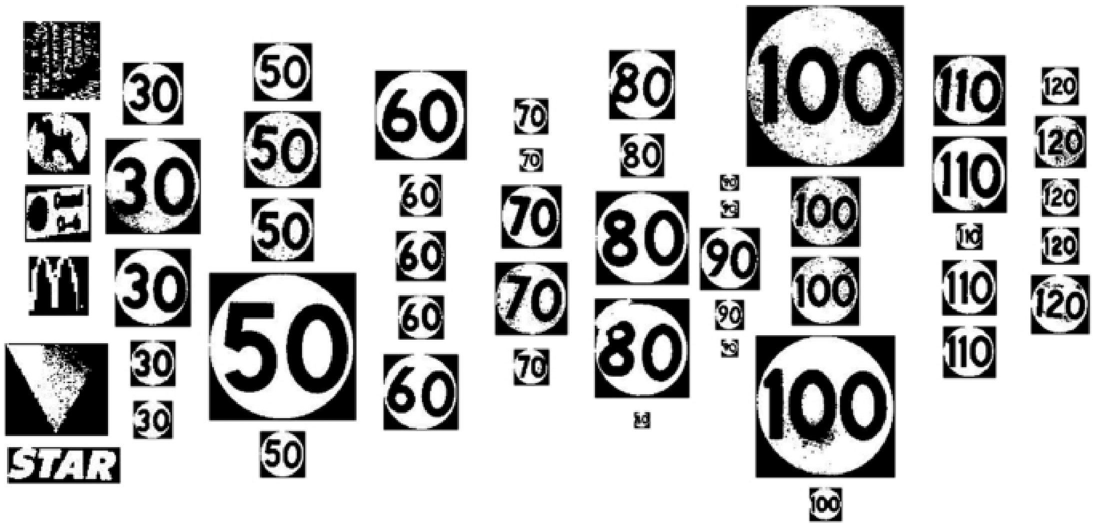


Figure 14.21 The training dataset includes non-traffic sign objects and pictograms.

block are concatenated to make the final set of features of the object under consideration as illustrated in Figure 14.20.

The result of the HOG algorithm is a discrete amount of features which describe the input image. The number of features depends on the number of cells and orientation bins.

As an example of this stage, HOG features are invoked to classify speed limit traffic signs [42]. The dataset comprises 1727 images from which a total of 1710 speed limit signs and 1025 nontraffic sign objects were extracted by the pictogram extractor. A number of images that comprise the dataset are shown in Figure 14.21. Gentle AdaBoost, a boosting algorithm which was introduced by Friedman [43], was trained by the HOG descriptors of the pictograms. This Gentle AdaBoost is a robust and stable version of the Real AdaBoost and performs slightly better than the latter on regular data and considerably better on noisy data. The algorithm uses adaptive Newton steps rather than exact optimization at each step to minimise the exponential criterion in order to stabilise the learning processing. The performance of the HOG descriptors for traffic sign recognition is evaluated in the following sections.

14.6.5.1 Effect of Number of Features

A Gentle AdaBoost classifier was trained and tested with different numbers of features. The number of features was based on different numbers of cells per block and a constant number of orientation bins. The number of orientation bins was nine, while the number of cells varied from 2×2 to 6×6 with an increment of 1 in each direction, that is, $2 \times 2, 3 \times 3, \ldots, 6 \times 6$. This gives 4, 9, 16, 25 and 36 cells in each block. A plot of the classification error versus the number of features indicates clearly that increasing the number of features deduced per block decreases the classification error. Figure 14.22 depicts the relationship between the classification error and the number of features. Since the extracted traffic signs which were exploited in this experiment were of different scales, it is obvious that HOG descriptors are scale invariant. This property is essential to traffic sign recognition because images or footage may contain traffic signs with different sizes depending on the distance between the vehicle and the traffic sign. To have a set of descriptors which performs with scale invariance means that time required for normalisation can be saved for real-time applications.

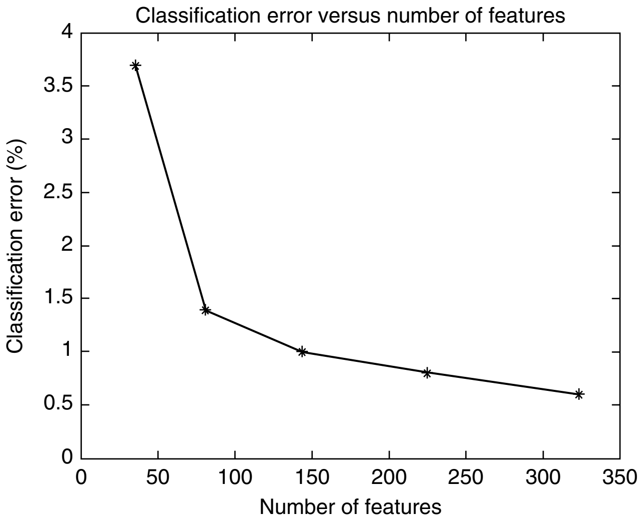


Figure 14.22 Classification errors versus number of features. Source: Fleyeh and Roch [42]. Reproduced with permission of International Journal for Traffic and Transport Engineering.

14.6.5.2 Classifying Disoriented Traffic Signs

Traffic signs are usually installed on poles which are always vertical with respect to the ground level. However, for many reasons such as the nature of the ground or environmental effects, these traffic signs depart from the vertical situation. The Gentle AdaBoost classifier was trained with 360 HOG features (6×6 cells and 10 bins). All of these features were derived from vertically oriented speed limit signs. The classifier was tested with HOG features derived from speed limit images which were rotated by different angles in clockwise and counter clockwise directions. Figure 14.23 depicts a plot of the classification rate versus the angle of rotation.

The plot shows a large drop in the classification rate when disoriented traffic signs were classified with the trained classifier, which means that HOG descriptors are rotation variant and cannot be utilised in these situations. As this is a crucial issue as far as traffic signs are concerned, a proper solution is essential in this case to avoid this kind of invariance.

The Gentle AdaBoost was trained again with a set of HOG descriptors which was derived from speed limit sign rotated by different angles between −90° and 90°. These HOG descriptors were derived in the same manner as described before. The classification rate did not drop as traffic signs rotate. Although there was a slight variation in the classification rate, the average classification rate was 92%. The plot of the classification rate versus the angle of rotation when Gentle AdaBoost was trained with rotated signs is depicted in Figure 14.24.

14.6.5.3 Training and Testing Time

Timings of training the Gentle AdaBoost classifier with 2735 descriptors together with the classification time of a speed limit sign using different numbers of HOG descriptors are depicted in Figure 14.25, showing a plot of training and testing times versus number of features. While training time increases with respect to the number of features, testing time is almost constant regardless of the number of features. The increment in the testing time between a low number of features and a high number of features is not crucial.

This system was tested on real traffic scenes collected on different environmental conditions when the vehicle was driven in the same speed of the traffic sign to be recognised. A number of speed limit results are shown in Figure 14.26.

Figure 14.23 Classification rate of the Gentle AdaBoost trained with vertically aligned signs versus angle of rotation of traffic signs. Source: Fleyeh and Roch [42]. Reproduced with permission of International Journal for Traffic and Transport Engineering.

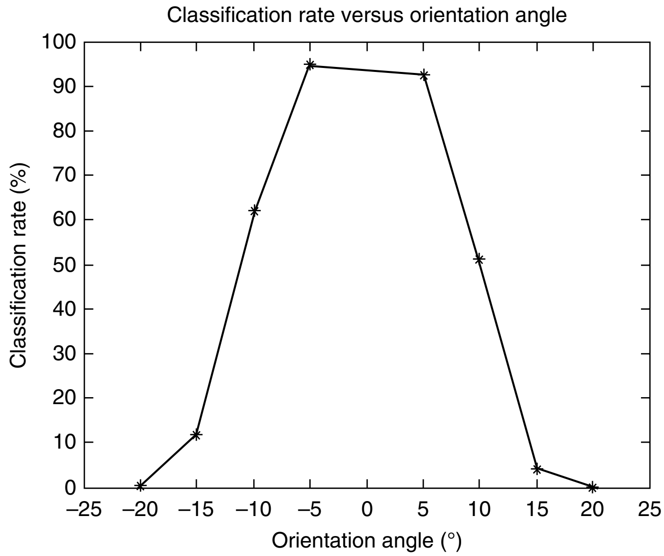
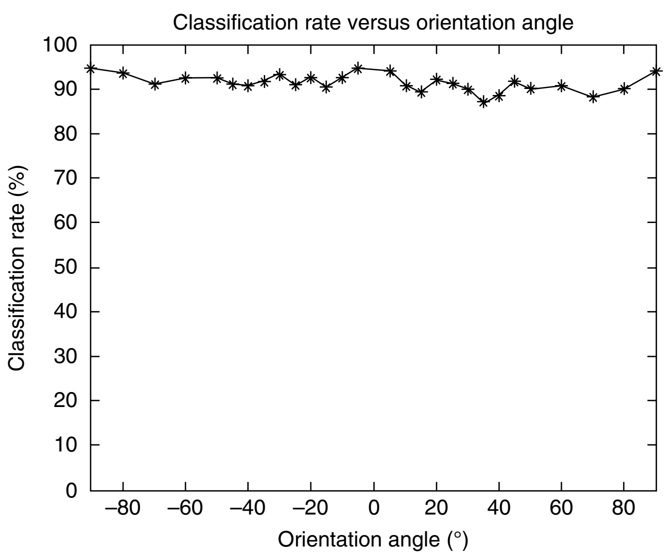


Figure 14.24 Classification rate of the Gentle AdaBoost trained with rotated signs versus angle of rotation of traffic signs. Source: Fleyeh and Roch [42]. Reproduced with permission of International Journal for Traffic and Transport Engineering.



14.7 Working Systems

Traffic sign recognition systems are adopted by many car manufacturers. Traffic recognition systems for speed limit applications similar to the one described earlier was introduced by BMW in 2008 and then followed by Mercedes-Benz. A second generation which can detect overtaking restrictions was introduced by many can manufacturers such as Volkswagen. Volvo used a system called road sign information (RSI) [44]. The system is designed to lower the risk that the driver misses any traffic sign. The vehicle is equipped with a camera and a traffic sign recognition system which looks for the traffic

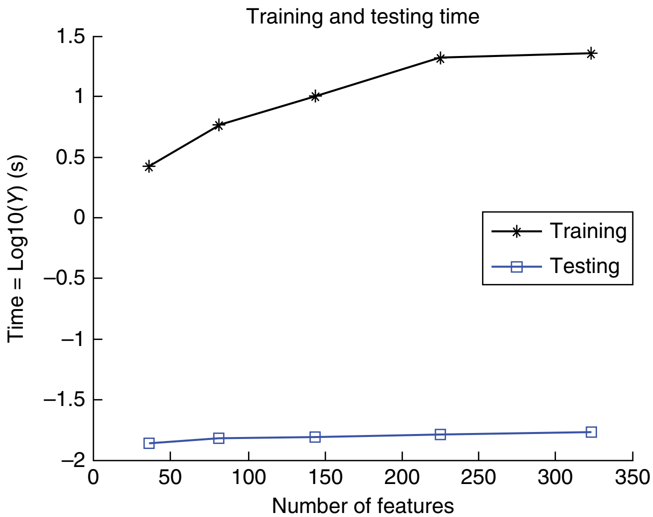


Figure 14.25 Training and testing time of the classifier with HOG descriptors. Source: Fleyeh and Roch [42]. Reproduced with permission of International Journal for Traffic and Transport Engineering.

Figure 14.26 Correctly classified speed limit traffic signs.

sign in the scene in front of the vehicle in real time; once the traffic sign is detected, it will be displayed as symbols on the speedometer of the vehicle. This technique, together with other new technologies such as adaptive cruising system, collision warning, automatic queue assistant system and automatic lane detection system, increases traffic safety which is one of the most important goals to achieve in the intelligent vehicles.

References

1 Fang, C., Chen, S. & Fuh, C. (2003) Road-sign detection and tracking. *IEEE Transactions on Vehicular Technology*, 52, 1329–1341.

2 Fang, C., Fuh, C., Chen, S. & Yen, P. (2003) A road sign recognition system based on dynamic visual model. *The 2003 IEEE Computer Society Conf. Computer Vision and Pattern Recognition*, Madison, WI, 18–20 June 2003, pp. 750–755.

3 de la Escalera, A., Moreno, L., Puente, E. & Salichs, M. (1994) Neural traffic sign recognition for autonomous vehicles. *20th International Conference on Industrial Electronics Control and Instrumentation, 1994*, IEEE, Bologna, Italy, 5 September 194, pp. 841–846.

4 Estevez, L. & Kehtarnavaz, N. (1996) A real-time histographic approach to road sign recognition. *IEEE Southwest Symposium on Image Analysis and Interpretation*, San Antonio, TX, 8–9 April 1996, pp. 95–100.

5 Miura, J., Kanda, T. & Shirai, Y. (2000) An active vision system for real-time traffic sign recognition. *2000 IEEE Intelligent Transportation Systems*, Dearborn, MI, 1–3 October 2000, pp. 52–57.

6 Vitabile, S. & Sorbello, F. (1998) Pictogram road signs detection and understanding in outdoor scenes. *Conference on Enhanced and Synthetic Vision*, Orlando, FL, April 1998, pp. 359–370.

7 Parodi, P. & Piccioli, G. (1995) A feature-based recognition scheme for traffic scenes. *Intelligent Vehicles '95 Symposium*, Detroit, MI, 25–26 September 1995, pp. 229–234.

8 Vitabile, S., Gentile, A. & Sorbello, F. (2002) A neural network based automatic road sign recognizer. *The 2002 International Joint Conference on Neural Networks*, Honolulu, HI, 12–17 May 2002, pp. 2315–2320.

9 Plane, J. (1992) *Traffic Engineering Handbook*. Englewood Cliffs, NJ: Prentice-Hall.

10 Jiang, G. & Choi, T. (1998) Robust detection of landmarks in color image based on fuzzy set theory. *Fourth International Conference on Signal Processing*, Beijing, China, 12–16 October 1998, pp. 968–971.

11 Lalonde, M. & Li, Y. (1995) Road sign recognition – survey of the state of art. Technical Report for Sub-Project 2.4, Center de recherche informatique de Montrèal, CRIM-IIT-95/09-35.

12 Wikipedia. (2014) Road Signs in the United States (online). Available: https://en.wikipedia.org/wiki/Road_signs_in_the_United_States (accessed 20 July 2014).

13 Paclik, P. (1984) ITS, Intelligent Transport System (online). Available: http://euler.fd.cvut.cz/research/rs2/articles/itsp.html (accessed 10 September 2016).

14 de la Escalera, A., Armingol, J. & Mata, M. (2003) Traffic sign recognition and analysis for intelligent vehicles. *Image and Vision Computing*, 21, 247–258.

15 Kehtarnavaz, N. & Kang, D. (1993) Stop-sign recognition based on color/shape processing. *Machine Vision and Applications*, 6, 206–208.

16 Volvo. (2011) Samlar den bästa säkerhetstekniken. *Aftonbladet* (online). Available: http://www.aftonbladet.se/kampanj/arkiv/volvo/innovation/article13635643.ab (accessed 10 September 2016).

17 Volvo. (2015) Säkerhet på en ny nivå (online). Available: http://www.volvocars.com/se/om-volvo/innovationer/intellisafe/driver-support (accessed 10 September 2016).

18 Mei, T., Liang, H., Kong, B., Yang, J., Zhu, H., Li, B., Chen, J., Zhao, P., Xu, T., Tao, X., Zhang, W., Song, Y., Wei, H. & Wang, J. (2012) Development of “Intelligent Pioneer” unmanned vehicle. *IEEE 2012 Intelligent Vehicles Symposium*, Alcalá de Henares, Spain, 3–7 June, pp. 938–943.

19 Blancard, M. (1992) Road sign recognition: a study of vision-based decision making for road environment recognition. In: Masaki, I. (ed.) *Vision-based Vehicle Guidance*. Berlin, Germany: Springer-Verlag.

20 Vitabile, S., Gentile, A., Dammone, G. & Sorbello, F. (2002) Multi-layer perceptron mapping on a SIMD architecture. *12th IEEE Workshop on Neural Networks for Signal Processing*, Martigny, Switzerland, 4–6 September 2002, pp. 667–675.

21 Vitabile, S., Pollaccia, G., Pilato, G. & Sorbello, F. (2001) Road sign recognition using a dynamic pixel aggregation technique in the HSV color space. *11th International Conference on Image Analysis and Processing*, Palermo, Italy, 26–28 September 2001, pp. 572–577.

22 Buluswar, S. & Draper, B. (1998) Color recognition in outdoor images. *International Conference on Computer Vision*, Bombay, India, 4–7 January 1998, pp. 171–177.

23 Luo, R., Potlapalli, H. & Hislop, D. (1993) Outdoor landmark recognition using fractal based vision and neural networks. *1999 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Yokohama, Japan, 26–30 July 1993, pp. 612–618.

24 Paclik, P. & Novovicova, J. (2000) Road sign classification without color information. *Sixth Annual Conference of the Advanced School for Computing and Imaging*, Lommel, Belgium, 14–16 June 2000.

25 Paclik, P., Novovicova, J., Pudil, P. & Somol, P. (2000) Road sign classification using Laplace kernel classifier. *Pattern Recognition Letters*, 21, 1165–1173.

26 Kang, D., Griswold, N. & Kehtarnavaz, N. (1994) An invariant traffic sign recognition system based on sequential color processing and geometrical transformation. *IEEE Southwest Symposium on Image Analysis and Interpretation*, Dallas, TX, 21–24 April 1994, pp. 88–93.

27 Perez, E. & Javidi, B. (2000) Composite filter bank for road sign recognition. *13th Annual Meeting IEEE Lasers and Electro-Optics Society*, Rio Grande, Puerto Rico, 13–16 November 2000, pp. 754–755.

28 Fleyeh, H. (2009) Traffic Sign Database (Online). Available: http://users.du.se/~hfl/traffic_signs/ (accessed 10 September 2016).

29 Stallkamp, J., Schlipsing, M., Salmen, J. & Igel, C. (2011) The German traffic sign recognition benchmark: a multi-class classification competition. *IEEE International Joint Conference on Neural Networks*, San Jose, CA, 31 July–5 August.

30 Larsson, F. & Felsberg, M. (2011) Using Fourier descriptors and spatial models for traffic sign recognition. In: Heyden, A. & Kahl, F. (eds.) *Image Analysis*. 17th Scandinavian Conference on Image Analysis, SCIA 2011. Berlin/Heidelberg: Springer, pp. 238–249

31 Buluswar, S. & Draper, B. (1994) Non-parametric classification of pixels under varying outdoor illumination. *Intelligent Robots and Computer Vision XIII: Algorithms and Computer Vision*, October 1994, Vol. 2353, pp. 529–536.

32 Gevers, T. & Smeulders, A. (1999) Color-based object recognition. *Pattern Recognition*, 32, 453–464.

33 Fleyeh, H. (2006) Shadow and highlight invariant colour segmentation algorithm for traffic signs. *2006 IEEE Conference on Cybernetics and Intelligent Systems*, Bangkok, Thailand, June 2006, pp. 108–114.

34 Fleyeh, H. (2004) Color detection and segmentation for road and traffic signs. *2004 IEEE Conference on Cybernetics and Intelligent Systems*, Singapore, 1–3 December 2004, p. 808–813.

35 Fleyeh, H. & Mumtaz, A. (2011) Adaptive shadow and highlight invariant colour segmentation for traffic sign recognition based on Kohonen SOM. *Journal of Intelligent Systems*, 20, 15–31.

36 Cleveland, W. & Devlin, S. (1988) Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83, 596–610.

37 Fleyeh, H., Bhuiyan, N. & Biswas, R. (2011) Prohibitory traffic signs detection using LVQ and windowed Hough transform. *Fifth Indian International Conference on Arificial Intelligence*, Tumkur, India, 14–16 December.

38 Fleyeh, H., Biswas, R. & Davami, E. (2013) Traffic sign detection based on AdaBoost based color segmentation and SVM classification. *IEEE EuroCon 2013*, Zagreb, Croatia, 1–4 July 2013.

39 Schulz-Mirbach, H. (1995) Invariant features for gray scale images. *17 DAGM – Symposium "Mustererkennung"*, Bielefeld. Reihe Informatik aktuell, Springer, pp. 1–14.

40 Fleyeh, H. & Davami, E. (2011) Eigen-based traffic sign recognition. *IET Intelligent Transport Systems*, 5, 190–196.

41 Dalal, N. & Triggs, B. (2005) Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE Computer Society Washington, DC, 20 June 2005, pp. 886–893.

42 Fleyeh, H. & Roch, J. (2013) Benchmark evaluation of HOG descriptors as features for classification of traffic signs. *International Journal for Traffic andTransport Engineering*, 3, 448–464.

43 Friedman, J., Hastie, T. & Tibshirani, R. (2000) Additive logistic regression: a statistical view of boosting. *Annals of Statistics*, 28, 337–407.

44 Wikipedia. (2014) Traffic Sign Recognition (online). Available: https://en.wikipedia.org/wiki/Traffic_sign_recognition (accessed 20 July 2014).