Chapter 2
Background on Traffic Sign Detection and Recognition

Abstract The automatic sign detection and recognition has been converted to a real challenge for high performance of computer vision and machine learning techniques. Traffic sign analysis can be divided in three main problems: automatic location, detection and categorization of traffic signs. Basically, most of the approaches in locating and detecting of traffic signs are based on color information extraction. A natural question arises: which is the most proper color space to assure robust color analysis without influence of the exterior environment. Given the strong dependence on weather conditions, shadows and time of the day, some autors focus on the shape-based sign detection (e.g. Hough transform, ad-hoc models based on Canny edges or convex hulls). Recognition of traffic signs has been addressed by a large amount of classification techniques: from simple template matching (e.g. cross-correlation similarity), to sophisticated Machine learning techniques (e.g. suport vector machines, boosting, random forest, etc), are among strong candidates to assure straightforward outcome necessary for a real end-user system. Moreover, extending the traffic sign analysis from isolated frames to videos can allow to significantly reduce the number of false alarm ratio as well as to increase the precision and the accuracy of the detection and recognition process.

Keywords Traffic sign detection · Traffic sign recognition · Color-based description · Shape-based description · Uncontrolled · Environments · Multi-class classification

Recognition of road signs is a challenging problem that has engagedthe attention of the Computer Vision community for more than 30 years. According to Paclik [1], the first study of automatedroad sign recognition was reported in Japan in 1984. Since then, anumber of methods have been developed for road sign detection andidentification. For years, researchers have been addressingthe difficulties of detecting and recognizing traffic signs.

The most common automatic systems for traffic signs detection and recognitioncomprise one or two video cameras mounted on the front of the vehicle

S. Escalera et al., *Traffic-Sign Recognition Systems*, SpringerBriefs in Computer Science,
DOI: 10.1007/978-1-4471-2245-6_2,

(e.g. a geovan). Recently, some geovans also have another camera at the rear and/or the side of the vehicle recording the signs behind or alongside the vehicle. The cars are fitted with a PC system for acquiring the videos, or specialized hardware for driving assistance applications.

Road signs have specific properties that distinguish them fromother outdoor objects. Operating systems for the automatic recognition of road signs are designed to identify these properties. Traffic sign recognition systems have three main parts:

1. Location of the region of interest and segmentation: usually, a number of binary masks are generated to separate the objects of interest from the background. Usually, color information is applied since traffic signs are characterized by a predetermined number of relatively constant colors (white, red, and blue). As a result, regions of interest are determined as connected components, some of which are traffic signs.
2. Detection by verification of the hypothesis of the presence of the sign: to detect signs most authors use knowledge of their shape (e.g. equilateral triangles, circles, etc.)
3. Categorization of the type of traffic sign: the final step is the recognition of the sign using a fixed database of all possible traffic sign models. Methods ranging from template matching to sophisticated machine learning apparatus can be used to achieve robust and efficient recognition of traffic signs.

The detection of the signs from outdoor images is the most complex step in the automatic traffic sign recognition system [2]. Many issues make theproblem of the automatic detection of traffic signs difficult (see Fig. 2.1) such as changeable light conditions which are difficult to control (lighting varies according to the time of the day, season, cloud cover and otherweather conditions); presence of other objects on the road (traffic signs are often surrounded by other objects producing partial occlusions, shadows, etc.); it is very difficult to generate all off-line models of all the possibilities since signs have a high number of degrees of freedom (their size depends on the camera distance and angle of views causing affine deformation; equally, age and accidents can also affect the signGs appearance). Hence, any robust automatic detection and recognition system must provide straightforward results that are not affected by perspective distortion, lighting changes, partial occlusions or shadows [3]. Ideally, the system should also provide additional information on the lack of visibility, poor conditions and poor placement of the traffic signs.

2.1 Color-Based Sign Detection

Sign detection using color is based on the five typical colors defined in standard traffic signs (red, blue, yellow, white and black). Most researchers look for robust color segmentation, paying special attention to non-homogeneous illumination, since errors in segmentation may be propagated in the following steps of the system. One of

Fig. 2.1 A real outdoor scene of a road with traffic sign

the first attempts to construct a real-time system for automatic traffic sign recognition appears in Akatsuka et al. [4] where a look-up table in Nrgb color space is used in order to specifically design the segmentation of the speed limit signs. The authors in [5] studied the changes in traffic sign colors according to the time of the day (sunrise, midday, sunset). They deduced that the variation of the outdoor illumination does not significantly affect the RGB component differences of the colors of TSs, and proposed a simple algorithm for the segmentation of the images as a first step towards an automatic TS detection system. Zadeh et al. in [6] proposed a careful analysis of the nature of variations of pixel values of the same color in the RGB space and designed sub-spaces corresponding to the highest rate of variation of the traffic sign colors.The sub-spaces are defined as canonical regions placed on straight lines from RGB (0,0,0) to the combination of primary colors corresponding to the TSs. Other color appearance models have also been presented although they have received less attention. For example, Escalera et al. in [7] operated on the ratios between the intensity of a given channel and the sum of all RGB channel intensities. They pointed out that the RGB-HSV conversion formulas are non-linear and hence the computational cost involved is excessively high. Ritter et al. in [8, 9] combined color of image regions to create a specific code for traffic sign detection and then applied different shape filters to detect the TS presence within the region of interest. False color regions were filtered by using special look-up tables in the RGBcolor space.

A different approach was adopted in [10] where the authors trained scale-specific road sign detectors using Haar wavelet features parameterized by color. The best features were selected by a cascaded AdaBoost framework from a large space of features defined over multiple color representations: plain R, G, and B channels, normalized R, G, and B channels, and a gray-scale channel. In this way, the most suitable color representation wasinferred automatically from the data rather than arbitrarily chosen by the authors. Despite the excellent final results, one of the main drawbacks of these approaches is the high computational load that makes it difficult to produce a real-time system. As the authors note, for a $384\times$ 288 video resolution a processing speed of only 10 frames per second is achieved.

Among the studies in which the color information was used to detect the traffic signs, a significant amount of work can be found based on non-RGB color spaces. The Hue-Saturation-Value (HSV) color model was adopted because it is based on human color perception and is considered largely invariant to illumination changes. In [11] the authors defined sub-spaces limiting the color values of stop signs. Liu et. al. in [12] used color quantization in the HSV color model to find ROIs, followed by border tracing and ROI scaling. Piccioli et al. [13] defined the regions of interest by clusteringsmall blocks of the image where the number of pixels with a hue in an appropriate range exceeded a predetermined threshold.

The authors in [14] use a HSV model to classify the test sign images into several distinctive categories. Fang et. al. in [15] extract color features by a neural network. The authors in [16] presented a novel parallel segmentation algorithm called color structure code based on a hierarchical region growing on a special hexagonal topology that extracts sets of color regions (in the HSI color space) from theimages. The authors in [17] estimated the appearance of sign-characteristic colors independently from real-life images taken under different viewing conditions. They adopted the CIECAM97 color model to segment the images. An original and robust approach to color-based detection and segmentation of road signs using IHLS color space was also proposed in [18]. Both these studies considered device-independent color appearance models, which may be a better option than the standard device-dependent RGB model.

Recent work has recorded significant advances applying advanced machine learning techniques. Nguwi in [19] segment traffic signs pixels in YCbCr color space by using a multi-layer perceptron trained on traffic signs vs. non-traffic signs regions of interest. Fang et al. in [20] applied a Spatio-Temporal Attentional neural network to extract information of interest related to the formation of focuses of attention. Sign detection is obtained by analyzing the combination of color and edge information. The image features were learned by Configurable Adaptive Resonance Theory and Hetero-Associative Memory networks to classify the category and the object within that category.

A complete comparison between different segmentation techniques is presented by Gómez-Moreno in [21] where special attention is paid to the different color spaces to represent TS image features. The authors showed that the best methods are those that are normalized with respect to the illumination, such as RGB or Ohta Normalized, and, surprisingly, there is no improvement in the use of Hue Saturation Intensity spaces. In addition, using a LUT with a reduction in the less-significant bits improves speed, while maintaining quality of the TS segmentation. Applying a support vector machine, the authors obtain reasonable results although, as they admit, improvements in achromatic colors are still required in order to achieve the level of performance a real system needs.

On the other hand, color segmentation may suffer from various phenomena such as distance from the target, weather conditions, time of day, or reflectance of the signsĠ surfaces. Some authors have preferred a strictly colorless approach. They apply genetic algorithms [22] or distance transforms (DTs) [23]. In [24], the images were transformed using wavelets and classified by a Perceptron neural network.

Even though changes in lighting conditions affect the color information, color remains a useful clue for detecting and recognizing traffic signs. Moreover, signs may have very similar appearances and color may be a very important characteristic for distinguishing between them. When there is a large number of classes in the traffic sign database, color carries very valuable discriminative information which should be used whenever possible. Finally, we should not forget that in some countries, e.g. Japan, there are pairs of signs which differ only in terms of color. Thus, developing robust color models taking into account the variance of color appearance is of great interest for final automatic driving systems.

2.2 Shape-Based Sign Detection

Recent advances in object recognition open up great opportunities for robust traffic sign detection in uncontrolled environments. Still, challenges remain, such as the detection of traffic signs in cluttered scenes, in varying conditions of brightness and illumination, affine distortions according to the point of view, partial occlusions or even other signs and other information attached to traffic signs. Another additional difficulty is the simplicity of the shape of the traffic sign, which means that it can be easily confused with other objects or parts of objects. Depending on its distance from the acquisition system, the traffic signGs size can vary and its spatial resolution may be very low (e.g. 30–40 pixels).

Given the regular geometric information provided by the traffic signs, one of the first attempts to address the problem was to apply Hough transform on the edge map of the region of interest [11]. To speed up the process Piccioli in [13] used color information to limit the region of interest followed by a geometrical analysis on the edge map to extract and detect circular and triangular shapes. After extracting straight lines using Canny edges the different segments of proper length and slope may be suitable for use as traffic signs. In [16] the authors propose a method that first segments the images based on their color and then applies a local and global growing technique to hierarchically organize the information and form the traffic sign candidates. Their convex hulls are compared to a predetermined set of basic traffic sign shapes to check that the region of interest represents a traffic sign. Moreover, the authors claim that the real-time performance they achieve makes the method very attractive for final integration in operating systems.

Again, machine learning approaches significantly improve the final results of traffic sign detection. In [7], the authors developed a simple approach based on color thresholding and shape analysis to detect the signs,followed by a neural network to classify them. In later work [3], they showed that the use of color LUTs in the HSI space to identify the regions of interest and genetic algorithms to detect the signs within the regions of interest produced much better final results. The main advantage of their approach based on genetic algorithms is that it allows for efficient traffic sign detection regardless of position, scale, rotation, partial occlusions, the presence of other objects, and variations in weather conditions. Final traffic sign recognition is achieved by aneural network. An alternative is presented by Gavrila et al. in [23] in which potential traffic signs are identified by a template-based correlation method using distance transforms,while the classification is based on radial basis function networks. A different machine learning approach is used in [25] where a support vector machine is responsible for segmenting the image in the RGB space followed by the detection of circular shapes, paying special attention to possible sign deformations.

2.3 Sign Recognition

Once the region of interest is determined and a traffic sign is detected, it should be recognized using a predetermined database of all the traffic signs in the system. Most of the state-of-the-art approaches can be divided into two strategies: template-based vs. classifier-based comparison. The most frequent similarity measure in traffic sign classification is normalized cross-correlation [13]. Here, a specific similarity measure between the gray-level region of interest and several templates of the traffic signs is estimated based on the normalized cross-correlation in order to identify the traffic sign. Using 60 circular signs and 47 triangular signs in their database, these authors report an accuracy of 98%.The main advantage of normalized cross-correlation isits simplicity, robustness to varying illumination conditions, and the ease of finding a statistical interpretation. It is also appealing from the point of view of implementation since it has been shown to be equivalent to an optical correlator [26]. Another template matching method can be observed in [27] who considered a database of 30 triangular signs and 10 circular signs and achieved an accuracy of 85%. Zadeh et al. in [6] presented an approach for model matching analyzing the area ratio of the colors in the regions of interest horizontally and vertically in order to differentiate between traffic signs with similar overall color ratios. The advantage of their approach is that the method is invariant to rotation and size of the region and so the number of potential models is substantially reduced, thus accelerating the systemĠs recognition time. Because of the global nature of template-based matching (e.g. cross-correlation similarity), the method may suffer from the presence of non-informative regions of interest pixels (e.g. due to occlusions). To cope with this problem, in [28] the authors propose a novel trainable similarity measure based on individual matches in a set of local image regions. Additionally, the set of regions relevant for a particular similarity measure can be refined by training procedures. A step forward in measuring the dissimilarity between different signs is presented by [29] where a special color distance transform enables robust comparison of discrete color image signs. The authors show that this method of feature selection combined with a one-vs-all nearest neighbor classifier performs better than alternatives like Principal Component Analysis or Adaboost offering adequate description of signs with little training effort.

The second group of approaches is based on more sophisticated machine learning techniques which in general achieve more robust final results when the images are analyzed under uncontrolled environments.The early work was based on neural networks such as the one presented by Krumbiegel et al. in [30] where a neural network is trained on the color, shape and texture of the traffic sign templates. Douvilee in [24] used amultilayer perceptron on the fast Fourier transform of the detected sign and a bank of filters, and found that a neural network achieves better results than the template matching procedure. Cyganek in [25] used two committee networks operating in the spatial domain and the log-polar representation. Each committee network was composed by several Hamming neural networks trained on a set of signs of reference. A san alternative, Nguwi et al. in [19] showed that a cascade of multilayer perceptron machines achieved better results than a support vector machine where Resilient Back propagation and Scaled Conjugate Gradient algorithms achieved a classification accuracy of 96% in near real time.

One of the main drawbacks of current methods for traffic sign recognition is the lack of a public domain database that contains enough examples of a wide set of traffic signs acquired in uncontrolled environments [2]. Most of the published works are presented on a limited number of signs with a small number of different examples. An exception is the study in [2] where a database of 1,500 road scene images is analyzed with traffic signs extracted and included in the board image data set and over 3,500 images of individual traffic sign boards. Images have been acquired from three European countries: the Czech Republic, Spain and the United Kingdom, and represent a great variability of illumination and meteorological conditions. In order to achieve robust results, a self-organizing map is used to detect potential road signs by analyzing the distribution of red pixels within the image. Traffic signs are detected from the distributions of the dark pixels in their pictograms. Following this, a hybrid system is introduced combining programmable hardware and neural networks for embedded machine vision leading to a robust and fast final prototype of the system and achieving very satisfactory results in near real time.

A real end-user system for traffic sign detection and recognition must offer high performance but should also be real-time. In [31] the authors apply the radial symmetry detector in order to discover the speed signs in urban scenarios and test it in a real environment. The detector is mounted inside a road vehicle and its performance, speed and stability are tested while the vehicle is moving. The detector is shown to run under a wide variety of visual conditions. Another study in which radial symmetry helps to achieve fast and robust detection is presented in [32]. The authors show that combining a Canny edge detector with the radial symmetry techniques and Adaboost as a classifier obtains very accurate final traffic sign recognition results in spite of the high variance of sign appearance due to noise, affine deformation, and reducedillumination.

Most studies on automatic detection and recognition of traffic signs are developed in a single frame. However, mobile mapping systems are also used to acquire videos or sequences of traffic scenes; thus, temporal information can prove very useful for increasing the accuracy of the detection and recognition process. In [15] the authors apply Kalman filters to track the signs until their size is large enough to ensure robust recognition results. A more recent study by the same authors [20] presents the evolution of the system based on a computational model of human visual recognition processing. The system consists of three basic parts: sensory, perceptual and conceptual analyzers. The sensory part extracts the spatio-temporal information from the video sequences. The information obtained is used by a spatio-temporal attentional neural network in the perceptual analyzer. This part decides whether the focus of attention corresponds to a horizontal projection of a road sign. The signs detected are normalized and correlated with all the prototypes previously stored in the traffic sign database. Excellent results have been achieved with tolerance of three pixels due to the spatio-temporal information embedded in the tracking system and the final recognition module.

References

1. Paclik, P.: Road sign recognition survey. Online, http://euler.fd.cvut.cz/research/rs2/files/skoda-rs-survey.html
2. Prieto, M., Allen, A.: Using self-organizing maps in the detection and recognition of road signs. Image Vis. Comput. **27**, 673–683 (2009)
3. de la Escalera, A., Armingol, J., Mata, M.: Traffic sign recognition and analysis for intelligent vehicles. Image Vis. Comput. **21**, 247–258 (2003)
4. Akatsuka, H., Imai, S.: Road signposts recognition system. In: The International Conference on SAE Vehicle Highway Infrastructure: safety compatibility, pp. 189–196 (1987)
5. Benallal, M., Meunier, J.: Real-time color segmentation of road signs. In: Proceedings of the IEEE Canadian Conference on Electrical and Computer Engineering (CCGEI) (2003)
6. Suen, C.Y., Zadeh, M.M., Kasvand, T.: Localization and recognition of traffic road signs for automated vehicle control systems. In: Proceedings of the SPIE Intelligent system and automated manufacturing, pp. 272–282 (1998)
7. de la Escalera, A., Moreno, L.E., Salichs, M.A., Armingol, J.M.: Road traffic sign detection and classification. IEEE Trans. Ind. Electron. **44**((6), 848–859 (1997)
8. Ritter, W., Stein, F., Janssen, R.: Traffic sign recognition using colour information. Math. Comput. Model. **22**(4-7), 149–161 (1995)
9. Ghica, R., Lu, S., Yuan, X.: Recognition of traffic signs using a multilayer neural network. In: Proceedings of the Canadian Conference on Electrical and Computer Engineering (1994)
10. Bahlmann, C., Zhu, Y., Ramesh, V., Pellkofer, M., Koehler, T.: (2005) A system for traffic sign detection, tracking and recognition using color, shape, and motion information. In: Proceedings of the IEEE Intelligent Vehicles Symposium, pp. 255–260
11. Kehtarnavaz, N., Griswold, N.C., Kang, D.S.: Stop-sign recognition based on colour-shape processing. Machin. Vis. Appl. **6**, 206–208 (1993)
12. Liu, Y.S., Duh, D.J., Chen, S.Y., Liu, R.S., Hsieh, J.W.: Scale and skew-invariant road sign recognition. Int. J. Imaging Syst. Technol. **17**((1), 28–39 (2007)

13. Piccioli, G., De Micheli, E., Parodi, P., Campani, M.: Robust method for road sign detection and recognition. Image Vis. Comput. **14**(3), 209–223 (1996)
14. Paclik, P., Novovicova, J., Pudil, P., Somol, P.: Road signs classification using the laplace kernel classifier. Pattern Recognit. Lett. **21**(13-14), 1165–1173 (2000)
15. Fang, C.Y., Chen, S.W., Fuh, C.S.: Roadsign detection and tracking. IEEE Trans. Veh. Technol. **52**((5), 1329–1341 (2003)
16. Priese, L., Klieber, J., Lakmann, R., Rehrmann, V., Schian, R.: New results on traffic sign recognition. In: IEEE Proceedings of the Intelligent Vehicles Symposium, pp. 249–254 (1994)
17. Gao, X.W., Podladchikova, L., Shaposhnikov, D., Hong, K., Shevtsova, N.: Recognition of traffic signs based on their colour and shape features extracted using human vision models. J. Vis. Commun. Image Represent. **17**((4), 675–685 (2006)
18. Fleyeh, H.: Color detection and segmentation for road and traffic signs. Proc. IEEE Conf. Cybern. Intell. Syst. **2**, 809–814 (2004)
19. Nguwi, Y.Y., Kouzani, A.Z.: Detection and classification of road signs in natural environments. Neural Comput. Appl. **17**((3), 265–289 (2008)
20. Fang, C.Y., Fuh, C.S., Yen, P.S., Cherng, S., Chen, S.W.: An automatic road sign recognition system based on a computational model of human recognition processing. Comput. Vis. Image Underst. **96**((2), 237–268 (2004)
21. Gómez, H., Maldonado, S., Jiménez, P.G., Gómez, H., Lafuente-Arroyo, S.: Goal evaluation of segmentation for traffic sign recognition. IEEE Trans. Intell. Transp. Syst. **11**(4), 917–930 (2010)
22. Aoyagi, Y., Asakura, T.: A study on traffic sign recognition in scene image using genetic algorithms and neural networks. In: Proceedings of the 1996 IEEE IECON 22nd International Conference on Industrial Electronics Control and Instrumentation **3**, 1838–1843 (1996)
23. Gavrila, D.: Multi-feature hierarchical template matching using distance transforms. In: Proceedings of the IEEE International Conference on Pattern Recognition, pp. 439–444, Brisbane, Australia (1998)
24. Douville, P.: Real-time classification of traffic signs. Real-Time Imaging, **6**(3), 185–193 (2000)
25. Cyganek, B.: Circular road signs recognition with soft classifiers. Computer-Aided Eng. **14**((4), 323–343 (2007)
26. Guibert, L., Petillot, Y., de de la Bougrenet Tochnaye, J.L.: Real-time demonstration on an on-board nonlinear joint transform correlator system. Opt. Eng. **36**((3), 820–824 (1997)
27. Hsu, S.H., Huang, C.L.: Road sign detection and recognition using matching pursuit method. Image Vis. Comput. **19**, 119–129 (2001)
28. Paclik, P., Novovicova, J., Duin, R.: Building road-sign classifiers using a trainable similarity measure. IEEE Trans. Intell. Transp. Syst. **6**(3), 309–321 (2006)
29. Ruta, A., Li, Y., Liu, X.: Real-time traffic sign recognition from video by class-specific discriminative features. Pattern Recognit. **43**, 416–430 (2010)
30. Krumbiegel, D., Kraiss, K.-F., Schrieber, S.: A connectionist traffic sign recognition system for onboard driver information. In: Proceedings of the Fifth IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems, pp. 201–206 (1992)
31. Barnes, N., Zelinsky, A., Fletcher, L.: Real-time speed sign detection using the radial symmetry detector. IEEE Trans. Intell. Transp. Syst. **9**(2), 322–332 (2008)
32. Escalera, S., Radeva, P. et al.: Fast greyscale road sign model matching and recognition. In: Vitria, J. (eds) editor Recent Advances in Artificial Intelligence Research and Development, pp. 69–76. IOS Press, (2004)

Springer

http://www.springer.com/978-1-4471-2244-9

Traffic-Sign Recognition Systems
Escalera, S.; Baró, X.; Pujol, O.; Vitrià, J.; Radeva, P.
2011, VI, 96 p. 34 illus., Softcover
ISBN: 978-1-4471-2244-9