

Framework for Machine Vision Based Traffic Sign Inventory

Petri Hienonen^{1,2}, Lasse Lensu¹, Markus Melander^{2,3},
and Heikki Kälviäinen¹(✉)

¹ Machine Vision and Pattern Recognition Laboratory (MVPR),
School of Engineering Science, Lappeenranta University of Technology (LUT),
Skinnarilankatu 34, 53850 Lappeenranta, Finland
{lasse.lensu,heikki.kalviainen}@lut.fi

² Vionice Ltd., Brahenkatu 4, 53100 Lappeenranta, Finland
{petri.hienonen,markus.melander}@vionice.fi

³ Finnish Transport Agency (FTA), Laserkatu 6, 53850 Lappeenranta, Finland

Abstract. Automatic traffic sign inventory and simultaneous condition analysis can be used to improve road maintenance processes, decrease maintenance costs, and produce up-to-date information for future intelligent driving systems. The goal of this research is to combine automatic traffic sign detection and classification with traffic sign inventory and condition analysis. This paper proposes a complete machine vision framework for the purpose and presents the results of its performance evaluation with three datasets: Traffic Signs Dataset, and two datasets collected for this research. The experimental results show that the system is able to detect, locate, and classify almost all the traffic signs, and is a suitable platform for traffic sign condition analysis.

Keywords: Traffic sign inventory · Detection · Classification · Localization · Distributed asset management · Condition analysis · Machine vision · Image processing

1 Introduction

Road maintenance operations and contracting require up-to-date traffic sign inventories, and the importance of accurate inventory information will even increase with the widespread deployment of intelligent transportation systems. Such inventories contain information on the sign type, location, direction, and condition. Compiling and keeping the inventories up-to-date requires a considerable amount of work. For example, in Finland the traffic signs are supposed to be inventoried every five years since missing or damaged signs cause danger to road users. The Finnish Transport Agency (FTA)¹ is responsible for approximately 78,000 km of highways. In total the Finnish road network is approximately 454,000 km long. FTA aims to automate the inventory and condition

¹ <http://www.liikennevirasto.fi/web/en>.

analysis of the traffic signs to be performed continuously during the normal road maintenance. The system would improve services to citizens and decrease road maintenance costs by real-time machine vision based monitoring of the signs. Currently, traffic sign inventory is maintained manually which is slow, laborious, and error prone.

A machine vision based solution that updates the inventory during normal road maintenance offers benefits that cannot be achieved with current processes: close to real-time asset management, increased road security, improved competitive bidding processes, objective evaluation of contracts, and efficient information management for intelligent driver assistance systems. This article presents the machine vision methods and implementation of an automatic system for traffic signs inventory and simultaneous traffic sign condition analysis. The first study by the authors has been published in [1].

A practical solution for data collection is to mount a digital camera into in-service road maintenance vehicles to provide an efficient method for monitoring the entire road network and its traffic signs on a daily basis and with relatively low costs. Automatic traffic sign detection and classification have been previously studied from the viewpoint of self-driving vehicles and driver assistance systems, but systems combining the inventory and condition analysis have not been presented. In addition, condition analysis of traffic signs has not been comprehensively studied before, with the exception of automatic reflectance assessment. The work differentiates between traffic signs and sign posts, the latter of which are not considered in this study. Moreover, two novel datasets were collected in winter conditions and annotated for the evaluation and for traffic sign condition analysis.

2 Machine Vision Based System

Road maintenance vehicles traverse the roads often, especially in the wintertime. Therefore, the approach harnessing the vehicles for data collection can provide comprehensive analysis of the road network, and at the same time, acquire multiple images of individual signs for further analysis. There are two constraints affecting the proposed system: the system should be easy to install on existing road maintenance vehicles, and it should be operable on low-cost mobile hardware.

Automatic Traffic Sign Inventory (TSI) and condition analysis consists of three high level tasks: (i) Traffic Sign Recognition (TSR): Detection (TSD) and Classification (TSC), (ii) sign location estimation, and (iii) condition analysis. TSD is the problem of finding the traffic sign from an image which is a binary classification between traffic signs and the background. TSC is a multi-class classification problem where the previously detected traffic sign patches are classified as sign types to determine the class label. Common-use cases for TSR are autonomous driving, assisted driving, and mobile traffic sign mapping.

2.1 Related Work

Research mainly focuses on TSR [2] or TSD only [3,4], usually on a certain subclass of signs, for example, speed limit signs [5]. The survey [5] conducted into TSR shows that comparing the methods is difficult. Different studies typically use different comparison metrics and data, consider the subsequent tasks of TSD [6], TSC and tracking, or only part of the processing tasks. A system for inventory purposes in Spain was proposed in [7], containing 51 sign types. In [8] 176 different sign types were considered exploiting street-level panoramic images where promising results for large-scale automated surveying were reported. These studies differ from our work since they do not consider GPS positioning, sign condition analysis, and winter weather conditions.

Recently, three large and publicly available traffic sign datasets for the detection and classification have been released: Belgian KUL traffic sign detection and KUL traffic sign classification benchmark 2011 datasets [9], Traffic Signs Dataset [10,11], and German GTSD and GTSRB datasets [2,4], but no datasets related to the traffic sign condition analysis exist. The German datasets were introduced for two competitions that benchmark different TSR methods which were used for selecting a TSR algorithm for the system presented in this article. The camera and installation setup need to be selected: either a single camera, a dual camera [9,12], or specialized equipment such as infrared cameras [13].

2.2 Environment

Traffic signs are standardized, and white or yellow is the second color of prohibitory signs. In Finland, Sweden, Poland and Iceland, the color is yellow for better visibility in snowy landscapes. In Finland, traffic signs come in three sizes (width of a side): small (400 mm), medium (640 mm) and large (900 mm).

TSR algorithms have to cope with natural, complex, and dynamic scenes (see Fig. 1). The appearance of traffic signs is affected by variations in lighting conditions due to, e.g., shadows, clouds and direct sunlight. Colors in captured images depend on the ambient light spectrum (daylight, headlights, infrastructure lighting) and viewing geometry (angle, distance). The color and reflectivity of traffic signs fade with time, and signs can be damaged, misaligned, or obstructed. Other



Fig. 1. Operating environment. (Color figure online)

objects with colors and shapes similar to traffic signs, such as advertisements, certain parts of vehicles, and buildings, make the recognition task more difficult.

2.3 Proposed System

The proposed system is presented in Fig. 2. A single camera setup was chosen for this implementation. The core components of the system are marked as gray. At the general level, object detection, object classification, and condition analysis all contain feature extraction, feature post-processing, and classification. The core components of the system and their purposes are as follows:

1. **Camera with known location:** A camera captures either still images or video, and the corresponding GPS location data is stored. For example, the camera can be integrated into a mobile phone with a built-in GPS receiver.
2. **Image pre-processing:** The captured images are processed to better suit to the next steps, including, e.g., motion de-blurring and color correction.
3. **Object detection:** The task of the detection module is to locate the objects (traffic signs) in the 2D image. The module outputs the location of possible signs in the image and the reliability of the detection, given by the classifier.
4. **Object classification:** The located signs (objects) are classified based on a predefined set of sign types.
5. **Localization:** The detections are combined with known camera parameters to estimate the sign location in with respect to the camera location.
6. **Trajectory prediction:** Information on the localized signs is further refined by predicting the space-time trajectories, used as a priori information for the next detection round. The relationship between the trajectories and the detections is asymmetric, and new detections can occur and the old ones vanish from the view in matching between the time steps.
6. **Global location assessment:** The sign locations are mapped onto the world coordinate system using interpolated/extrapolated GPS coordinates and 3D-localized signs. The GPS locations are estimated with ellipsoid projections on a surface.
7. **Condition evaluation:** The condition of the signs is analyzed. The sign image is segmented from the background, image features related to the sign condition are extracted, and the condition category is determined.

3 Traffic Sign Recognition and Localization

The algorithms chosen for the system are shown in Fig. 3. Object detection has to be done fast whereas there is more time for object classification, considering variations of the environment affecting the appearance of the traffic signs. During TSD, a binary classifier is used to discriminate traffic signs from the background. TSC is a multi-class classification problem where the previously detected traffic sign patches are classified as sign types to determine the class label.

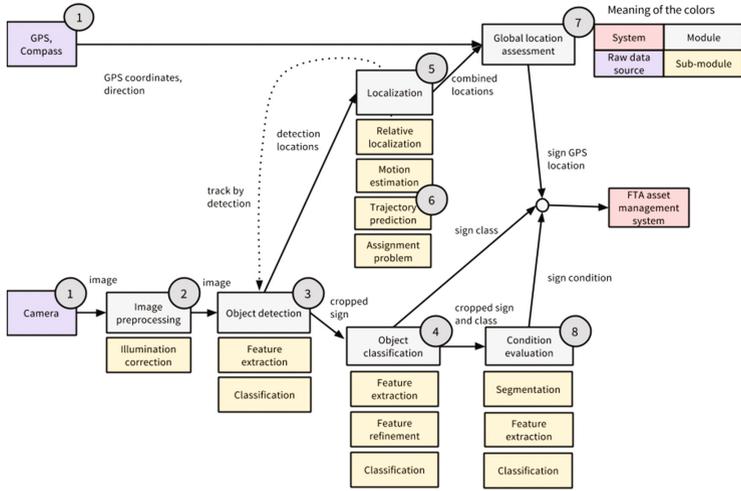


Fig. 2. The proposed system and the relations between different components.

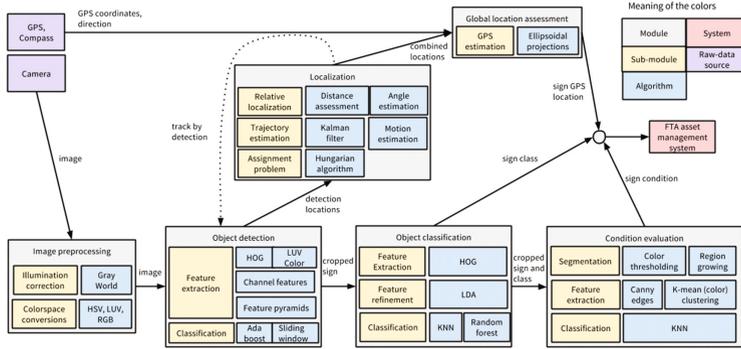


Fig. 3. Algorithms of the proposed system.

The TSR approaches utilize three prominent features: color, inner shape, and contour shape. Due to diverse natural lighting conditions, the use of color information is difficult and many heuristics have been proposed to make the color useful feature [14, 15]. Related to shape, there are two general approaches: (i) parametrized methods such as Hough Transform (HT) for different shapes and (ii) model-based methods utilizing Haar-like features [16].

Dense HOG features have been used to capture the shape features of traffic signs [2, 4, 17]. The research conducted in [18] shows a comparison of HOG feature parameters, different scales, and their performance as basic shape features for TSR. It has been demonstrated that the color is an important feature in detection, but less important in classification. Dealing with high dimensionality reduction techniques or sparse modeling can be used as PCA and KNN [19].

LDA directly deals with the classes, whereas PCA tries to find the principal components from all the data without taking into account the class structure. Sparse representation and encoding methods, such as SPM [20] and LLC [21] have been used for traffic signs resulting in state-of-the-art performance.

3.1 Traffic Sign Detection

The detection can be performed in two ways: the computationally complex sliding window [18] approach and the computationally inexpensive color thresholding method [22–24]. The sliding window method was chosen for the proposed system since of good results in pedestrian detection [25] and traffic sign detection [4, 18, 26]. The detection is done by defining a score at different sliding window positions and scales in the image. If the score is greater than a threshold, the area is considered as a detected bounding box [26]. Due to the large number of classifications required per image, a robust classifier is needed. The AdaBoost classifier [27] has been shown to perform well in such situations [18, 28], and is fast to train, and performs especially well with large feature sets. The ACF approach [25] is used for feature pyramids in the proposed system.

3.2 Traffic Sign Classification

Similar to TSD, HOG features are used to describe the shape of the traffic sign. Numerous classification algorithms have been used for TSC, including KNN [18], Random Forests [2, 17], Neural Networks (NN) [2, 29], and different variations of SVM [18]. Image features and their representation have been shown to be more important for the performance than selecting a specific classifier [18].

3.3 Location Assessment

The accurate location of the traffic sign is determined in two steps: (i) The location of the sign is first estimated relative to the car and (ii) the relative location is transformed into the global coordinate system. In the localization, the location of the sign relative to the camera has to be determined. The side length of a traffic sign is known which can be used for the determination of the location accuracy. After the distance to the sign is known, the estimate has to be refined using the geometry of the scene. For global location assessment, Karney's implementation for geodesic calculations [30] was used. The accuracy of GPS can be improved to be within 1–2 m in urban settings using sensor fusion, and below 0.10 m by using the Real Time Kinematics correction signal [31].

3.4 Traffic Sign Condition Analysis

The machine vision approach can be divided into three parts: (i) definition of the exact location of the sign surface (segmentation), (ii) selection and extraction of features that correlate with traffic sign condition, and (iii) determination of the

condition of the traffic sign with classification. Next, the current human-performed condition analysis criteria are presented, followed by a machine vision solution for the same criteria. The machine vision approach for condition analysis was introduced in [1] and the comprehensive results are to be reported in the future.

4 Experimental Results

4.1 Datasets and Evaluation

The implemented system was tested with three datasets: Traffic Signs Dataset published by the Computer Vision Laboratory at Linköping University, Sweden [10, 11] (Dataset 1), Finnish Winter Dataset (Dataset 2) and Lappeenranta Road Signs Dataset (Dataset 3). The two latter ones were collected in this research.

Dataset 1 was used in the TSD and TSC experiments since the Swedish and Finnish traffic signs are very similar. Dataset 1 consists of 20,000 images from video sequences of which around 20% have been annotated. The dataset consists of continuous video sequences, recorded during the summer and a single tour, in daytime in varying illumination conditions and different driving scenarios (rural, urban, and highway). The dataset contains 16 sign classes as shown in Table 1. Annotations smaller than 25×25 pixels were ignored due to too low resolution.

Dataset 2 contains approximately 20 h of video material for the system testing in difficult environmental conditions and to demonstrate the functionality of the

Table 1. Annotated signs available in the Traffic Signs Dataset.

Name	Sign category	Count
Priority road	Priority	496
Give way	Priority	145
Stop and give way	Priority	4
No standing or parking	Prohibitory	79
No parking	Prohibitory	102
Allowed direction left or right	Mandatory	27
Pass this left of right	Mandatory	347
Speed limit 30	Prohibitory	21
Speed limit 50	Prohibitory	163
Speed limit 60	Prohibitory	21
Speed limit 70	Prohibitory	223
Speed limit 80	Prohibitory	112
Speed limit 90	Prohibitory	30
Speed limit 100	Prohibitory	100
Speed limit 110	Prohibitory	29
Speed limit 120	Prohibitory	23
Total		1922

system shown in Fig. 5. The dataset was collected since suitable videos captured in snowy winter conditions with GPS location information were not available. The videos were recorded in an urban environment using a Garmin Virb Elite camera (with GPS) attached to a road maintenance vehicle inside the driver’s cabin. The camera contains a CMOS sensor with a spatial resolution of 1980×1080 pixels and records 30 fps. The higher resolution, the better for the condition analysis. The camera’s horizontal field of view is 151° . The exposure time for the video is controlled automatically based on average lighting of the view.

Dataset 3 consists of 325 still images for traffic sign condition analysis. The dataset contains 397 condition and class-annotated traffic signs. The dataset was introduced in [1].

4.2 Traffic Sign Detection

The approach was evaluated using Dataset 1. Common evaluation criteria include the number of True Positives (TP), False Negatives (FN), and False Positives (FP). The accuracy of detections can be measured based on the relative overlap between the Ground Truth (GT) bounding boxes (BBs) and the detected BBs. This measure is known as Intersection of Union (IoU) [26]. The performance of the detection accuracy is measured using the two main indicators False Positives Per Image (FPPI) and the miss rate. FPPI describes how many false positives are found on the average per image. The miss rate describes what percentage of the real signs are missed (FN) during the detection: $1 - \text{True Positive Rate (TPR)}$ where $\text{TPR} = \text{TP}/(\text{TP} + \text{FN})$. In the tests, the detection is considered to be successful when IoU is more than 0.5 which is considered to be good enough [26].

The results are shown in Fig. 4 and Table 2. The detectors were trained for each sign category separately and for all the three sign categories. Figure 4

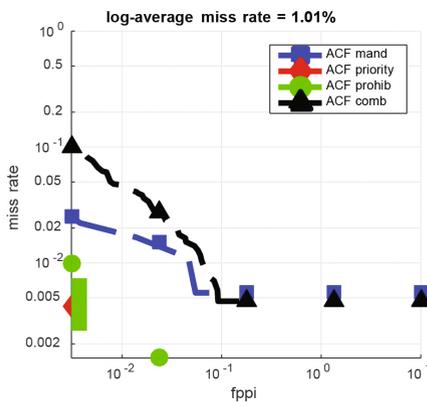


Fig. 4. Logarithmic FPPI/miss rate curve of the sign detection on Dataset 1 using Adaboost, and the HSV and HOG detectors for three subcategory detectors and the combined detector with IoU 0.7.

Table 2. TSD using Dataset 1 and three separately trained detectors.

Type	Mand.	Prior.	Prohib.	Combined
Training	193	353	520	1066
Test	181	292	383	856
FN	1	0	0	4
Error	0.006	0.014	0.024	

presents the miss rates on a logarithmic scale, showing very good detection performance.

4.3 Classification

The evaluation was also done using Dataset 1. The classification performance was computed by comparing the classification results to GT. The performance was evaluated using the HOG and grayscale features. Grayscale features are just normalized gray-level values of the image patch. Dimension reduction techniques tested were LDA and PCA, and the KNN, Naive Bayes and Random Forests classifiers. For KNN, $k = 5$ was selected by experimentation. PCA was used so that the set of features explains 95% of the variance in the features. The depth of the forest was limited to 64. The classification results, computed by using the 10-fold cross validation, are presented in Table 3.

Table 3. TSC using Dataset 1.

Feature	Dim. red.	Classifier	Mean error
Gray-scale	PCA	KNN	0.1957
Gray-scale	LDA	KNN	0.1251
HOG	None	KNN	0.0295
HOG	PCA	KNN	0.0295
HOG	LDA	KNN	0.0146
Gray-scale	None	Random Forest	0.1387
HOG	None	Random Forest	0.0424
HOG	LDA	Random Forest	0.0493
HOG	PCA	Random Forest	0.1215
HOG	LDA	Naive Bayes	0.0187
HOG	PCA	Naive Bayes	0.0586

4.4 Distance and Location Evaluation

During the collection of data, no GT data for the assessment of localization accuracy was collected. The distance estimation to the sign is one of the error-prone parts of traffic sign localization. The distance estimation accuracy was evaluated by comparing six independent separate images taken on six different traffic signs specifically for the purpose of computing distance estimation accuracy with GT measured using a laser distance meter. Images were taken with the same Garmin Virb camera as described earlier. The camera manufacturer does not give the exact focal length or width of view parameters for the camera, and they were approximated by using calibration. The camera performs lens correction for images automatically. The BBs for calculating the heights were placed manually, and the averages of the height and the width were computed for the projection. All the tested signs were 640 mm wide. The absolute and relative average errors were -0.31 m and $-1,9\%$ only.

5 Discussion

The framework for traffic sign inventory with simultaneous sign condition analysis were proposed and tested using three different datasets. Two datasets were collected during the research project, the one for condition analysis and the other for the detection, localization, and classification in difficult winter conditions. TSD uses the rigid, HOG+color feature detector which reached detection of 96.00% of the signs and runs around 15fps on a 640×480 frame on a mobile i5-4200u processor. The best results for TSC were obtained using a HOG+LDA+KNN combination, classifying 98.55% of the signs correctly. The

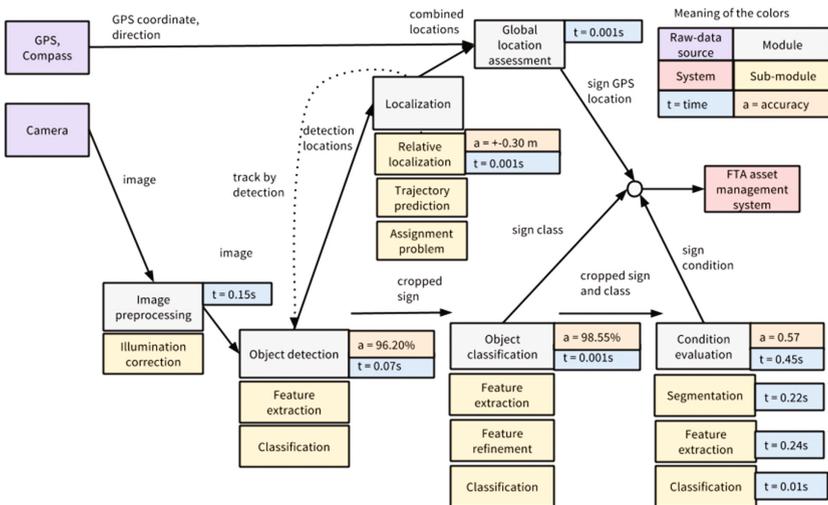


Fig. 5. Implemented system performance on a i5-4200 cpu.

mean error of the condition analysis phase per sign was 0.583 [1]. GPS localization is accurate, mostly because the signs are detected correctly. Figure 5 shows the performance of the proposed system.

In the future, system performance has to be verified in practice, including challenging weather conditions. To evaluate the performance in extreme conditions (fog, snowfall, low light, and rain) and to estimate the localization error more accurately, a new dataset is needed. This can be developed for the proposed system by extending Dataset 2. The TSI process can be further improved by adding more information specific to the environment and traffic signs, such as the movement of the vehicle and scene understanding. The TSD and TSC results can be combined with information from multiple frames. An improved classification model could produce appropriate a posteriori probabilities for TSC. Moreover, experiments can be extended by utilizing, for example, Convolutional Neural Networks (CNN) [32].

6 Conclusion

This research was a first step in automating and combining traffic sign condition analysis with TSI and, thus, reducing road maintenance costs and increasing quality of the traffic sign inventory data. The research shows that machine vision solutions are accurate enough for implementing a TSI system for automatic asset management. The sign detector and classifier performed very well and close to real-time. The best results were obtained using HOG and color features in the detection, and the HOG-LDA-KNN combination in the classification. In the future, the system should be tested more in challenging winter conditions.

Acknowledgements. The authors would like to thank the Finnish Transport Agency for funding of the TrafficVision research project.

References

1. Hienonen, P.: Automatic traffic sign inventory and condition analysis. Master's thesis, Lappeenranta University of Technology, Finland (2014)
2. Stallkamp, J., Schlipsing, M., Salmen, J., Igel, C.: Man vs. computer: benchmarking machine learning algorithms for traffic sign recognition. *NN* **32**, 323–332 (2012)
3. Baro, X., Escalera, S., Vitria, J., Pujol, O., Radeva, P.: Traffic sign recognition using evolutionary adaboost detection and forest-ECOC classification. *IEEE ITS* **10**(1), 113–126 (2009)
4. Houben, S., Stallkamp, J., Salmen, J., Schlipsing, M., Igel, C.: Detection of traffic signs in real-world images: the German traffic sign detection benchmark. In: *Proceedings of IJCNN* (2013)
5. Mogelmose, A., Trivedi, M., Moeslund, T.: Vision-based traffic sign detection and analysis for intelligent driver assistance systems: perspectives and survey. *IEEE ITS* **13**, 1484–1497 (2012)
6. Bahlmann, C., Zhu, Y., Ramesh, V., Pellkofer, M., Koehler, T.: A system for traffic sign detection, tracking, and recognition using color, shape, and motion information. In: *Proceedings of IEEE IV* (2005)

7. Maldonado-Bascon, S., Lafuente-Arroyo, S., P. Siegmann, H.G.M., Acevedo-Rodriguez, F.: Traffic sign recognition system for inventory purposes. In: Proceedings of IEEE IVS (2008)
8. Hazelhoff, L., Creusen, I.M.: Exploiting street-level panoramic images for large-scale automated surveying of traffic signs. *MVA* **25**, 1893–1911 (2014)
9. Timofte, R., Zimmermann, K., Van Gool, L.: Multi-view traffic sign detection, recognition, and 3D localisation. *MVA* **25**, 633–647 (2011)
10. Larsson, F., Felsberg, M.: Using Fourier descriptors and spatial models for traffic sign recognition. In: Heyden, A., Kahl, F. (eds.) SCIA 2011. LNCS, vol. 6688, pp. 238–249. Springer, Heidelberg (2011). doi:[10.1007/978-3-642-21227-7_23](https://doi.org/10.1007/978-3-642-21227-7_23)
11. Larsson, F., Felsberg, M., Forssen, P.E.: Correlating Fourier descriptors of local patches for road sign recognition. *IET Comput. Vis.* **5**, 244–254 (2011)
12. Gu, Y., Yendo, T., Tehrani, M., Fujii, T., Tanimoto, M.: Traffic sign detection in dual-focal active camera system. In: Proceedings of IEEE IV (2011)
13. Gonzalez, A., Garrido, M., Llorca, D., Gavilan, M., Fernandez, J., Alcantarilla, P., Parra, I., Herranz, F., Bergasa, L., Sotelo, M., Revenga de Toro, P.: Automatic traffic signs and panels inspection system using computer vision. *IEEE ITS* **12**, 485–499 (2011)
14. Houben, S.: A single target voting scheme for traffic sign detection. In: Proceedings of IEEE IV (2011)
15. Fleyeh, H.: Color detection and segmentation for road and traffic signs. In: Proceedings of IEEE CIS (2004)
16. Viola, P., Jones, M.J.: Robust real-time face detection. *IJCV* **57**, 137–154 (2004)
17. Zaklouta, F., Stanculescu, B., Hamdoun, O.: Traffic sign classification using k-trees and random forests. In: Proceedings of IJCNN (2011)
18. Mathias, M., Timofte, R., Benenson, R., Gool, L.J.V.: Traffic sign recognition - how far are we from the solution? In: Proceedings of IJCNN (2013)
19. Martinez, A.M., Kak, A.C.: PCA versus LDA. *IEEE PAMI* **23**, 228–233 (2001)
20. Zhu, Y., Wang, X., Yao, C., Bai, X.: Traffic sign classification using two-layer image representation. In: Proceedings of ICIP (2013)
21. Lu, K., Ding, Z., Ge, S.: Sparse-representation-based graph embedding for traffic sign recognition. *IEEE ITS* **13**, 1515–1524 (2012)
22. de la Escalera, A., Moreno, L., Salichs, M., Armingol, J.: Road traffic sign detection and classification. *IEEE IE* **44**, 848–859 (1997)
23. Fleyeh, H.: Shadow and highlight invariant colour segmentation algorithm for traffic signs. In: Proceedings of CCIS (2006)
24. Broggi, A., Cerri, P., Medici, P., Porta, P., Ghisio, G.: Real time road signs recognition. In: Proceedings of IEEE IV (2007)
25. Dollar, P., Appel, R., Belongie, S., Perona, P.: Fast feature pyramids for object detection. *IEEE PAMI* **36**, 1532–1545 (2014)
26. Dollar, P., Wojek, C., Schiele, B., Perona, P.: Pedestrian detection: an evaluation of the state of the art. *IEEE PAMI* **34**, 743–761 (2012)
27. Breiman, L.: Random forests. *Mach. Learn.* **45**, 5–32 (2001)
28. Dollar, P., Tu, Z., Perona, P., Belongie, S.: Integral channel features. In: Proceedings of BMVC (2009)
29. Sermanet, P., LeCun, Y.: Traffic sign recognition with multi-scale convolutional networks. In: Proceedings of IJCNN (2011)
30. Karney, C.F.: Algorithms for geodesics. *J. Geod.* **87**, 43–55 (2013)
31. Martí, E.D., Martín, D., García, J., de la Escalera, A., Molina, J.M., Armingol, J.M.: Context-aided sensor fusion for enhanced urban navigation. *Sensors* **12**, 16802–16837 (2012)
32. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *Nature* **521**, 436–444 (2015)