Robotics and Autonomous Systems 62 (2014) 16–24

Contents lists available at ScienceDirect

Robotics and Autonomous Systems

journal homepage: www.elsevier.com/locate/robot

Real-time traffic sign recognition in three stages

Fatin Zaklouta *, Bogdan Stanciulescu

Robotics Center, Mines ParisTech, 60 Bd Saint-Michel, 75006 Paris, France

A R T I C L E   I N F O

Article history:
Available online 8 August 2012

Keywords:
Traffic Sign Recognition (TSR)
Advanced Driver Assistance Systems (ADAS)
Intelligent transport systems
Color segmentation
Feature space reduction
German Traffic Sign Recognition Benchmark (GTSRB)

A B S T R A C T

Traffic Sign Recognition (TSR) is an important component of Advanced Driver Assistance Systems (ADAS). The traffic signs enhance traffic safety by informing the driver of speed limits or possible dangers such as icy roads, imminent road works or pedestrian crossings. We present a three-stage real-time Traffic Sign Recognition system in this paper, consisting of a segmentation, a detection and a classification phase. We combine the color enhancement with an adaptive threshold to extract red regions in the image. The detection is performed using an efficient linear Support Vector Machine (SVM) with Histogram of Oriented Gradients (HOG) features. The tree classifiers, K-d tree and Random Forest, identify the content of the traffic signs found. A spatial weighting approach is proposed to improve the performance of the K-d tree. The Random Forest and Fisher's Criterion are used to reduce the feature space and accelerate the classification. We show that only a subset of about one third of the features is sufficient to attain a high classification accuracy on the German Traffic Sign Recognition Benchmark (GTSRB).

1. Introduction

Advanced Driver Assistance Systems (ADAS) play an important role in enhancing car safety and driving comfort. Some of their components include navigation systems to provide directions as well as up-to-date traffic information and vision-based systems such as lane departure warning systems and Traffic Sign Recognition (TSR).

The latter enhances traffic safety by informing the driver of speed limits or possible dangers such as icy roads, imminent road works or pedestrian crossings. TSR algorithms face three main difficulties:

i. poor image quality due to low resolution, bad weather conditions, over- or under-illumination,
ii. rotation, occlusion and deterioration of the signs,
iii. limited memory and processing capacities in real-time applications such as ADAS.

Some examples of the first two difficulties are illustrated in Fig. 1.

We start this paper by giving an overview of the existing TSR approaches in Section 2. Our three stage approach is described in Section 3 and the overall performance of the proposed system is presented in Section 4. The effect of the feature selection on the performance of the traffic sign classification is evaluated in Section 5. We conclude this paper and present an outlook on further possible improvements in Section 6.

* Corresponding author.
E-mail addresses: fatin.zaklouta@ensmp.fr, fatin.zaklouta@mines-paristech.fr (F. Zaklouta), bogdan.stanciulescu@mines-paristech.fr (B. Stanciulescu).

0921-8890/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.robot.2012.07.019

2. Existing approaches

Most traffic recognition algorithms divide the problem into three stages:

i. a rough segmentation to determine the location of the signs
ii. category determination and
iii. candidate classification to identify the content of the extracted traffic signs using various machine learning techniques.

The segmentation phase is not a mandatory step. However, it is often deployed in approaches using color images. This section gives a brief overview of some of the techniques used in the three TSR stages.

Some approaches, such as [1,2] reduce the memory and processing requirements by using tracking. The candidates found are tracked over several frames to reinforce the decision made or to reduce the search space in the subsequent frame and therewith accelerate the performance. Some implementations, such as [3], only examine every n-th frame, to speed up the overall system.

2.1. Determining the location using segmentation

As traffic signs need to be easily perceivable, they are brightly colored in red, blue, yellow. Hence, the detection is often based on a pre-segmentation of the image to reduce the search space and retrieve Regions of Interest (ROIs).

Since the direct thresholding of the RGB channels is sensitive to changes in illumination, the relation between the RGB (Red Green Blue) colors is often used. In [3], the color enhancement is used to extract red, blue and yellow blobs. This transform emphasizes the

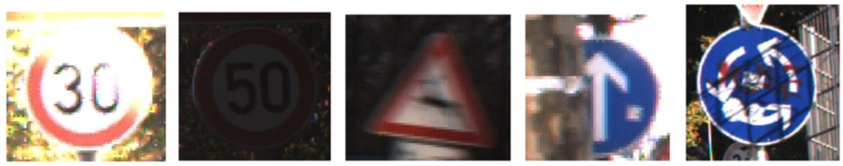

Fig. 1. Examples for difficulties facing the Traffic Sign Recognition (TSR) task: over-illumination, under-illumination, rotation, occlusion and deterioration of the signs.

pixels where the given color channel is dominant over the other two in the RGB color space.

In [4], chromatic and achromatic filters are used to extract the red rims and the white interior of the speed limit and warning traffic signs respectively. The HSI model (Hue Saturation Intensity) is used in [5] as it is invariant to illumination changes. Empirically determined fixed thresholds define the range of each HSI channel in which lie the red and blue traffic sign candidates. It is pointed out in [6], however, that HSI is computationally expensive due to its nonlinear formulae.

2.2. Category determination using shape detection

Several detection algorithms are based on edge detection, making them more robust to changes in illumination. These are shape-based methods which exploit the invariance and symmetry of the traffic signs. They are however often sensitive to the quality of the preprocessing, such as edge extraction or color segmentation.

Franke et al. [7] use Distance Transform (DT) and Template Matching (TM) to detect circular and triangular signs. Similarly, Ruta et al. [3] use the Color Distance Transform, where a DT is computed for every color channel separately. The advantage of matching DTs over edge images is that the similarity measure is smoother and robust to slight rotations. It is, however, sensitive to affine rotations and occlusions.

In [4], four Support Vector Machines (SVMs) are trained on the Distance to Border (DtB) vectors to classify the shape of an extracted candidate ROI. In [1], the FFT signatures of candidate signs are used to detect relevant shapes. This feature is robust to rotation and scaling, yet not to occlusion and deterioration.

The Hough Transform is also widely used to detect regular shapes such as circles and triangles [8,9,5]. The processing time is decreased by the simpler Radial Symmetry Detector [10], yet it is limited to circular traffic signs. Ruta et al. [9] refine the Hough Transform result using the Confidence-Weighted Mean Shift to eliminate redundant detections. The Hough transform is combined with an iterative process of median filtering and dilation to refine the candidate set in [5].

Many recent approaches use gradient orientation information in the detection phase. Gao et al. [11] classify the candidate traffic signs by comparing their local edge orientations at arbitrary fixation points with those of the templates. In [12], Edge Orientation Histograms are computed over shape-specific subregions of the image. In [13,14], the Regions of Interest (ROI) obtained from color-based segmentation are classified using the Histogram of Oriented Gradients (HOG) feature. To integrate color information in the HOG descriptor, Creusen et al. [15] concatenate the HOG descriptors calculated on each of the color channels. The advantages of this feature are its scale-invariance, the local contrast normalization, the coarse spatial sampling and the fine weighted orientation binning.

2.3. Classification techniques

The classification techniques used to determine the content of the detected traffic signs belong to two general categories: learning and Nearest Neighbor approaches. The learning consists of finding an optimal separation between two or more classes. It includes, amongst others, one-vs-all SVM classifiers, Adaboost and Neural Networks. The Nearest Neighbor approaches seek the most similar existing training sample to the given unknown. They include template matching and tree classifiers such as K-d trees and Random Forests.

Xie et al. [13] train the HOG descriptors of each class using one-vs-all SVMs. The Forest-ECOC Adaboost classifiers achieved high performance rates in [16]. Multi-layer Perceptrons (MLPs) yield high accuracy rates in [17,18]. They also achieve low false positive rates when identifying the characters in speed limit signs in [19]. The performance of the Neural Networks is improved by pre-selecting the color–shape features using PCA and Fisher Linear Discriminant in [20].

The K-d tree is used in [5] to identify the content of the sign. The Random Forests used in [21] outperform the one-vs-all SVMs on their dataset. They also generate an ensemble of SVMs using bagging and a boosted ensemble of Naive Bayes classifiers, which improve the performance of the non-ensemble version. However, these do not outperform the Random Forests.

The advantage of the tree classifiers is that they are easy to train and update. The learning approaches tend to be biased towards over-represented classes and generally require a large training set. The Best Bin First Approximate Search [22] enables a rapid search in the K-d tree. The randomness and ensemble voting of the Random Forests make them robust to outliers and unbalanced datasets.

3. Our approach

We propose a speed and danger traffic sign recognition approach consisting of three stages: (i) segmentation, (ii) shape detection and (iii) classification. These are illustrated in Fig. 2. The image segmentation reduces the search space to the red regions that potentially contain a traffic sign. We improve the red color enhancement approach used by Ruta et al. in [3] by introducing a global threshold. The circular and triangular signs are detected using one linear SVM/HOG detector each. The candidates found are further efficiently classified using tree classifiers. We further introduce spatial weighting techniques to improve the accuracy and feature space reduction for resource optimization.

The performance of each of these three stages is compared to the state-of-the-art techniques. The classification step was also put to the test at the live German Traffic Sign Benchmark Challenge [23], where our approach was ranked 3rd in terms of accuracy and proved to be suitable for embedded systems as it runs in real-time and is resource efficient.

3.1. Image segmentation

Potential ROIs are extracted in the segmentation phase. We implement the red color enhancement, proposed by Ruta et al. in [3,24], the chromatic filter [4] and introduce the morphological filters. After applying one of these filters, the image is thresholded using an empirically determined threshold or the adaptive threshold which we propose. The resulting binary mask is used to