Road Sign Detection and Recognition using Mean Shift and Fourier Descriptors

Thesis submitted in partial fulfilment of the requirement for the degree of

Bachelor of Computer Science and Engineering

Under the Supervision of

Dr. Jia Uddin

By

Nazmus Sakib (13101069)

MD. Moshiur Rahman Aunik (12201082)

Mahmud Hassan Sabit (13101065)

Mohammad Towsif Patwary (12301021)

School of Engineering and Computer Science

August 2017

BRAC University, Dhaka, Bangladesh

Declaration

We hereby declare that this thesis is based on results obtained from our own work. Due acknowledgement has been made in the text to all other material used. This thesis, neither in whole nor in part, has been previously submitted to any other University or Institute for the award of any degree or diploma.

Signature of Supervisor **Signature of the Authors**

Dr. Jia Uddin

Nazmus Sakib(13101069)

MD. Moshiur Rahman Aunik(12201082)

Mahmud Hassan Sabit(13101065)

Mohammad Towsif Patwary(12301021)

Acknowledgement

First and foremost, we would like to thank Almighty Allah for enabling us to initiate the research, to put our best efforts and successfully conclude it.

Secondly, we submit our heartiest gratitude to our respected Supervisor Dr. Jia Uddin for his contribution, guidance and support in conducting the research and preparation of the report. Every last involvement of his, starting from instilling in us the deadliest of fears to the kindest words of inspiration has permitted us to effectively complete the paper. We are truly grateful to him.

We revere the patronage and moral support extended with love, by our parents as well as our friends. They helped us with their direct or indirect suggestions which aided in achieving our goal. We would also like to acknowledge the assistance we received from numerous resources over the Internet especially from fellow researchers' work.

Last but not the least, we thank BRAC University for providing us the opportunity of conducting this research and for giving us the chance to complete our Bachelor degree.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AUC	Area Under Curve
DFT	Density Functional Theory
FAR	False Alarm Rate
FD	Fourier Descriptor
GTSDB	German Traffic Sign Detection Benchmark
HOG	Histogram of Oriented Gradients
HIS	Hue, Saturation, Intensity
MISE	Mean Integrated Square Error
ROI	Region Of Interest
RGB	Red, Blue, Green
STS	Swedish Traffic Sign
SVM	Support Vector Machines
YUV	Luminance, blue–luminance, red–luminance

Abstract

This paper presents an automated system for road sign detection and shape recognition based on geometric attribute and color information of the road sign. Road sign detection and recognition technique is important to support a driver. It can be the part of autonomous vehicles or road sign maintenance system to maintain processing and time efficiency. This technique determines road sign information and its exact position. Failure detection by the driver of any traffic sign may increase accident risk significantly. It still have some performance limitations though it has been studied for many years. The proposed system consists of three steps: (1) The initial RGB image taken from a camera is pre-processed using Mean Shift clustering algorithm. The clustering technique determines the color information, (2) In order to discard non-interest regions the clustered image is post-processed using a random forest classifier to segment the clustered image, (3) A shape based classification is performed using Fourier Descriptors and SVM's. The proposed method is applied on both German Traffic Sign Detection Benchmark and Swedish Traffic Signs Datasets. It yields to satisfactory results when compared with other shape recognition process using machine learning technique and some state-of-the-art technique. Experiments with both German and Swedish datasets where RGB images are captured with different cameras yielded a sign detection ration between 95% and 97% and the false alarm ration between 5% and 3%. The rate of the detection depends on the size of the database.

CHAPTER 01

Introduction

Traffic sign detection and recognition is a key enabling technology to control and guide traffic to favor road safety. It regulates traffic and report traffic information to drivers on different aspects about road perambulation. This technology has the ability to detect the exact location of the road signs and what exactly the road signs stands for. Automatic traffic sign detection and recognition systems are very helpful for intelligent vehicle development and road maintenance. This task is very complex and it requires high accuracy in real time because generally traffic signs are detected from live video during fast movement of vehicles. The aim of the traffic sign detection systems are to help system users to detect a road sign and interpret. Special attention has been devoted while working on this driver support system to the robustness and flexibility of traffic sign detection and recognition. The road signs can be classified according to their shape and color. The goal of this proposed method is to detect road sign with low-light RGB image. This method recently attracted high attention of research.

1.1Motivation

Road signs detection and recognition is a very important task for intelligent driver support system. It is a key enabling technology within the framework of a variety of intelligent systems. It detects and recognizes the road signs and informs the driver from traffic situation of the roads. This technology determines where and what exactly the road sign is. This task will be very crucial in situations like in the nights or when the driver cannot see the road clearly. Road signs are generally detected from live video during fast movement of vehicles and it is a very complex task requiring high accuracies in real-time. The sign may be occluded with some objects such as a tree or may shadows make it darker. It may be corrupted or scuffed. The detection system should overcome these difficulties.Robust method for road sign detection and recognition by G. Piccioli, E. De Michelib, P. Parodi and M. Campani, color analysis, segmentation and geometrical analysis were used for Shape detection and Triangular road signs understanding[19].Spatial kernel K-harmonic means clustering for multi-spectral image segmentation by Q. Li, N. Mitianoudis and T. Stathaki, robust kernelbased KHM metric is employed to reduce the effect of outliers and noise for better image segmentation[17]. Real-time traffic sign recognition from video by class-specific discriminative features by AndrzejRuta, YongminLi,XiaohuiLiu, specific discriminative features results an efficient and faster road sign recognition system[11].

1.2Contribution Summary

In this paper, we presented a new technique for traffic sign detection using Mean Shift Clustering, Random Forest classifier and Fourier Descriptors. Mean Shift Clustering algorithm preprocessed the initial image into different clusters using color information. From the given clustered image random forest classifier determines the region of interest. Images are classified by their shape using Fourier Descriptors and SVM, which are able to classify different complex shape. Because of Fourier Descriptors invariance with respect to scaling, shift and rotation, they are used as SVM input features.

1.3Thesis Outline

- Chapter 2 provides the Background study in details including the algorithms and techniques used in the system
- Chapter 3 discusses the Literature Review of related works in this field
- Chapter 4 describes the proposed model along with implementation details
- Chapter 5 presents the results of the experiment along with performance analysis and comparisons
- Chapter 6 concludes the paper specifying the limitations and challenges while planning future development of the project

CHAPTER 02

Background Analysis

2.1 Color Segmentation

Road sign detection and recognition process works usually via color and shape information. This task requires much concentration during a long time and different types of errors can be originated with manual operating technique by any human operator for his or her poor visibility conditions[10]. Road signs are always designed with some specific colors like red, green or blue and some specific shapes like circles, rectangle or triangle[9]. The accuracy of the detection and recognition depends on the wide appearance and enough visibility of the sign of RGB image taken from the camera. Sometimes challenging working conditions occurs when the signs are captured improperly.

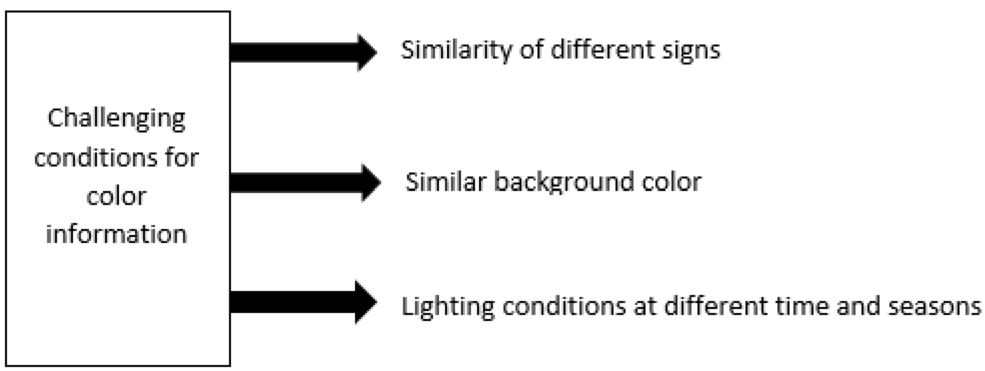


Fig. 2.1: Difficulties to identify colors of road signs

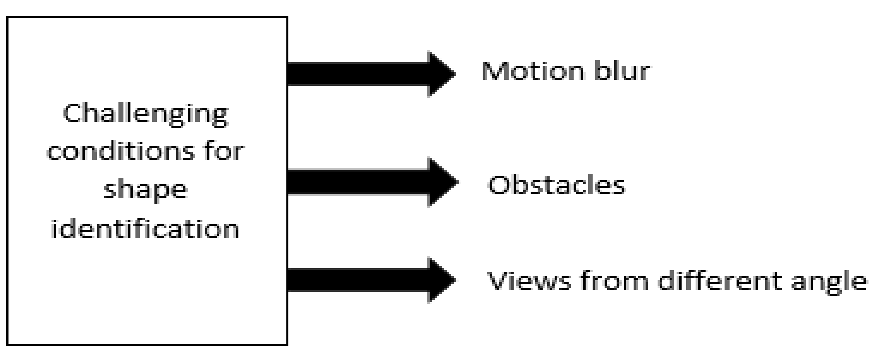
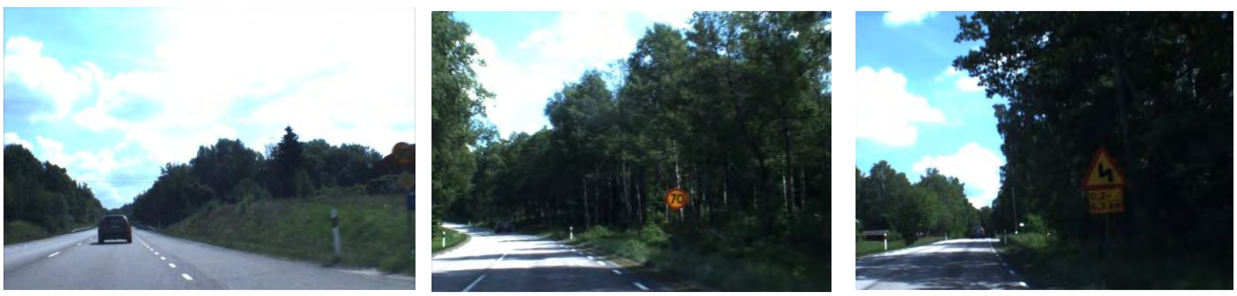


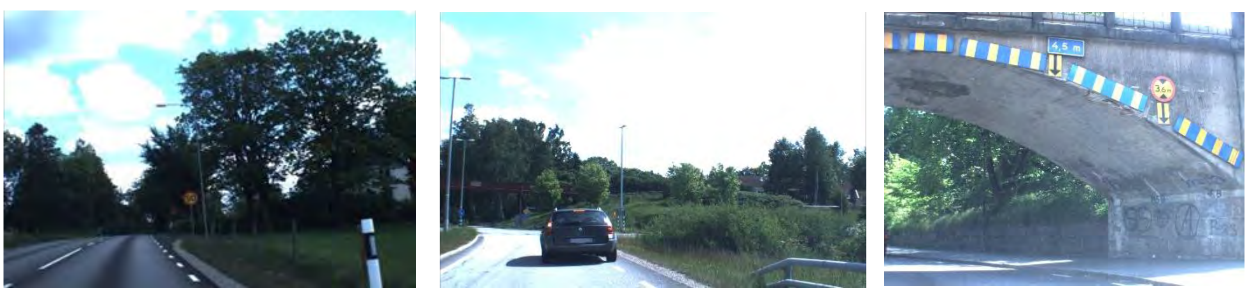
Fig. 2.2: Difficulties to identify shape of road signs

2.1.1 Mean Shift Clustering

For each data point, mean shift defines a window around it and computes the mean of data point. Then it shifts the center of window to the mean and repeats the algorithm till it convergence Mean shift is a nonparametric iterative algorithm or a nonparametric density gradient estimation using a generalized kernel approach. Mean shift is the most powerful clustering technique[27]. Mean shift is used for image segmentation, clustering, visual tracking, space analysis and mode seeking. Mean shift segmentation is an advanced and efficient technique for clustering based segmentation.



(a) (b) (c)



(d) (e) (f)

Fig. 2.3: Examples of some challenging scenes for a traffic sign detection system: (a) Motion blur; (b) Occluded by trees; (c) Lighting variation; (d) Cloudy weather; (e) Sign hiding behind the car; (f) Similar background.

2.1.2 Algorithm

- Let {xi} i=1…n be the original image points, {zi}i=1… n the points of convergence, and {Li} i=1…n a set of labels Mean Shift Segmentation[27].
- For each i = 1…n run the mean shift procedure for xi and store the convergence point in zi.
- Identify clusters {Cp } p=1…m of convergence points by linking together all zi which are closer than 0,5 from each other in the joint domain.
- For each i = 1…n assign Li = {p | zi ϵ Cp}. Eliminate spatial regions smaller than □ M pixels.

2.1.3 Random Forest Classifier

Random forests is a classification and regression algorithm originally designed for the machine learning community. This algorithm is increasingly being applied to satellite and aerial image classification and the creation of continuous field's data sets, such as, percent tree cover and biomass. Random forests has several advantages when compared with other image classification methods[28]. It is non-parametric, capable of using continuous and categorical data sets, easy to parameterize, not sensitive to over-fitting, good at dealing with outliers in training data, and it calculates ancillary information such as classification error and variable importance.

2.1.4 Algorithm

The Random Forests algorithm was developed by Leo Breiman and Adele Cutler. Random Forests grows many classification trees. Each tree is grown as follows:

- If the number of cases in the training set is N, sample N cases at random - but with replacement, from the original data. This sample will be the training set for growing the tree.
- If there are M input variables, a number M is specified such that at each node, m variables are selected at random out of the M and the best split on these m is used to split the node. The value of m is held constant during the forest growing.
- Each tree is grown to the largest extent possible. There is no pruning.

2.2Shape Detection

In the detection stage, image segmentation technique are used where it performed a deeper analysis of color information for separating color and intensity information. The detection process confirms the region of interest based on color information[23]. In the recognition stage, the signs are labeled with their symbolic information based on its color and shape.

2.2.1 Fourier Descriptor

Fourier descriptors are a way of encoding the shape of a two-dimensional object by taking the Fourier transform of the boundary, where every point on the boundary is mapped to a complex number.Fourier transform is invertible, all the information about the shape is contained in the Fourier descriptors[5]. A common thing to do with Fourier descriptors is to set the descriptors corresponding to values above a certain frequency to zero and then reconstruct the shape. The effect of this is a low-pass filtering of the shape, smoothing the boundary.

2.2.2 Algorithm

The Fourier descriptors of a shape are calculated as follows.

- Find the coordinates of the edge pixels of a shape and put them in a list in order, going clockwise around the shape.
- Define a complex-valued vector using the coordinates obtained.For example: (3,4) → 3+4i
- Take the discrete Fourier transform of the complex-valued vector.

Fourier descriptors inherit several properties from the Fourier transform.

- Translation invariance: no matter where the shape is located in the image, the Fourier descriptors remain the same.
- Scaling: if the shape is scaled by a factor, the Fourier descriptors are scaled by that same factor.
- Rotation and starting point: rotating the shape or selecting a different starting point only affects the phase of the descriptors.

2.2.3 SVM Classifier

Experimental results show that SVMs achieve significantly higher search accuracy than traditional query refinement schemes after just three to four rounds of relevance feedback.Classification of images can also be performed using SVMs[3]. This is also true of image segmentation systems, including those using a modified version SVM that uses the privileged approach as suggested by Vapnik.SVMs are effective when the number of features is quite large. It works effectively even if the number of features are greater than the number of samples. Non-Linear data can also be classified using customized hyperplanes built by using kernel trick. It is a robust model to solve prediction problems since it maximizes margin[3]. The biggest limitation of Support Vector Machine is the choice of the kernel. The wrong choice of the kernel can lead to an increase in error percentage. With a greater number of samples, it starts giving poor performances. SVMs have good generalization performance but they can be extremely slow in the test phase. SVMs have high algorithmic complexity and extensive memory requirements due to the use of quadratic programming.

2.2.4 Algorithm

In Linear Classifier, A data point considered as a p-dimensional vector(list of p-numbers) and we separate points using (p-1) dimensional hyperplane. There can be many hyperplanes separating data in a linear order, but the best hyperplane is considered to be the one which maximizes the margin i.e., the distance between hyperplane and closest data point of either class[3].

The Maximum-margin hyperplane is determined by the data points that lie nearest to it. Since we have to maximize the distance between hyperplane and the data points. These data points which influences our hyperplane are known as support vectors.

Vapnik proposed Non-Linear Classifiers in 1992. It often happens that our data points are not linearly separable in a p-dimensional(finite) space. To solve this, it was proposed to map p-dimensional space into a much higher dimensional space[3]. We can draw customized/non-linear hyperplanes using Kernel trick. Every kernel holds a non-linear kernel function.This function helps to build a high dimensional feature space. There are many kernels that have been developed.

For implementing support vector machine on a dataset, we can use libraries. There are many libraries or packages available that can help us to implement SVM smoothly. We just need to call functions with parameters according to our need.

In Python, we can use libraries like sklearn. For classification, Sklearn provides functions like SVC, NuSVC&LinearSVC.SVC() and NuSVC() methods are almost similar but with some difference in parameters. We pass values of kernel parameter, gamma and C parameter etc. By default kernel parameter uses “rbf” as its value but we can pass values like “poly”, “linear”, “sigmoid” or callable function.LinearSVC() is an SVC for Classification that uses only linear kernel. In LinearSVC(), we don’t pass value of kernel, since it’s specifically for linear classification.In R programming language, we can use packages like “e1071” or “caret”. For using apackage, we need to install it first. For installing “e1071”, we can type install.packages("e1071") in console. e1071 provides an SVM() method, it can be used for both regression and classification. SVM()method accepts data, gamma values and kernel etc.

CHAPTER 03

Literature Review

3.1 Previous Works and Technical Overview

Road sign detection technique usually works with two types of processing technique. Firstly, color based process and secondly, shape based process. Many authors have explained the segmentation process by thresholding on RGB images with preprocessing with Vector Filter algorithm or at pixel level [4]. Vector Filter algorithm is more elaborated and accurate [5]. RGB space have some limitation with lighting change. Insufficient light or excessive light hampers the segmentation process. That's why some author used color spaces that are more accurate with lighting conditions than RGB color space. The transformation of H and S component was explained in [6] where the transformation was non-linear. In [19], the author segmented the red, blue and yellow color using color enhancement technique. RGB color space was also used in [16] where the initial image was segmented at first and the region of interest was extracted using novel environment selection strategy. Many author proposed robust traffic scene understanding algorithm using RGB-D data. RGB-D is more robust than RGB color space [22]. They used spatial information to identify objects in the scene. To calculate more discriminative object they integrated depth information. In [7,8], the author used two types of color space for chromatic and achromatic scene. H and S components were used to segment images with chromatic color and RGB color space was used to threshold signs with prevalent achromatic color. The L*a*b* color space was used to extract the region of interest with the Gabor Filter process in [9] where only a*b* components are used to detect traffic signs. Authors in [17] and [15] used HSI color space for the segmentation stage where the thresholds define the range of each HSI channel. This channel contains red and blue sign candidates. The YUV space was used for thresholding segmentation in [10] and [11]. The authors [8, 9] proposed methods with HSI and YUV spaces for more elaborated segmentation process. The HSV color space is used to extract red bitmap in [21]. The algorithm performed the process to improve the accuracy of the extracted red bitmap by detecting neighboring pixels of any particular road sign. For gray-level images systems usually focus on edges detection process. In [12] the presented method was a shape based approach where a black band detector was used to highlight the regions of interest. To implement a gradient geometric model to detect triangular signs, bisector detection with a transformation for angle vertex was used in [13]. Use of geometric information in any detection method is a very common technique. To classify blobs as circles, squares and triangles the authors in [14] used a signature defined as the distance from the center of the blob to its edge as a function of the angle. Both Haar-like and Hog feature used in [20] to analyze traffic images with Adaboost and SVM classifier.

3.2 Methodology

3.2.1 Data collection

The main focuses of this proposed method are as follows. First, the mean shift clustering is used to preprocess the initial image for more accurate segmentation process and to reduce execution time. Random forest classifier is used to extract the regions of interest from the clustered image. Second, Fourier Descriptor also used to characterize the boundaries of ROIs for better accuracy. These regions are used as input to detect shape of the signs with SVMs. The proposed method is applied on GTSDB[2] and STS[1] data sets and showed more accurate result than other good methods.

3.2.2 Tools used

We have used MATLAB R2010a Simulation tools for data analysis.

CHAPTER 04

Proposed Model

4.1 System Design

The traffic sign detection system is detailed in this section. The point to be noted is the sign must be inside the image frame. The distance must be suitable as it should be distinguishable by the naked eye. The basic methodology flowchart is depicted in Figure 4. The proposed method is experimented on German Traffic Sign Detection Benchmark and Swedish Traffic Sign datasets.

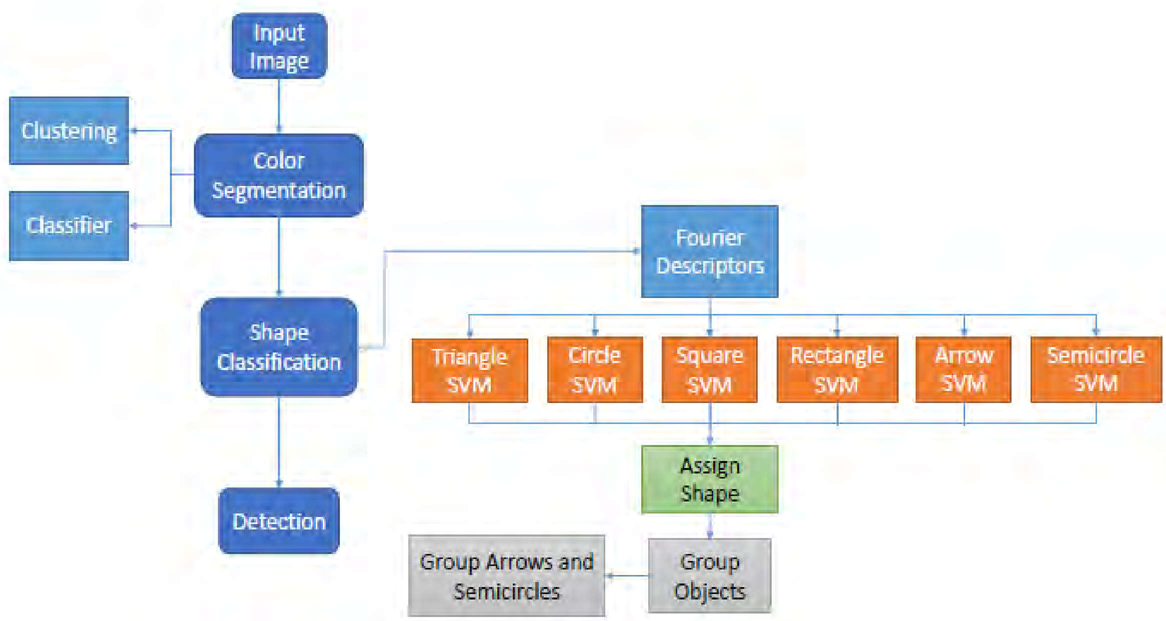


Fig 4.1. Block diagram of the proposed model

The procedure is given below:

1. Segmentation: The image is segmented based on its color information using mean shift clustering process. It divides the initial RGB image taken by camera into clusters. Only one pixel by cluster is send to random forest classifier. Processing is not performed pixel by pixel.

2. Classification: To reduce the region of interest the random forest classifier is trained with GTSDB[2] and STS[1] data sets. In this stage the 3D color value is converted to binary value using random forest classifier.
3. Shape Detection: Clustered images may extracted into several ROIS. The Fourier Descriptors is used to specify every region boundary and SVM classifiers are used to recognize shapes of signs in each ROIs.

4.2 Color Segmentation

The aim of this stage is to extract the region of interest to detect the traffic sign and to gather color information. Usually traffic signs can be differ from their surroundings by their characteristic color. Some traffic signs which are only black and white are easy to detect. Some traffic signs has chromatic color. Signs have some common properties within one target category. Some signs have triangular red borders like danger signs. Prohibitory signs have circular red borders. Mandatory signs have white arrows with blue background. Some signs have only white backgrounds like derestriction signs. The goal of the color segmentation step is find these particular color from the images. In city areas some extra signs like weather condition can affect the color segmentation process. Daylight, clouds, rotated signs or shadows can also make the segmentation process difficult. Other objects with similar color as road sign may appear. These difficulties occur frequently in urban areas. That's why color segmentation process should perform before detecting shape. Machine learning technique transforms the original image into binary one. They map the positive colors like red, blue and white to complete white and other negative colors to black. So machine learning technique is more preferable to perform color segmentation here. The random forest classifier is used here to map 3D color value to binary value.

Using GTSDB[2] and STS[1] data sets random forest classifier is trained. We performed different types of color spaces but the efficient result was found by using RGB color space. By knowing that the process is a time consuming technique, we used pixel by pixel processing to find more efficient and less time consuming outputs. In the preprocessing step mean shift clustering technique is used. Fukunaga and Hostetler in [25] developed the mean shift estimate of the gradient of a density function and mode seeking iterative associative procedure. By applying mean shift image segmentation or high quality edge preserving filtering can be obtained in spatial-ranged domain. The idea is similar with iterative shifting a fixed size window within to the average of the data points. The color information of the image are preserved due to the nonparametric character of the analysis which does not assume a priori any particular structure of the data. To calculate the density gradient let X_i where i=1….n be an arbitrary set of n points in d-dimensional Euclidean spaceR^d. With kernel K(x) and window radius h, the multivariate kernel density is estimated.

$$\hat{f}(x) = \frac{1}{nh^d} + \sum\nolimits_{i=1}^{n} K\left(\frac{x-x_i}{h}\right) \tag{1}$$

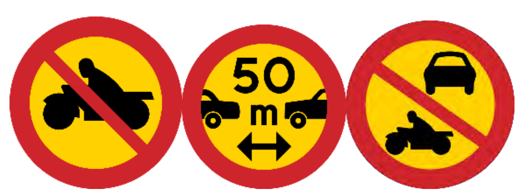
The minimum mean integrated square error (MISE) is the Epanechnikov kernel where C_d is the volume of the unit d-dimensional sphere [26]. To define the estimate of the density gradient the differentiable kernel is used.

$$\hat{\nabla} f(x) = \hat{\nabla}\hat{f}(x) = \frac{1}{nh^d}\sum\nolimits_{i=1}^{n} \nabla K\left(\frac{x-x_i}{h}\right) \tag{2}$$

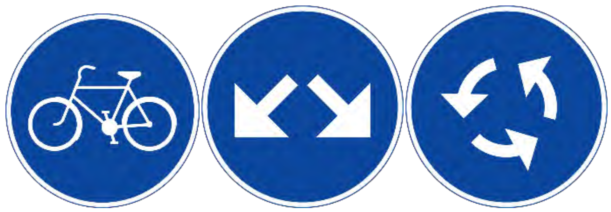
The segmentation technique associates with each pixel in the image the closest local mode in the density distribution of the joint domain. It requires the fusion of the regions associated with nearby modes to segment the image onto clusters[26].

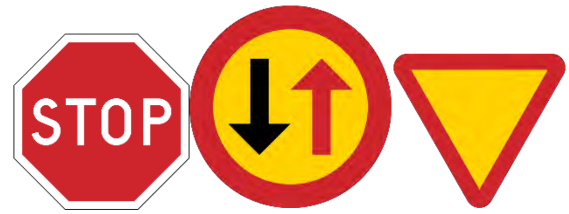
Swedish traffic signs of different categories and different colors are shown below:

(a)



(b)





(c) (d)

Fig 4.2. Examples of Swedish traffic signs. (a) Warning signs with red border; (b) Circular prohibitory sign with red border; (c) Blue color mandatory signs with white border; (d) Prohibitory sign.

Mean shift performs clustering technique to divide initial image into clusters. By each cluster only one image is send as input to random forest classifier. Processing the image pixel by pixel is more time consuming. The resolution of the spatial and the range domains of mean shift are taken as two parameters to use collected data from GTSDB[2]. Than the datasets are trained to random forest classifier.

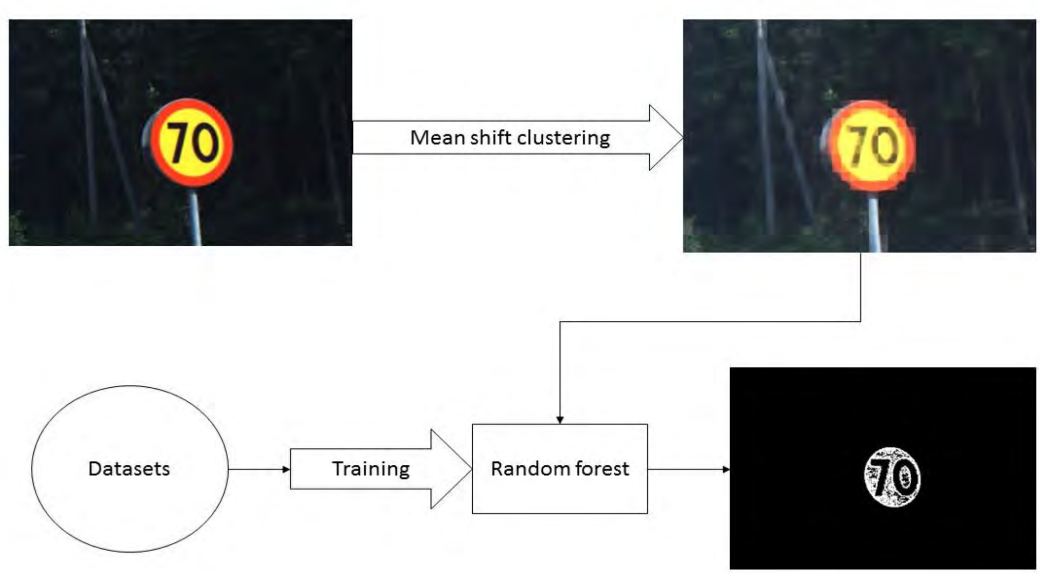


Fig 4.3.The process of color segmentation

In Fig 4.3, the initial RGB image is segmented using mean shift clustering. After that at a time only one pixel is fed to random forest classifier. The segmentation process is completed with the binary image found with only black and white color using aspect ratio and size constraints. The aspect ratio between 0.5 and 1.9 is considered as the best region of interest. The size of the region of the interest is confirmed between $\frac{H\times W}{25}$ and $\frac{H\times W}{3}$. Based on GTSDB[2] and STS[1] data sets the size and the aspect ratio thresholds are selected.

4.3 Shape Detection

FDs are computationally more efficient than Shape Context descriptors, we propose to use FDs for shape characterization. There are three main types of contour-based descriptors: (1) Global descriptors like circularity and eccentricity, which can only discriminate shapes with large dissimilarities; (2) Shape signatures, like centroid distance function, tangent angle or chord length function, which represent the boundary as a one dimensional function derived from the boundary pixels and are sensitive to noise; And (3) spectral descriptors as the Fourier Descriptors, obtained by applying the Fourier Transform on the pixel coordinates of the region contour in a polar-raster way. The first step to compute the FDs is representing every region boundary as a sequence of Cartesian coordinates. Since every region has a particular size, corresponding boundaries may have a different number of pixels. Hence, to characterize every region with the same number of descriptors, every boundary was resampled to Nc points, what might bring a shape smoothing. In the context of traffic sign characterization, another contour-based descriptor named "distance to borders" (DtB) has been used in [18]. This method computes the distance between every pixel of the region contour and the side of the bounding box encompassing it, and it is quite sensitive to rotation. Shape Context has been successfully used as a descriptor in complex object matching [24], and [23] applied it for signs recognition in blue traffic signs. The Discrete Fourier Transform (DFT) of the sequence {cn} is computed as

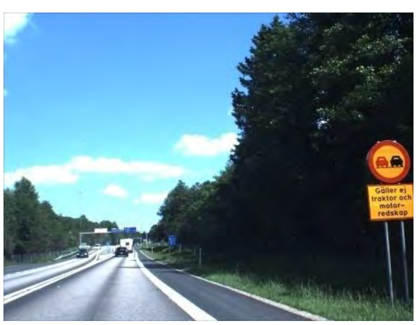
$$C_u = \frac{1}{N_c}\sum_{n=0}^{N_c-1} c_n e^{\frac{-j2xun}{N_c}}, u = {}^{-N_c}/_2, \dots, {}^{N_c}/_2 - 1 \quad (3)$$

An arbitrary pixel (x0,y0), the (x,y) coordinates of the Nc boundary pixels are read in a polar-raster way, transformed to complex numbers cn = xn + jyn with n = 0,...,Nc − 1 and collected in the set {cn}. The complex coefficients Cu are the shape FDs. The rotation information of the boundary is described by the FDs phase, so rotation invariance is achieved by considering only the magnitude of the coefficients. We chose Nc = 128 in order to accelerate the DFT computation. These details are not helpful for shape discrimination, the number of FDs can be reduced by discarding the high-frequency coefficients. Low frequency FDs contain information about the general shape features, while finer boundary details are described by the higher frequencies. We found that a number of 21 coefficients was enough to reconstruct any sign boundary in the spatial domainwhile allowing to discriminate among different shapes after the preliminary experiment.

(a)

(b)

(c)



(d)

(e)

(f)

(g) (h) (i) (j) (k)

Fig 4.4.Normalized magnitude of the FDs of interest for different traffic signs with different size and orientation: (a-f) Original images; (g-k) Segmented ROIs.

The coefficients |C0| and |C−1| were always equal to 0 and 1, respectively, both were removed from the set of features, resulting in 19 coefficients. Translation and scale invariance were achieved by setting C0 = 0 and normalizing the FDs by the magnitude of C−1 using FDs properties. Different road signs with different size and orientation, together with their associated normalized FDs sequence is shown in Figure 7.

To design the classifiers, the normalized FDs sequence was computed for 300 objects of the training database. Each sequence was manually labeled with its shape. From the sequence 40 objects for each shape class and 60 for the non-shape class. Three schemes were proposed for shape classification, namely, voting k-NN, linear SVM, and non-linear SVM. With respect to the SVM schemes (linear and nonlinear), six parallel two class SVM classifiers were designed, each one to identify a shape of interest. In order to reduce the computational burden, free parameters were set to the same values for the six classifiers by conducting a grid-search for $C \in [1,100]$ and $\sigma \in [0.001,10]$, choosing the pair (C,σ) yielding the best accuracy on the validation set. The best accuracy was achieved for $k = 1$ for the voting k-NN scheme. A high rate of semicircular and arrow-shaped regions were incorrectly classified.

Semicircular and arrow-shaped objects were assigned to the non-interest shape class by the linear SVM approach. The nonlinear-SVM scheme only misclassified some non-interest shape objects, which were classified as one of the six shapes of interest. Hence, the non-linear SVM scheme was chosen for the shapes classification stage due to its low missing rate. Best results were provided by the nonlinear SVM scheme ($C = 60$, $\sigma = 0.01$) in comparison to linear SVM ($C = 50$).

For final grouping, two sequential steps were carried out: Firstly, embedded regions were grouped together. Secondly, nearby semicircles were grouped to form a unique circular sign, and nearby arrow-shaped regions were combined to form a unique panel. To determine whether two semicircles/arrow-shaped ROIs were in the same traffic sign, the mass center of every ROI was obtained and distance between mass centers was computed. Situations are common for de-restriction signs and curve indication panels, respectively. ROIs with mass centers close enough were combined in the same traffic sign.

CHAPTER 05

Experimental Results

In this section, the performance of the proposed road sign detection system is evaluated. The proposed method is tested on GTSDB[1] and STS[2] datasets. Using 2.5GHz intel core i5 processor the experiments were performed. To assess the performance the proposed system is compared with recent state-of-the-art method. The merit figures used for the assessment and the statistical results are described.

The procedure is revisited and results are illustrated with two complete examples. A training set with 500 images from GTSDB for system design. A test set from STS data set with 200 images for independent evaluation of the system performance. MatlabR2016a software was used for the experiments. All the images are normalized into 640×480 pixels using bilinear interpolation from both data sets.

The methodology used in detection module is shown in Fig5.1. First of all the initial image is segmented based on color information. Some blobs are discarded based on their aspect ratio and size as an alternative process. This section accelerates the detection and reduces the search space as the number of ROIs to be reduced. This segmentation stage successfully detects the road sign shown in Fig 5.1(b) and 5.1(e). To reduce the ROIs the random forest classifier and Fourier Descriptors are used. Then we refer to the shape detection process. This stage processed the extracted region of interests and classified the shape. The final results are shown in Fig 5.1(c) and 5.1(f). Two representative examples are next presented. Example 1 corresponds to identification of traffic signs in Fig 5.2. In order to get some insight into the segmentation and post-processing stages, Fig 5.2(d)–5.2(f) shows the segmented regions for the three colors of interest here, i.e. red, yellow, and white. Images after the subsequent post-processing stages are presented in Fig 5.2(g)5.2(i), and the final result is presented in Fig5.2(c), where circular and triangular co-located signs are separated, and the signs are correctly identified. This image illustrates that regions are filtered according to their shape and correctly detected as circle and triangle signs. Note that although the small circular sign in the right part of the image is far from the camera, it complies with the size requirements and is still properly detected, whereas the blue circular sign on the left is too small and it is not detected. The final grouping stage allows to detect the two triangular regions as a single traffic sign, since the yellow region is embedded into the red one. Example 2 shows the sign identification process in Fig5.3(a), which conveys a circular sign.

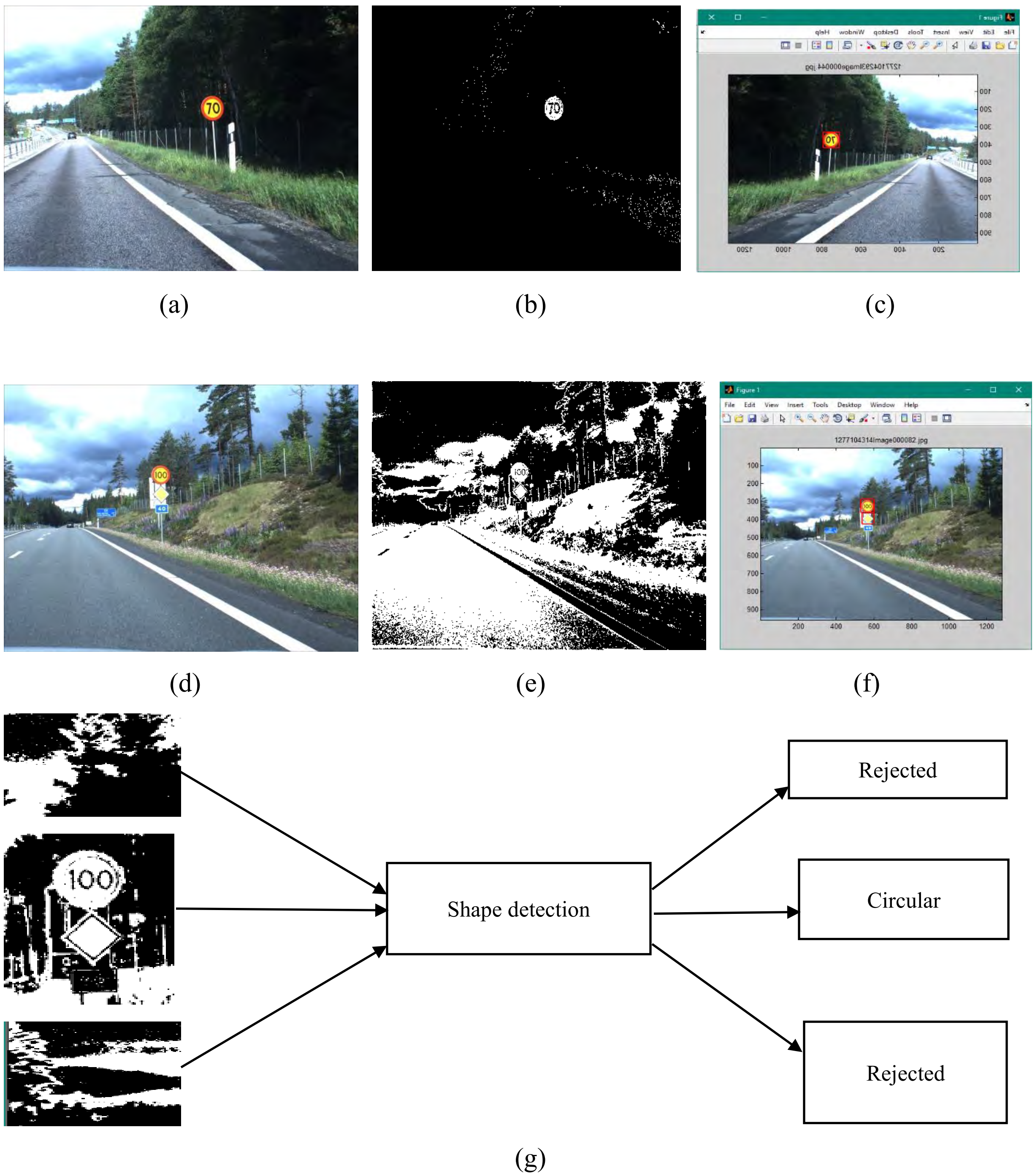


Fig 5.1.The methodology used in detection module. Original image (a) and (d), Segmented image (b) and (e), Final result (c) and (f), Shape classification (g).

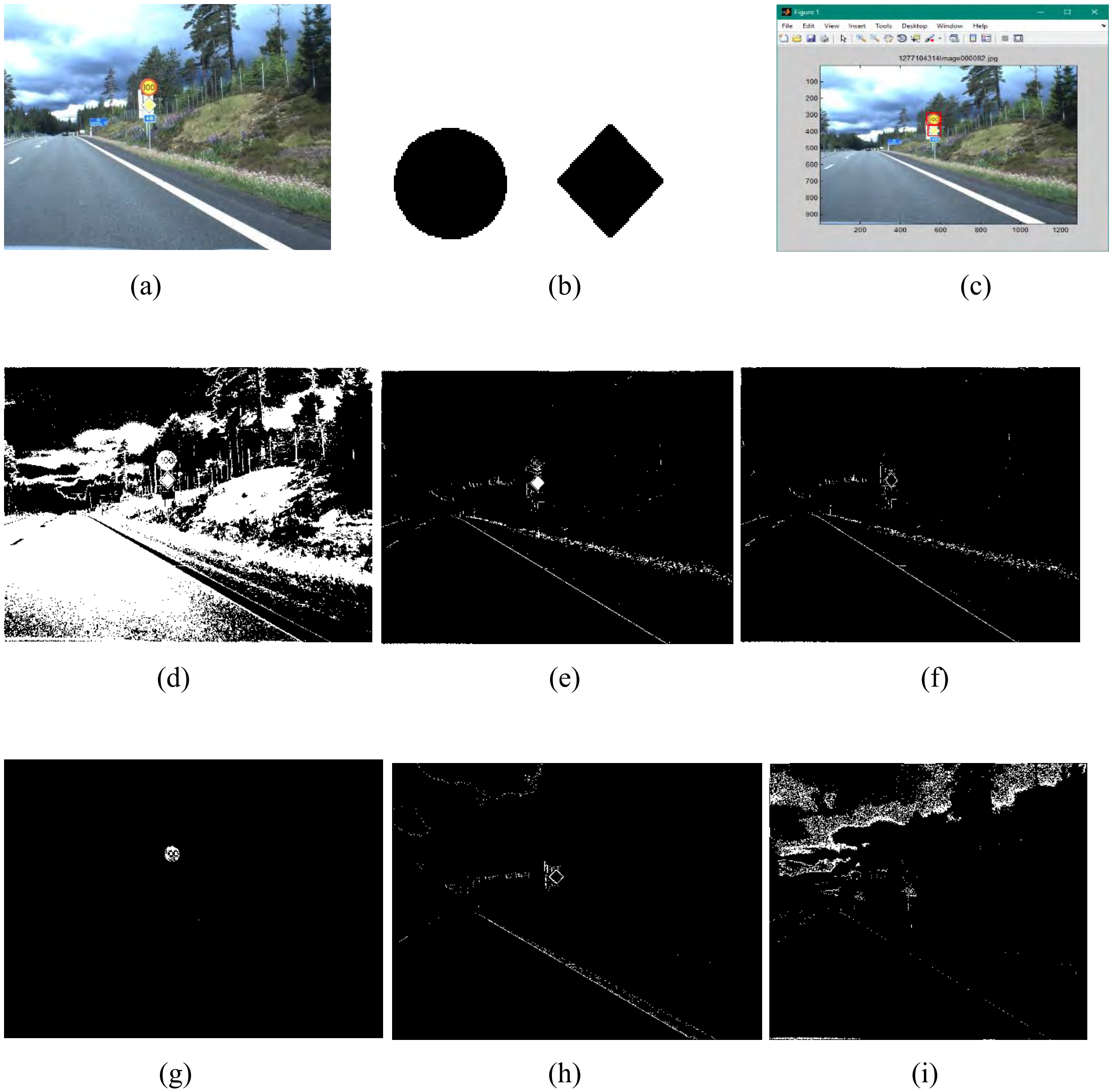


Fig 5.2.Example 1. (a) Original image (b) Shapes classification (c) Result (d-f) Segmentation process (g-i) Post-processing.

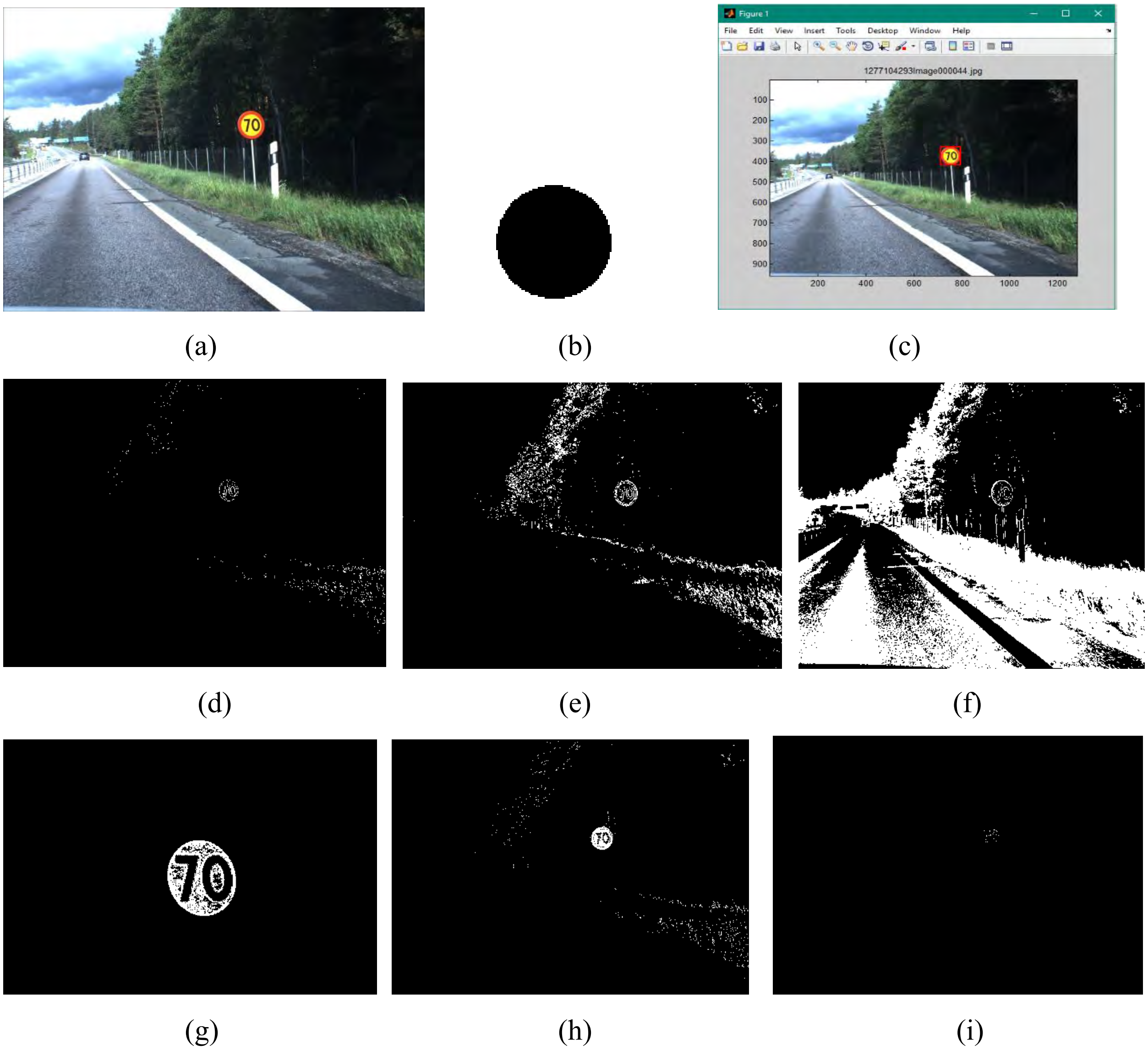


Fig 5.3. Example 2. (a) Original image (b) Shape classification (c) Final result (d-f) Segmentation process (g-i) Post-processing

Images in Fig 5.3(d)-5.3(f) correspond with the output of the segmentation stage for red, yellow, and white colors, and those in Fig 5.3(g)-5.3(i) are the outputs of the subsequent post-processing stage. Note that segmented regions corresponding to sky and vegetation are discarded by the classification stage. Finally, the four arrow-shaped regions are correctly grouped together as misc. sign sown in Fig 5.3(c).

A deeper analysis of all results showed that they corresponded to achromatic signs, which are the most difficult signs to identify, and more, they are occasionally over-segmented. The highest FAR corresponded to rectangular signs, since they can be confused with similar small regions in the scene, such as advertising panels or the back of the trailer trucks. TABLE I presents the performance for each sign shape. All merit figures shown that the best performance was obtained for triangle and circle signs, because these signs have a color and shape that is quite discernible from the rest of the scene.

TABLE I: Performance evaluation of all test images after segmentation and shape classification.

	Signs	FAR	FSR	DR
Triangle	40	0.7	99.7	100
Square	60	1.7	98.3	100
Circle	39	4.2	96.3	100
De-restriction	36	9.8	92.2	98.2
Rectangle	12	3.4	88.6	93.6

The precision and recall values are computed for the proposed technique using following formulas:

$$\text{Recall}= \frac{true\ positives\ detected}{totall\ true\ positive\ detected} \times 100 \tag{4}$$

$$\text{Precision}=\frac{true\ positives\ detected}{all\ detections} \times 100 \tag{5}$$

If the corresponding bounding box overlaps with at least 50% of the area covered by the sign presented in the image is considered true positive. Otherwise the sign is not considered as true positive.

The recall, precision and AUC values of proposed technique is shown in TABLE II and TABLE III. The value are calculated using GTSDB[1] and STS[2] data sets.

TABLE II: The comparison between recall, precision and AUC values between the proposed method and the method in [6] using GTSDB data sets.

	The proposed method	The method in [6]
Recall	94.26%	91.07%
Precision	96.32%	90.13%
AUC	96.22%	93.69%

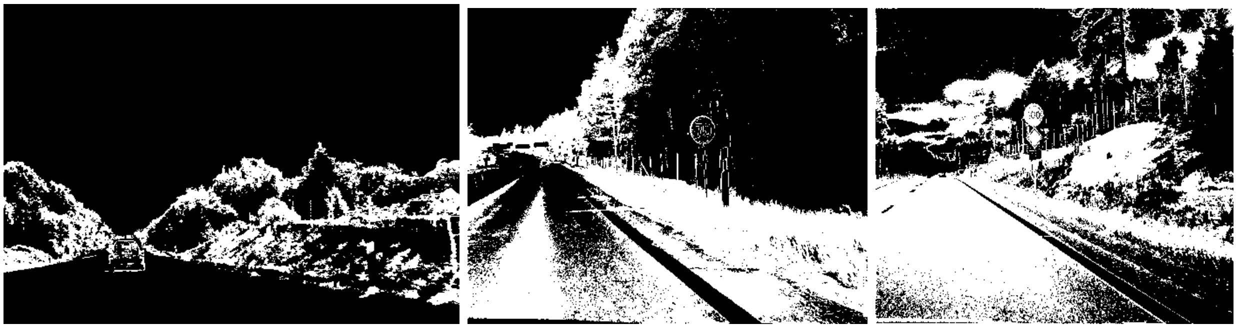
TABLE III: The comparison between recall, precision and AUC values between the proposed method and the method in [6] using STS data sets.

	The proposed method	The method in [6]
Recall	93.63%	93.27%
Precision	92.07%	90.27%
AUC	96.36%	94.05%

Some detection examples are shown in Fig 5.4. The original images are depicted in 5.4(a)-5.4(c). Each images contains different traffic signs from different category. The first image is a prohibitory sign. The second image is a warning sign and the third image is mandatory sign. There corresponding segmented results are shown in Fig 5.4(d)-5.4(f). The shape classification information depicted in the third row from 5.4(g)-5.4(i). The ROIs are reduced in the shape detection section. At most 3 or 4 region of interests are shown in every segmented image. The triangular shape has been detected from 5.4(a). Circulars signs are detected from 5.4(b) and 5.4(c).

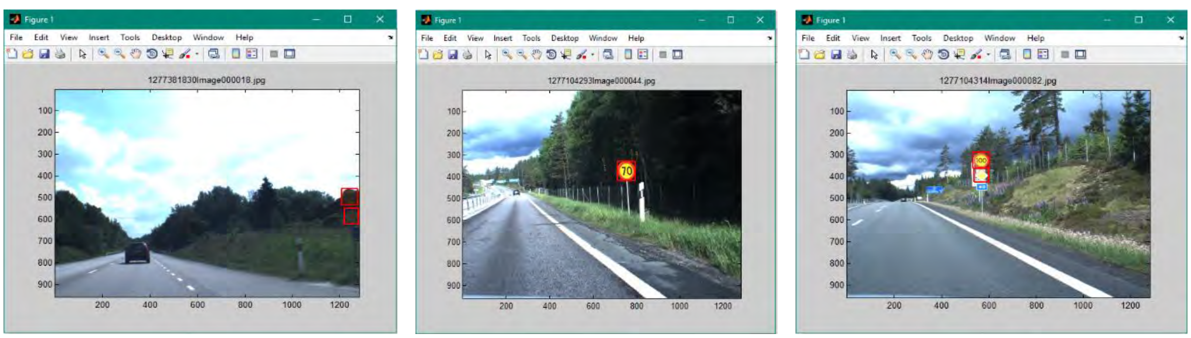
The proposed method is applied on different traffic signs from different weather including low light condition. Some examples of successful outputs of given results are illustrated in Fig 5.5.However, the road signs could not be detected are shown in Fig5.6. The road signs were blur for their motion and low light condition failure results are shown in in Fig 5.6(a) and 5.6(b).

(a) (b) (c)



(d) (e) (f)

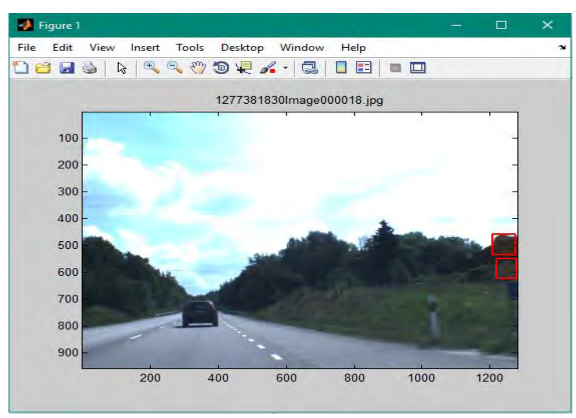
(g) (h) (i)



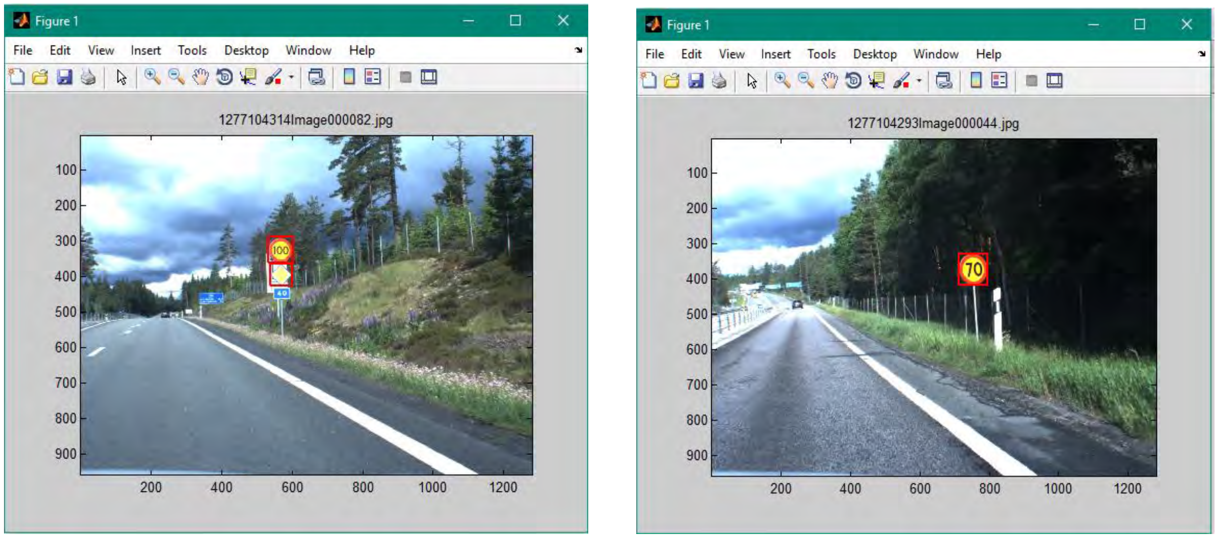
(j) (k) (l)

Fig 5.4. (a-c) Original image (d-i) segmentation and post processing (j-l) Result found

(a)



(b)



(c) (d)

Fig 5.5. Detection examples

(a) (b)

Fig 5.6. Miss-detection examples

CHAPTER 06

Conclusion and Future Work

6.1 Conclusion

In this paper we are using two steps road sign detection method. Firstly we segment the initial image based on the color information. Then we have used mean shift cluster algorithm to transform the image into cluster. Then the clustered image is carried out to the random forest classifier algorithm for segmenting those cluster based on their color. After that we classify the ROIs from the previous steps result according to the shapes. Secondly we have used Fourier Descriptor to detect the shape. SVM based classification algorithm is used to represent the shape. Our proposed model has proven robust against motion blur, scale, rotation and shadows. Our proposed model has provided higher accuracy. The final contribution of this work is the whole procedure which provides with good performance in non-uniform light, motion blur, shadows even when images are taken into different cameras. Our proposed model shows a very high DR which was our primary objective with a low FAR. The highest far was for the black and white sign for which we faces most of the difficulties to determine from other achromatic objects. Experimental results of the GTSDB and STS database show that the proposed method achieves AUC of 96.22% and 96.36% respectively.

6.2 Future Work

In this paper, we worked with road sign detection and recognition. We are planning to use audio information in this traffic sign detection method and to integrate temporal information. We are also planning to improve the robustness of this detection method to similar background color and a tracking algorithm taking into account the spatio-temporal correlation between consecutive video frames would improve the results of this work.

REFERENCE

[1] (2011) Swedish traffic signs database. http://www.cvl.isy.liu.se/research/datasets/traffuc-signs-dataset/

[2] (2013) German traffic sign detection and recognition benchmark datasets. http://benchmark.ini.rub.de/

[3] Bascon SM, Rodriguez JA, Arroyo SL, Caballero AF, Lopez-Ferraras F (2010) An optimization on pictogram identification for the road-sign recognition task using svms. Computer Vision and Image Understanding 114(3):373-383

[4] Kuo WJ, Lin CC (2007) Two-stage road sign detection and recognition. In: Multimedia and Expo, 2007 IEEE International Conference on, IEEE, pp 1427-1430

[5] Larsson F, Felsberg M (2011) Using fourier descriptors and spatial models for traffic sign recognition. In: Image analysis, Springer, pp 238-249

[6] Lillo-Castellano J, Mora-Jimenez I, Figuera-Pozuelo C, Rojo-Alvarez J (2015) Traffic sign segmentation and classification using statistical learning methods. Neurocomputing 153:286-299

[7] Lopez LD, Fuentes O (2007) Color-based road sign detection and tracking. In: Image analysis and recognition, Springer, pp 1138-1147

[8] Mercedes-Benz (2014) Techcenter: Traffic sign assist. http://techcenter.mercedes-benz.com/_en/traffic_sign_assist/detail.html

[9] Opel-International (2014) Opel innovation for your safety. http://www.opel.com/experience_opel/innovation/safety.html

[10] Overett G, Pettersson L, Andersson L, Pettersson N (2009) Boosting a heterogeneous pool of fast hog features for pedestrian and sign detection. In: Intelligent Vehicles Symposium, 2009 IEEE, IEEE, pp 584-590

[11] Ruta A, Li Y, Liu X (2010) Real-time traffic sign recognition from video by class-specific discriminative features. Patter Recognition 43(1):416-430

[12] Traver VJ, Bernardino A (2010) A review of log-polar imaging for visual perception in robotics. Robotics and Autonomous Systems 58(4):378-398

[13] Xia Y, Xu W, Zhang L, Shi X, Mao K (2015) Integrating 3d structure into traffic scene understanding with rgb-d data. Neurocomputing 151:700-709

[14] Opel-International. Opel innovation for your safety [online] (Oct 2014). http://www.opel.com/experience opel/innovation/safety.html.

[15] J. Stallkamp, M. Schlipsing, J. Salmen, C. Igel, Man vs. computer: Bench- marking machine learning algorithms for traffic sign recognition, Neural Networks 32 (2012) 323–332. doi:10.1016/j.neunet.2012.02.016.

[16] R. Gonzalez, R. Woods, Digital Image Processing, 3rd Edition, Pearson Prentice Hall, 2008.

[17] J. Abukhait, I. Abdel-Qader, J. Oh, O. Abudayyeh, Occlusion-invariant tilt angle computation for automated road sign condition assessment, in: IEEE Intl Conf on Electro/Information Technology, 2012, pp. 1–6. doi:10.1109/EIT.2012.6220740.

[18] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jiménez, H. Gómez-Moreno, F. Lopez-Ferreras, Road-sign detection and recognition based on support vector machines, IEEE Trans on Intelligent Transportation Sys- tems 8 (2) (2007) 264–278.

[19] P. Sermanet, Y. LeCun, Traffic sign recognition with multi-scale convolution networks, in: Intl Joint Conf on Neural Networks, 2011, pp. 2809-2813.doi:10.1109/IJCNN.2011.6033589

[20] H. Gomez-Moreno, S. Maldonado-Bascon, P. Gil-Jiménez, S. Lafuente Arroyo, Goal evaluation of segmentation algorithms for traffic sign recognition, IEEE trans on Intelligent Transportation System 11 (4) (2010) 917-930.

[21] G. Overett, L. Petersson, L. Andersson, N. Pettersson, Boosting a heterogeneous pool of fast hog features for pedestrian and sign detection, in: Intelligent Vehicles Symposium, 2009, pp. 584–590.

[22] R. Belaroussi, J. Tarel, Angle vertex and bisector geometric model for triangular road sign detection, in: Workshop on Applications of Computer Vision, 2009, pp. 1–7.

[23] K. Li, W. Lan, Traffic indication symbols recognition with shape context, in: 6th Intl Conf on Computer Science Education, 2011, pp. 852–855.

[24] J. Yu, D. Liu, D. Tao, H. Seah, On combining multiple features for cartoon character retrieval and clip synthesis, IEEE Trans on Systems, Man and Cybernetics, Part B: Cybernetics 42 (5) (2012) 1413–1427.

[25] K. Fukunaga, L.D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition", IEEE Trans. Info. Theory, vol. IT-21, 32-40, 1975.

[26] B.W. Silverman, Density Estimation for Statistics and Data Analysis, New York: Chapman and Hall, 1986.

[27] AliyariGhassabeh, Youness (2015-03-01). "A sufficient condition for the convergence of the mean shift algorithm with Gaussian kernel". Journal of Multivariate Analysis. 135: 1–10.

[28] Shi, T., Horvath, S. (2006). "Unsupervised Learning with Random Forest Predictors". Journal of Computational and Graphical Statistics. 15 (1): 118–138.