

**HO CHI MINH CITY UNIVERSITY OF TECHNOLOGY AND
EDUCATION**

FACULTY OF HIGH-QUALITY TRAINING



HCMUTE



PROJECT OF ARTIFICIAL INTELLIGENCE

**DESIGN AND IMPLEMENT A SYSTEM TO DETECT
VIETNAM TRAFFIC SIGNS IN REAL-TIME**

NGUYỄN HOÀI AN

18146001

MAJOR: ARTIFICIAL INTELLIGENCE

ADVISOR: NGUYEN TRUONG THINH

HO CHI MINH CITY, 6/2022

Preface

Abstract:

This report is the end-of-semester report on the topic of artificial intelligence. It details the completed work for a period of fifteen weeks at Ho Chi Minh University of Technology and Education. Work has been report-oriented, so reporting structured with 5 separate chapters instead of linear product development line. The project is mainly about Vietnam traffic signs discovered but includes a chapter on pedestrians detection is good. A comprehensive survey of traffic The signature detection system has been implemented and it shows a lack of work with Vietnamese signs and a lack of publicity database for these. Therefore, a dataset with nearly 2000 annotated signs has been created. The dataset is unique, not just because it contains Vietnamese signs, but also because it includes video. This report also details investigations into the use of aggregated training data for traffic sign detection devices but concludes that the composite image is not suitable for real-world training images. A pure model based on a shape-only detection system also presented as a building block for a Detection system. Finally, for a person walking two floors detection systems have been developed and documented. This system extends the pre-neural system and produces better detection with fewer false positives.

Acknowledgments:

First of all, I would like to thank Professor Nguyen Truong Thinh, my advisor on the topic of artificial intelligence, for teaching me a lot of knowledge. He is a good person because he has helped me a lot in this subject. You taught me what artificial intelligence is and how to apply it in practice. The subject has helped me a lot in my career path.

Contents

1 Introduction	1
1.1 Survey of the state of the art in traffic sign detection	2
1.2 Traffic Sign Dataset	3
1.3 Synthetic training data	5
1.4 Model based sign detection	8
2 Survey of the state of the art in traffic sign detection.	9
2.1 Introduction	9
2.2 Using traffic sign detection in driver assistance	10
2.3 On traffic signs	7
2.4 Sign detection versus classification	11
2.5 Literature search	12
2.6 Results	14
2.7 Discussion	16
3 Traffic Sign Dataset.	17
3.1 Introduction	18
3.2 Methods	18
3.3 Results	18
3.4 Discussion	19
4 Synthetic training data	19
4.1 Introduction	19
4.2 Methods	19
4.3 Results	19
4.4 Discussion	20

5 Model based sign detection	21
5.1 Introduction	22
5.2 Methods	22
5.3 Results	23
5.4 Discussion	23

1

Introduction

During my time at Ho Chi Minh University of Technology and Education, I focused 100% on absorbing all the knowledge that my teachers taught me. So this project is not developing a product like conventional artificial intelligence projects. Instead, I did research in the area of traffic sign detection and pedestrian detection. Therefore, this report does not document the development of a solution to a particular problem. All work presented in this report was done by myself with input from two supervisors Nguyen Truong Thinh.

1.1 Survey of the state of the art in traffic sign detection

Chapter 2 is a survey of the cutting edge of traffic sign detection. It presents dominant detection strategies and discusses their merits. It also provides an overview of the existing public image databases and presents open issues in the field.

This chapter provides an overview of the signature detection status. Instead of processing the entire Traffic Sign Recognition (TSR) stream, the focus is on detecting signs only. In recent years, TSR has made a great effort, mainly from Europe, Japan and Australia and the developments described.

The detection process is divided into segmentation, feature extraction, and detection. Many segmentation methods exist, mainly based on the evaluation of colors in different color spaces. As for features, there are also plenty of options. The choice is made along with the choice of the detection method. By far the most popular features are edges and gradients, but other options such as HOG and Haar wavelets have been investigated. The detection phase is dominated by the Hough transform and its derivatives, but for the HOG and Wavelet Haar features, SVM, neural networks, and cascading classifiers were also used.

Arguably, the biggest current problem for signature detection is the lack of use of public image databases to train and test the system. Now, every new approach is presented using a new dataset for testing, difficult to compare between papers. This gives the TSR attempt a somewhat scattered look. Recently, several databases have been made available, but they are not yet widely used and include only signs of compliance with the Vienna Convention.

This leads to the main unanswered question in signature detection: Is a model-based shape detector superior to a learned approach or vice versa? Systems using both approaches exist, but are difficult to compare because they both use different data sets.

Many contributors cite the driver assistance system as their main driving force in creating the system, but so far only very little effort has gone into the field of combining the TSR system with other aspects of the assist driving, and notably, no studies have included knowledge of driver behavior to modulate TSR performance system for drivers.

Other open issues include a lack of research in finding non-European signs and signs found that are difficult to relate to their surroundings.

1.2 Traffic Sign Dataset

One result of the survey was the lack of a public database of traffic signs in Vietnam. Because of this, it was decided to create one. This database creation and structure is the topic of chapter 3.

This chapter describes the assembly of a Traffic Sign Dataset, collected during driving through urban environments in a Chinese city. It is a dataset of 2000 signs annotated in 28 classes collected and modified on a huge dataset of traffic signs in China. A toolkit for creating annotations has been created, with a strong focus on traceability, so each caption can be retrieved back to its source video.

The dataset fills the gap as it is the first public dataset with Vietnam traffic signals. It also includes full video footage, which can enable the development and testing of detection systems using tracking.

1.3 Synthetic training data

In chapter 4, synthetic training data is mentioned. Creating a sign database is tedious and time consuming, and since all traffic signs are based on well-defined patterns, it is a good idea to research whether the generated training data is aggregated. can be used to detect signage instead of real-world data. Chapter 4 description results of this study.

This chapter describes experiments to evaluate the performance of synthetic training data for traffic sign detection against real-world data. A program has been developed to generate synthetic training data that will simulate real-world signage images based on a drawn pattern. Several detectors were then trained with both synthetic data and real-world data. A waterfall application with a lot of features already used as a detection framework.

The results show that the aggregated data generated here is not of sufficient quality to compare with real-world training data, so unfortunately the aggregated generated data cannot to replace carefully collected real world data.

1.4 Model based sign detection

However, not all signature detection approaches involve training data. Chapter 5 describes another way in which signs are detected using a theoretical model of their shape.

This chapter describes a purely model-based detector, based on the shape of the traffic signs. This method has been tested on a subset of the data set and shows a pretty good per-signal detection result, with many errors. the positive side, however. The tracking system used to reduce false positives is too simple and can even affect overall detection performance. The detector itself is not very useful but will work well as a component of a larger system. This is consistent with how it is used in existing documents.

Survey of the state of the art in traffic sign detection

2.1 Introduction

The field of Traffic Sign Recognition (TSR) systems has received increasing research interest in the past decade, but the task of recognizing signs in Vietnam has so far remained unexplored. TSR is a task with clearly defined different applications.

1. Highway maintenance: Check the presence and condition of signs along major roads.
2. Sign inventory: Similar to the above task, create an inventory of symbols in the city environment.
3. Driver assistance system: Assists the driver by notifying current restrictions, limitations, and warnings.
4. Intelligent Autonomous Vehicle: Any self-driving vehicle used for driving on public roads

In recent years, speed limit detection systems have been introduced into the forefront of many vehicle models by many manufacturers, but a more general sign detection solution and integration into other vehicle systems remain. has not materialized. The most modern condition

The TSR system does not use driver information, or driver input, to enhance performance. Extensive studies of Human-Machine Interaction are needed to present TSR information carefully and to inform drivers without distraction or confusion.

2.2 Using traffic sign detection in driver assistance

Nearly all the articles surveyed point to driver assistance as the main driving force behind the creation of the system but despite this, very few studies actually, involve drivers as well because smart cars have not really been given due attention. To work with drivers, the TSR study needs to take into account the driver's imaging system. This may include factors such as the visual saliency of the markings, the driver's attention span, and the perceived load. For example, a driver may have to pay attention to a sign, but fail to recognize it and not remember its information. Although drivers always pay attention to speed limit signs and recall their information, they are less likely to confuse or distract from driving.

The implications of using TSR in smart vehicle systems are clear. Instead of focusing on detecting and recognizing all the signs of a certain class of vehicles, which is already the goal of an automatic transmission, the task is now to detect and highlight those that are not. The driver was not seen. This gives way to different TSR models, up to now driver attention and interaction problems. It simply recognizes all the signs present. In costume. Instead, it simply highlights the type of sign that is easily overlooked, like the pedestrian warning in the study.

2.3 On traffic signs

Traffic signs are markers placed along roads to inform drivers of road conditions and limits or directions. They convey a lot of information but are designed to do it efficiently and quickly. This also means that they are often designed to stand out from their surroundings, making the task of detection fairly well defined. Instructional and informational markings are not particularly interesting from the point of view of the driver assistance system.

The design of traffic signs is standardized through legislation in Vietnam. There, shapes are used to classify different types of signs: Circular signs are prohibited signs including speed limits, triangular warning signs, and rectangular warning signs or auxiliary signs in combination with one of the standard shapes. In addition to these signs, octagonal signs are used to signal a full stop, downward triangles, and countries of various types, e.g. to announce city limits.

2.4 Sign detection versus classification

The task of TSR is usually divided into two different phases: Detection and classification. Detection involves locating marks in the input image, while classification is about determining what type of mark the system is looking at. The two tasks can often be handled completely separately, but in some cases, the classifier relies on the detector system to provide information, such as sign shape or sign size. This also means that if necessary, different classifiers can be run, depending on the shape of the detected sign. In this report, the following term is used: Detection of signs indicating

where the position of a sign in the input image is determined. Sign classification is the task of finding out exactly what type of sign it is. Sign recognition is the entire process from data collection to final output, including detection and classification.

2.5 Literature search

The literature search for this survey was carried out in four stages:

1. Look for reports or previous experiments.
2. Choose a set of several articles and read them to get an initial overview of the field.
3. View references from the original selection of articles, as well as other articles written by the same author, to get a fuller understanding.
4. Search the public database of traffic sign images. The first step is to find out if there have been any surveys. This step is actually taken before it is decided to do a detailed survey. If any comprehensive survey documents are already in place, further steps will not be necessary. As mentioned earlier, the literature contains only two surveys of TSR, none of which are very comprehensive. Step 2 is done by looking at references in existing surveys as well as performing a search in several scientific databases. For step 3, a larger set of articles was selected. This is done by keeping track of references in the previous group of articles, searching for other publications by known authors, and simply performing more general searches. After reading the papers, it is clear that hardly any papers are used the same dataset for testing, so it is necessary to determine if any public tuples exist. That's step four. Some datasets do exist, but they are very new and not well advertised, so they are mostly found using references in some papers.

2.6 Results

Approaches in the detection phase have traditionally been divided into two categories:

- Methods based on color.
- Shape-based method.

The color-based approach takes advantage of the fact that traffic signs are designed to be easily distinguishable from their surroundings, often colored in conspicuously contrasting colors. These colors are extracted from the input image and used as the basis for detection. Like signs with specific colors, they also have very well-defined shapes that are searchable. Shape-based methods omit color in favor of the characteristic shape of the sign. Each method has its pros and cons. The color of the sign, although theoretically well-defined, varies greatly depending on the available light, as well as the age and condition of the sign.

On the other hand, finding specific colors in an image is pretty straightforward. The shape of the mark always changes with light and age, but parts of the mark may be obscured, making detection difficult, or the mark may rest on a similarly colored background, impairing detection. edge detection that most shape detection devices rely on. Dividing systems in this way can be problematic. Almost all colors based on these approaches will form after having considered a color. Others use shape detection as their primary method, but also integrate some color aspects. Here, the detection step is divided into the feature extraction step and the actual detection step, which works on the extracted features. Many methods are shape-based only with no segmentation steps.

An overview of all the surveys and their methods is listed in the survey. To conserve space, these tables are not reproduced here. It contains each system and lists the segmentation method, feature type, and detection method used. They do not constitute a ranking in any way. Apart from this division, the two tables are structured in the same way: The type of sign-on paper describes what kind of sign the paper's author was trying to figure out, while the type of sign can be the types of signs. signals that this method can be extended to include, often a very broad group. Real-time is about the running speed of the system if that information is available. Any system with a frame rate faster than 5 fps is considered to have real-time potential. Rotation invariant indicates whether the technique used is suitable for the rotation of signs. Model vs. descriptive training if the detection system is based on a theoretical model of markers (such as a predefined shape), if it uses a learned class of classifiers, or if it uses use a combination of the two. The test image type is the image resolution the system is designed to

operate at. Low-resolution images are usually video frames, while high-resolution images are still images.

Very few papers use a common database to test their performance, and the papers detect many different types and numbers of markers. So the numbers shouldn't be compared directly, but they give an idea of the performance. Not all articles provide all reported measurements. In other cases, these exact measurements are not given but can be calculated from other given numbers. When metrics are available, the best detection rate obtained by the system is reported along with the corresponding false-positive measurement. The detection rate is per frame, meaning 100% detection is achieved only if a signature is found in every frame it is present. Only detecting the signature in a few frames is not enough. This is how the results are presented in most papers, so it is the metric chosen here, even if a real-world system would perform well if only one signal was detected time.

The purpose of the segmentation step is to get a rough idea of where the markers might be, and thus narrow the search space for the next steps. Not all authors use this step. Since segmentation has traditionally been performed based on color, the authors believe that this is not part of signature detection that does not typically have any direct segmentation steps to detection.

Of the papers that used segmentation, all but Gu et al. (2011); Keller et al. (2008) use color to some extent. Usually, segmentation is done with color and then shape detection is done at a later stage. Gu and associates. (2011) reverse the usual order, so they use radial symmetry voting (see section 2.6.6) to segment and

a color-based approach to detection. Keller et al. (2008) also run radial symmetric voting as preprocessing, but track it with a cascade classifier using Haar wavelets In general, color-based segmentation is based on the threshold of the input image in a zero color space. Since many people believe that the RGB color space is very fragile with regard to light variations, these methods are led by the HSI space (or its close cousin, the HSV space).

Some authors are not satisfied with the performance of HSI, as it does not model color temperature changes in different weathers, but only changes in light intensity. This allows them to have a wide variety of color temperature variations. The RGB space was used by Timofte et al. (2009); Prisacariu et al. (2010), but they use adaptive thresholding in an attempt to combat instability caused by light variations.

Of particular interest to this color space, the discussion is an excellent paper by GomezMoreno et al. (2010), which showed that HSI-based segmentation offers no significant benefit over normalized RGB, but methods using color segmentation in general performs much better than shape-only methods. However, they have difficulty with the white signs. For a long time, it was generally assumed that the RGB color space was a bad choice for segmentation, through rigorous testing they showed that nothing was gained by switching to the HSI color space instead of the RGB space. standardized. As the authors wrote: “Why use a complex and nonlinear transformation if a simple normalization is good enough?”. Deguchi et al. have proposed a color-based model that does not rely on thresholds. (2011), using a cascade classifier trained with AdaBoost, similar to the one proposed by Viola and Jones (2001), but on Local Ratings, Model features instead of Haar wavelets. In addition, Ruta et al. (2010) use color-based search, which, although closely related, is not directly threshold-based. Here, the image is segmented into the colors that may exist on the signs. The customizing process is less destructive than thresholding in that it doesn't remove the pixels directly, instead, it maps them to the closest pixel to the color associated with the signature. In a more recent contribution (Ruta et al., 2011), they replace the color customization method with the Quad-tree interest area search algorithm, which searches for interesting areas using the method of interest. iterative search method for colored markers. One the realm is located in Houben (2011), who uses the learned probabilistic color preprocessing process.

2.6.1 Features and modeling

While different features are available from the visual literature, feature set selection is often coupled with a detection method, although some feature sets may be used with several detection methods. show different. The most popular feature is the edges- sometimes edges are obtained directly from the raw images, and sometimes edges are from the pre-segmented images. In fact, edges are always found using Canny edge detection or some similar method and they are used as unique features by Ruta et al. (2010, 2011); Loy and Barnes (2004); Barnes and Loy (2006); Barnes et al. (2008); Nunn et al. (2008); Meuter et al. (2011); Garcia-Garrido et al. (2011); Gonzalez et al. (2011); Timofte et al. (2009); Liu et al. (2002 year); Kuo and Lin (2007); Moutarde et al. (2007); Belaroussi and Tarel (2009); Ren and associates. (2009); Chiang et al. (2010); Qingsong et al. (2010); Deguchi et al. (2011); Houben (2011). Prisacariu et al. (2010) combine edges with Haar-like features and Ruta et al. (2007); Hoferlin and Zimmermann (2009) only looked at certain color-filtered edges.

While edges are the most popular feature choice, there are other options. The Directional Gradients (HOG) chart is one. It was first used to detect people in images, but has been used by Alefs et al. (2007); Pettersson et al. (2008); Overett and Petersson (2011); Gao et al. (In 2006); Xie et al. (2009) to detect markers. HOG's headquarters on how to graph gradient directions on arrays of images and compare them with known histograms for searched objects. HOG was also used by Creusen et al. (2010), but they enhanced the HOG feature vectors with color information to make them even more robust. Some articles by Bahlmann et al. (2005); Keller et al. (2008); Prisacariu et al. (2010); Baro et al. (2009) use wavelet-like features Haar, Bahlmann, et al. (2005) only on certain colors, and Baro et al. (2009) in the form of the so-called dissociative dipole with wider structural options than traditional Haar wavelets.

2.6.2 Detection

The detection stage is where the markers are actually found. In many ways, this is the most important and often the most complicated step. The choice of detection method is a bit more limited than in the previous two stages because this method has to work with features from the previous stage. As a result, the decision is often made in reverse: The desired detection method is chosen and the feature extraction stage is designed to provide what is needed to perform the detection. As is known from the previous section, the most common feature is the edges and this reflects the most popular choice in the detection method. Using the Hough transform to process edges is one of the options, as done by Moutarde et al. (2007); Garcia-Garrido et al. (2011); Gonzalez et al. (2011); Ren and associates. (2009). Moutarde et al. (2007) use a proprietary and undisclosed algorithm to detect rectangles in addition to the Hough transform used for circles. That said, Hough transformations are computationally expensive and not suitable for systems with real-time requirements. Therefore, the most common methods are derivatives of the radial symmetry detector first proposed by Loy and Zelinsky (2003) and first put to use by Barnes and Zelinsky (2004) to detect signals. The voting algorithm for most likely sign centers in an image is based on symmetry edges and is itself inspired by the Hough transform. The basic principle can be seen in the figure. 2.9. In a circle, all edge gradients intersect at the center. The algorithm finds gradients of magnitude above a certain threshold.

2.7 Discussion

In the previous sections, different methods and philosophies for each stage were presented. This section discusses the current state of the art and outlines ideas for future research directions. Currently, the problem in the TSR is the lack of use of a standardized signage image database. This makes comparisons between contributions very difficult. In order to make meaningful progress in this area, the development of such databases is crucial. So far, studies have only taken one method that they believe has potential, perhaps tested several solutions. There's no way to compare performance with other systems, it's not clear which approach works best, so every new team starts from scratch, implementing what they think might work best. Two attempts to remedy this situation are worth mentioning: The signature database presented earlier and the segmentation assessment performed by Gomez-Moreno et al. (2010). As mentioned earlier (section 2.6.1), several databases of public signs have emerged recently, but are not yet widely used. Gomez-Moreno et al. (2010) compared different segmentation methods on the same data set containing a total of 552 markers in 313 images. They also propose a way to evaluate the performance of segmentation methods. That paper provides a very good starting point for determining which segmentation method to use.

However, these two efforts are needed, and a public database of signs from areas not covered by the Vienna Convention is needed. The database that includes videos of the markers will also be very beneficial to the development of the TSR system, as many detectors use a system of tracking signs. This is, to some extent, included in the KUL Dataset. Not using a public database may not fully explain why so few comparative studies of the methods exist. Another reason is that TSR systems are long, complex chains of various methods, where it is not always possible to interchange each module. For example, when a detection method cannot be swapped for something else, it becomes difficult to determine whether other solutions might be better. This is addressed, if more papers divide their work into more clear stages, ideally used in this survey, plus a similar set of stages for classification. This was successfully done by Gomez-Moreno et al. (2010), as they experimented with different segmentation methods while keeping the feature extraction, detection, and classification stages constant.

3

Traffic Sign Dataset

3.1 Introduction

When training and testing a system like this, it is generally preferable to use data from existing public databases as opposed to collecting data for the particular project. Not only is it saving time and effort, it is also convenient for comparing results to previous works in the area.

3.2 Methods

This section describes the methods used to obtain the Traffic Sign Dataset. It describes both considerations about the actual data content, and the technical means with which it was collected. During this section there will be no distinction between training and test data. In the end, splitting the dataset into separate training and test pools is discussed.

3.2.1 Dataset content and structure

The design and content of the real-world dataset were inspired by other available datasets presented earlier. The more similar the different tuples are packaged, the easier they are to use, so some effort has been made into this. All significant datasets now contain several images annotated with the type and location of the markers. That's the bare minimum needed for any traffic sign dataset. However, the STS and KUL datasets have one significant advantage: They include full frames, making them useful for both detection and classification systems. This is very important, especially when this project is being discovered. Furthermore, the STS dataset includes additional annotation data: Information about whether a sign is visible, blurred, or obscured and whether the sign is on an existing or secondary road. As described in chapter 2, the task of determining whether a recognized signature really belongs to the pathway is currently not well explored, and the inclusion of this information in the dataset lays the foundation for this effort. This force. All datasets save their annotations in comma-separated text files (CSV files) in slightly different formats. The dataset from this report - named the labeled dataset- tries to get the best out of it each dataset adds more and more data at the same time. No

existing database includes video to any great extent. They all have a large collection of single annotated images, but this means they cannot be used to test detectors against timing information. Many systems already use different monitoring plans to minimize the number of false positives, and it is likely that in the future, findings using transient data will emerge more often. Thus, the dataset includes video as well as independent frames.

4

Synthetic training data

4.1 Introduction

Many signature detection systems (see chapter 2) rely on large amounts of training data to function. Over the past two years, several datasets on traffic signs have emerged, but none yet include anything other than the Vienna Convention signs or - with the contribution of this project - Vietnam signboard. Because of the different signs between arrival areas regions, and in many cases from country to country, an interesting proposition is to use synthetically generated training data instead of real images, saving a lot of time and effort in acquiring collect data. Synthetic training data is still not widely used in the field of TSR, but because traffic signs have a clear legal form, it is possible to randomly skew the samples to simulate real-world variations.

4.2 Methods

The work in this chapter can be divided into three parts: Synthetic data generation, machine learning-based detector creation, and detection performance evaluation. Each is described in more detail below.

4.3 Results

The bars are trained on twenty-seven different types of signs: Stop signs, pedestrian crossings, speed limits 5, and signals ahead,.... They are chosen because of their wide range of shapes and colors, and because they are well represented in the data set. Initial, the detectors were trained with a subset of the data set and with a corresponding number of synthetic training images. Finally, for classes where more real-world training images can be collected, real-world detectors are also trained.

4.4 Discussion

Aggregate data frequently underperforms real data. In all cases, real-world detectors have a detection rate above 70%, while synthetic detectors never exceed 30%, except in the case of a permanent synthetic stop-sign detector. Only use 10 stages. That number reaches 58.3% but at a cost seriously high false-positive rate. Provide more training data in aggregation

5

Model based sign detection

5.1 Introduction

To date, this report has mainly been concerned with machine learning-based detection approaches using training sets. But as mentioned in Chapter 2, model-based approaches are also common. This chapter presents the most common model-based approach. FRS is the most common shape-based detector in traffic sign detection and is used alone or in combination with color thresholding. For this project, it was taken as a standalone detector and essentially a shape detector. In the current version, there is no mechanism to eliminate the detection of shapes that match the pattern of a sign shape, but not the sign. It was chosen for a number of reasons:

- Previous experiments have shown that color-based detection methods are rather unreliable with varying lighting conditions, so one wanted to test a shape-based detector.
- There is no guarantee that the input data will be color video, this is also calculated using a purely shape-based approach.
- U.S. speed limit signs are monochrome with a white background making them difficult to detect with color-based approaches (Gomez-Moreno et al. (2010) have shown that it can be used with achromatic decomposition, although with mixed results).
- FRS is the most popular signature detection algorithm.

5.2 Methods

5.2.1 Radial Symmetry Voting

In a real application, the gradients have to be found first and that can be done using the Canny edge detector (Canny, 1986) or by complicating the image with an x-directed Sobel kernel and y and combining them. It's basically Canny's algorithm, but there's no zero maximum cancellation. Once the gradients are found, all gradients with magnitudes below some threshold are discarded. This

reduces the computational load of the algorithm and is allowed because the expected shapes are significantly different from their background. Then just assume each non-zero gradient is part of the circle and a vote is taken accordingly, as described above. The votes are tracked on a separate vote image of the same size as the input image. Because votes are made for particular circle sizes, the algorithm must be run multiple times if it needs to search for circles of different sizes. Then all points with enough votes or a fixed number of points with the most votes are selected for the center of the circle.

5.3 Results

No attempt was made to optimize the detection parameters, they were only chosen from what seemed reasonable, some errors were due to the sign being too large. Slightly better results are likely to be obtained if the settings are tweaked carefully. In addition, a limit has been set, so the detector only outputs the best 5 shapes in each frame (if more than 5 shapes pass requires two-frame tracking). This makes it possible, although unlikely, that the marks found were a less certain shape than the others in the frame.

5.4 Discussion

The detector has a large number of false positives and cannot work on its own, but it is not intended for that. It is important to remember that this is a general shape detector so some false detectors may actually be true positives in the sense that they are actually such a shape, just not an indication. However, a qualitative assessment of performance shows that this is rarely the case. Detectors are often fooled by the branches and trunks of trees that form patterns in the sky. The detection rate per frame is not impressive, but in most cases, the mark is at least found at some point in the track. Judging from the detection rate per frame, the stop sign is the hardest to find, which isn't much of a surprise, since it's the most complex of shapes. However, many stop signs are of very poor quality, which could explain the underperformance. The best performance is achieved for forwarding signaling signals, possibly because the combination of diagonals is not found. Qualitative evaluation of the detection results shows that the detector performs much better on sharp, shake-free, meaningful images when the detector is in operation by looking at the sharp transitions formed by strong contrast. As with any detection method, the better the source document, the better it performs.

