UNIVERSIDAD CARLOS III DE MADRID

TESIS DOCTORAL

Advanced Driver Assistance System based on Computer Vision using Detection, Recognition and Tracking of Road Signs

Autor:
Juan Pablo Carrasco Pascual

Directores:
Arturo de la Escalera Hueso
José María Armingol Moreno

DEPARTAMENTO DE
INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

Leganés, Julio 2009

TESIS DOCTORAL

ADVANCED DRIVER ASSISTANCE SYSTEM BASED ON COMPUTER VISION USING DETECTION, RECOGNITION AND TRACKING OF ROAD SIGNS

Autor: Juan Pablo Carrasco Pascual

Directores: Arturo de la Escalera Hueso
José María Armingol Moreno

Firma del tribunal calificador:

Firma

Presidente:

Vocal:

Vocal:

Vocal:

Secretario:

Calificación:

Leganés, 6 de Julio de 2009

This thesis is submitted to the Departamento de Ingeniería de Sistemas y Automática of the Escuela Politécnica Superior of the Universidad Carlos III de Madrid, for the degree of Doctor of Philosophy. This thesis is entirely my own work, and, except where otherwise indicated. describes my own research.

A mi madre

Acknowledgements

En un trabajo de esta envergadura no sólo han tenido participación los componentes del grupo de investigación sino también todos aquéllos que han hecho este recorrido más fácil, a todos ellos está dedicada esta tesis.

A Arturo y José María. El entorno humano que rodea la investigación es una importante motivación extra, y ellos siempre han procurado que al menos dentro del grupo me sintiese lo más a gusto posible. Fuera del entorno de investigación han respetado mi independencia y mis decisiones y no he conocido en estos cuatro años una postura tan abierta al diálogo y tan respetuosa con los derechos y la justicia como la de ellos. Esto es digno de elogio ya que supone que, por encima de nuestro trabajo, hay una dimensión humana inviolable que es la que realmente nos hace importantes, no respetarla significa no respetarnos a nosotros mismos. Espero que mi trabajo y, sobre todo, mi actitud en este tiempo les sirva de homenaje.

A Michael, que engañó a mis tutores para que aceptasen a un físico como doctorando, parece que finalmente no se equivocó...

A mis compañeros del grupo, Paco Pepe, Basam, María, Dani, Marco y Nando... en esta última parte de la tesis sobre todo, han estado conmigo al 110%. A Carlos y a María, echaré de menos los saludos mañaneros de uno y los abrazos de la otra.

A mis excompañeros y amigos de Electrónica que se lo merecen todo, en especial a Ricardo, Víctor, Rui, Vincent, Belenchu, Charo y Cristina.

Ad Alberto Broggi e Massimo Bertozzi per avermi accolto nel Vislab. A tutti i componenti del gruppo, in particolare a Pierpa e Paolo M. con i quali ho lavorato più a stretto contatto e ho sempre potuto contare su di loro. A Barbara. A Isabella, che sempre si è mostrata disponibile e mi ha sostenuto sia in ambito lavorativo che personale, non ho parole per esprimere la mia gratitudine per la sua amicizia, senza di lei la mia permanenza sarebbe stata sicuramente diversa. A Manuela, l'allegria fatta persona, a lei devo tutto

l'italiano che conosco.

Für Dietrich Paulus, Direktor der AGAS an der Universität Koblenz, wegen seiner Großzügigkeit, seiner uneingeschränkten Verfügbarkeit zur Verbesserung meiner Arbeit und der Kletterausflüge. Für alle Mitglieder der Universitätsgruppe, durch deren warmherzige Aufnahme mein Aufenthalt zu den besten Monaten meiner Promotion wurde, indem sie mich bei meiner Arbeit unterstützt und mich auch außerhalb der Arbeit zu Unternehmungen eingeladen haben. Für Marina, Benjamin, Johannes H., Detlev, Johannes P., Vinh, Wolfram, Peter, Marcin, Frank, Kerstin, Olaf, Antje, Christian und vor allem Stephan, ein Freund, den man sich nicht besser vorstellen kann.

A mis amigos de siempre, los que aguantan todos mis defectos y me regalan todas sus virtudes: Jacobo, José, Ana, Iván, Álex, Fer, Antonio, Inma, Juanma, Jesusfer y María, Jesús, Bea, Thomas, Marta, Marijose y Silvia.

A los Arcilla-Cobián, mi referente en muchos sentidos y mi segunda familia, nunca podré pagarles tanto cariño y atención como me han dado.

A mi familia al completo, especialmente a Mayte y Begoña y sobre todo a mis tíos Belén y Juan, los mejores: me han dado apoyo, cariño, comprensión, me han aguantado en su casa cuando las jornadas laborales se alargaban hasta límites insanos, y siempre han tenido la palabra justa en el momento exacto para hacer que me sintiera mejor sin esconder la realidad. Cuando sea mayor, más mayor quiero decir... quiero ser como ellos.

Por último a mi madre, la más importante. Todo cambiará y seguro que ella seguirá ahí. De ella aprendí a no rendirme y a luchar por lo que quiero con uñas y dientes. Le debo todo lo que soy y este trabajo le está especialmente dedicado.

Resumen

Los accidentes de tráfico son un grave problema socioeconómico. Obviamente el coste humano es imposible de evaluar y el económico supone un continuo e ingente gasto de dinero por parte de los gobiernos. Se han propuesto diferentes soluciones para paliar los efectos de los accidentes, una de las cuales, los Sistemas Avanzados de Ayuda a la Conducción, son el marco en el que se encuadra el presente trabajo. Estos sistemas, como su nombre indica, asisten al conductor ofreciéndole información del entorno o actuando en determinadas circunstancias para la salvaguarda de los ocupantes del vehículo, o para facilitar la conducción.

El sistema que se propone en esta tesis es una plataforma multipropósito original en su concepción, cuyo fin más inmediato es reconocer las señales de tráfico de prohibición, peligro, ceda el paso, obligación e indicación. La información obtenida de ese reconocimiento se integra dentro de un módulo de aviso al conductor. Lo que se pretende es que el conductor conozca en todo momento si está contraviniendo alguna norma de circulación derivada de una velocidad o maniobra inadecuada para el tipo de señal que se ha reconocido.

Dado que este sistema está embarcado en un vehículo, deberá cumplir dos requisitos de especial importancia: funcionar en tiempo real y tanto en entorno urbano como en autopista. El primero cobra especial relevancia si se piensa en que la seguridad de los ocupantes del vehículo y de los peatones puede depender de los avisos que permitan al conductor anticiparse a un peligro. La segunda, que es una aportación original, garantiza que el sistema funcionará en vías donde la velocidad es mayor y por tanto también la probabilidad y gravedad de un posible accidente.

El sistema, como se decía anteriormente, sirve al desarrollo de otras aplicaciones, como es el caso del inventariado automático de señales de tráfico, tan en auge actualmente.

El sistema de asistencia a la conducción propuesto condensa todas aquellas etapas que históricamente se han ido sumando a este tipo de aplicaciones: detección de la señal, clasificación (extracción de la información del símbolo) y seguimiento de la señal.

En el documento se tratan en profundidad e introductoriamente los distintos espacios de color y componentes de los mismos que mejor se adecúan a la búsqueda de señales en la imagen. La técnica de emparejamiento de patrones con la que se detectan las señales se optimizará mediante este estudio, así como el de los modelos que se utilizan para cada tipo de señal.

En la parte de reconocimiento se hará uso de las dos herramientas más importantes utilizadas hasta el momento en reconocimiento de objetos: emparejamiento de modelos y redes de neuronas. Se hará una comparación de las dos técnicas, con especial atención al preprocesado de la imagen, de manera que los resultados sean óptimos.

Finalmente se aborda el problema del seguimiento de la señal. La propuesta es un modelo basado en el movimiento de la cámara que, solidaria al vehículo, cambia su posición respecto a una señal de tráfico que se toma como referencia.

El sistema está implementado en una plataforma experimental bajo condiciones reales. Esto y la variedad de experimentos que se muestran para cada una de las partes en que se divide el sistema, corroboran su utilidad y efectividad.

Abstract

Road traffic accidents are a serious socio-economic problem, where the cost of human life is impossible to evaluate, and cause massive and continuous government spending. Different solutions have been proposed to reduce the effects of accidents, one of which, Advanced Driver Assistance Systems, forms part of the framework which encompasses the current investigation work presented in this thesis. These systems, as their name suggest, assist the driver by providing vital information on the traffic environment or by acting under specific circumstances to safeguard the occupants of the vehicle, or to facilitate driving.

A multitask driver assistance platform is originally presented as part of the research work of this thesis, among other tasks, it has been designed to recognize road signs in both urban and non-urban environments. The road signs that have been considered are: prohibition, danger, yield, obligation and indication. The information obtained from the road sign recognition process forms part of a complete module that advises drivers on traffic circulation requirements. The primary objective of this work has been to increase driver awareness, at all times, on the legal limitations which have established for road safety. The two areas which have been considered in this thesis are the velocity of the vehicle and the velocity of the vehicle corresponding to incorrect driving manoeuvres both of which are controlled using the information contained within road signs.

The assistance platform which has been designed forms an integral part of the vehicle, thus it must satisfy two important requirements: first, provide the driver with real time information from road signs and secondly, to operate in both urban and non-urban environments. Real time information is important for the safety of drivers, passengers, and pedestrians where the information provided warns the driver well in advance of any danger so that the appropriate manoeuvres can be made to correct the speed of the vehi-

cle. One important aspect of the work presented here is that the system has also been designed for non-urban environments, such as: national roads, toll roads and motorways where there is a higher probability of more serious and fatal accidents occurring due to the increased speed.

There is a wide range of possible applications for road sign recognition systems, another area of interest which has motivated the work carried out for this thesis has been an automatic road sign inventory system.

From the beginning of research in automatic road sign detection applications, many different stages have been proposed, such as: signal detection, road sign recognition, and road sign tracking. The work presented in this thesis provides an in depth analysis of each of these three stages and this has allowed a more robust and complete system to be designed.

In this thesis an exhaustive review is presented on color spaces and their characteristic color components which are best suited to the task of searching for road signs within an image. The technique of template matching using patterns with road signs has been optimized in this research, also included is an analysis of the most adequate models required to efficiently detect each type of road sign.

As part of the development of the recognition stage of the system, the two currently most important tools used in object recognition have been studied, these are: template matching and neural networks (NN). A comparative analysis of both of these techniques has been performed, where emphasis has been placed on image preprocessing to optimize the results.

The final stage of this system addresses the problem of road sign tracking. The method proposed is a model based on the movement of the camera within the vehicle with respect to the road sign which is taken as the reference point.

All the different stages of the system, which have been developed, form part of an experimental platform used on-board a test vehicle. To investigate the viability of this system the experimental trials have been carried out under real conditions. This methodology has been used for each stage, and the results presented corroborate the advantages and effectiveness of a multitask driver assistance platform.

Contents

Chapter 1

Introduction

Road traffic accidents are one of the main health risk problems. Globally, as the number of vehicles on the road increases so too does the number of fatalities and injuries. As a consequence of road accidents, every year approximately 1.2 million people are killed and 50 million disabled or injured [59]. Not only are road traffic accidents the eleventh cause of death in the world, but it is the only cause of death among the first twelve which is not related to illnesses or diseases. [table 1.1]. Injury assessment data from the World Health Organization (WHO) has shown that in the year 2000, 23% of all deaths have resulted from road traffic accident related injuries making it the main cause of mortality amongst other death related injuries. (Fig. 1.1). Future predictions show that the number of deaths and casualties worldwide from traffic accidents will increase to 65% by the year 2020 [96][114].

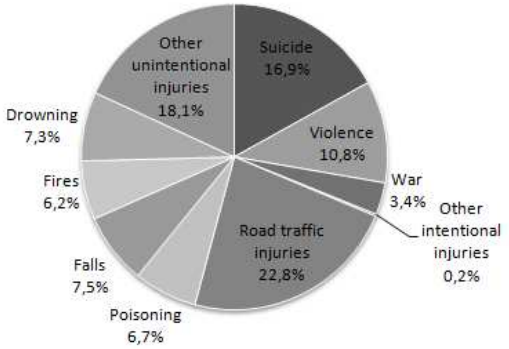


Figure 1.1: Distribution of injury global mortality by cause.

Rank	Cause	Proportion of total (%)
Deaths		
1	Ischemic Heart Disease	12.6
2	Cerebrovascular Disease	9.6
3	Lower respiratory infections	6.6
4	HIV/AIDS	4.9
5	Chronic obstructive pulmonary disease	4.8
6	Perinatal conditions	4.3
7	Diarrhoeal disease	3.1
8	Tuberculosis	2.8
9	Trachea, bronchus, lung cancer	2.2
10	Malaria	2.1
11	Road traffic injuries	2.1
12	Diabetes Mellitus	1.7

Table 1.1: The 12 leading causes of mortality

The nature and dynamics of this problem was first discussed in 1962 [116] but it was not until 1974, when the WHA27.59 resolution [29] declared road traffic accidents as a serious health problem, and urged the countries involved in the World Health Assembly to tackle the problem as well as enforcing "manufacturers to apply safety principles in the development of new types of vehicles". As a result of continuous awareness, from this period onwards, in 2003 the *World report on road traffic injury prevention* [59] was implemented by the United General Assembly [30]. This document contains a comprehensive study on the magnitude, risk factors and impact of road traffic injuries and considers prevention techniques to reduce the impact of accidents. In this report the efforts needed to respond to this problem have been summarized including the following tasks:

- a scientific approach to the problem.
- provision, careful analysis and interpretation of data.
- setting-up of targets and plans.
- creation of national and regional research capacity

– institutional cooperation within the different sectors.

As part of the scientific approach, one of the main areas of investigation includes prevention methods and minimizing the severity of injuries from road accidents. This is carried out using special design techniques and implementation of new technologies.

Socioeconomic Costs of road traffic accidents

The economic costs are colossal when considering that high-income countries spend almost 2% of the Gross National Product (GNP) on injuries due to road traffic accidents [87] this was estimated to be 518 billion US$ in the year 2000. According to the WHO, this mainly affects the most productive group with ages between 15 and 44. This situation is particularly serious for low-income countries where the capacity to generate earnings relies on more physical activities. Ironically, there has been little investment in road safety research and development when compared to other areas of diseases and other injury related illnesses [table 1.2].

Disease or injury	US$ millions
HIV/AIDS	919-985
Malaria	60
Diarrhoeal diseases	32
Road traffic crashes	24-33
Tuberculosis	19-33

Table 1.2: Estimated global Research and Development funding for Selected topics [59]

It should also be considered that the consequences of road traffic accidents are far greater than those expressed in financial terms. The psychosocial costs are not quantifiable, this is not only due to the emotional stress caused when a loved one is lost but also the psychological and physical condition of the injured individuals: chronic physical pain from ankles, knees and cervical spine, the trauma associated with permanent disfigurement, loss of eyesight, and brain damage etc. all of which result in a considerable load on family members[35].

The main risk factor: speed

The speed of vehicles on the road is directly related to the risk factor associated with accidents and is also responsible for the consequences [33]. Excess velocity on the road within areas of a determined speed limit and inappropriate driving speed are the main cause of road traffic accidents, i.e. driving at an excess velocity when considering parameters such as: the driver, volume of traffic, condition of the vehicle and road. As much as one third of road related fatalities are due to excess vehicle speed.

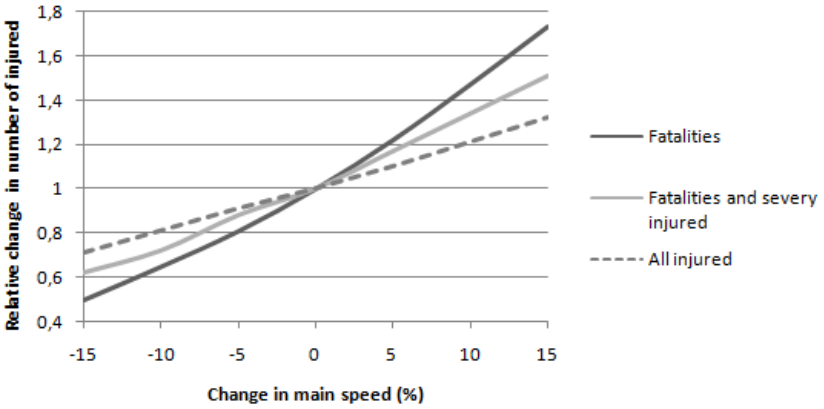


Figure 1.2: The number of injured people versus the change in the main speed according to the power model [115]

Much research work and study has been performed in this area providing interesting results. In [115] it has been determined that the probability of an accident involving injury is proportional to the square of the speed, a serious accident to the cube of the speed and fatal accidents are related to the fourth power of the speed, this is shown in Figure 1.2. Moreover, an increment of 1Km/h in mean vehicle speed typically results in a $\pm 3\%$ increase in the occurrence of injury related accidents (or an increment of $\pm$4-5% of fatal accidents). It can be concluded from these results that there is a more likely event of loss in control of the vehicle and failure to anticipate possible dangers. As the speed of the vehicle increases so too does the distance traveled during the time the driver takes to react also the distance required to stop the vehicle increases, see appendix B for further information. The probability of error and the consequences of the error are both increased. From a physical point of view, the effect of velocity when an impact occurs is devastating to

the body as it must absorb the kinetic energy, this is also dependent on the square of the speed. For example, in a collision against a stationary object at 50 km/h a 5 kg baby experiments the same change in energy as if it had fallen from a height of 10 meters (although the greater part of this energy is absorbed by the deformation of the vehicle).

Measures taken for reducing speed

Different initiatives have been taken in order to reduce the speed where necessary, these include physical obstacles which force the driver to reduce the speed of the vehicle. Other methods are based on the use of devices placed on the road and on-board vehicle technologies that inform the driver or take control of the vehicle.

- Engineering Treatments

Some of the most common methods used to slow down vehicles are the positioning of obstacles on the road, this forces the driver to reduce the vehicle to the appropriate speed that is required to pass the obstacle. Fig. 1.3 depicts some of these devices.

- Speed humps: Generally considered to be the most traditional traffic calming devices, these are an extremely effective means of controlling vehicle speed, slowing them down to velocities between 20-40 km/h.
- Speed cushions : these devices are used to slow vehicles down without affecting the speed of emergency vehicles. Speed cushions slow vehicles down to velocities between approximately 25-40 km/h.
- Radar Speed Signs: This device displays the speed of oncoming vehicles using highly visible LEDs this increases the drivers awareness on the current vehicle speed.
- Alley Bumps: These have been developed specifically to address speeding and cut-through traffic in city alleyways.
- Pedestrian platforms: A section or area of the roadway specifically textured or raised to slow down vehicles that are intended to provide a crossing point for pedestrians.

  - Roundabouts: Raised islands placed on intersections that slow vehicles down by forcing them to drive around a circle. In some countries they substitute traffic lights as they allow more fluent and safer driving.

- Technology

According to the resolution stated in [59], research must take an active part in the development of technologies to reduce the level of road traffic accidents and their consequences. Some of these initiatives are:

  - Road Speed Limiter (RSL) is a device that is installed in vehicles to limit it's speed to a preset value, this makes it impossible for the vehicle to exceed the legal maximum velocity. This device is typically installed in trucks and buses.

  - Event Data Recorders (EDR) and Vehicle Status Data Recorders (VSDR): these devices obtain data from airbags, yaw angle, stability sensors, antilock brakes, traction controllers, throttle controls and engine monitors. They are only activated when the sensors detect that the vehicle is about to crash. The latest version of these devices are VSDRs, these devices are continually active and collect additional data such as wheel and engine speed. It has been reported that their use has resulted in fewer collisions and encourage drivers to drive with increased caution [31].

  - Advanced Driver Assistance Systems (ADAS) this is an on-board vehicle device which focuses on the driving process. One of the main objectives of this technology has been to increases driver awareness by providing useful information.

    There are a large variety of systems on the market, some of those are shown in Figure 1.4 for example Blaupunkt has its own version of a navigator which is used to guide drivers to a user defined destination [7]. A device developed by Mercedes that monitors the space available on both sides and behind the vehicle and that is capable of detecting other vehicles. It notifies the presence of other close by vehicles when changing lane [10]. The parking assistant developed by Bosch [11] employs a total of six sensors on the front and rear of the vehicle and measures the parking space which is indicated to the driver using

(a)

(b)

(c)

(d)

(e)

(f)

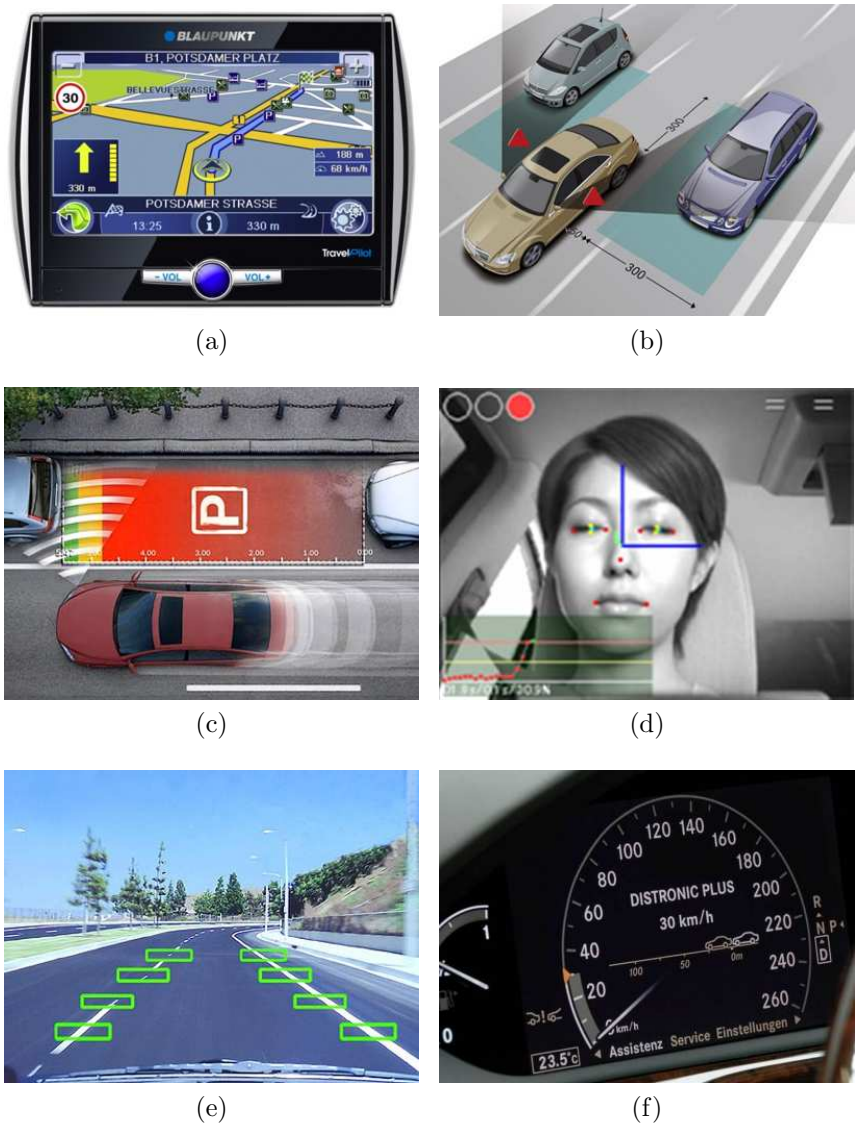
Figure 1.3: Engineering treatments for lowing speed: a) Speed humps. b) Speed Cushions. c) Radar Sign. d) Alley bump. e) Pedestrian Platform. f) Roundabout.

an acoustic emission whose characteristics are in function of the size of the space available. A driver monitoring system has been developed for Nissan which monitors the attention of the driver and detects possible symptoms of drowsiness [9]. The lane departure system of Iteris [20] warns drivers of unintentional lane changes in areas where the lanes are marked. It detects both continuous and discontinuous lines even when the road markings are not clearly visible. Mercedes has constructed a multipurpose device, DistronicPlus, which among other features has an adaptive cruise control function. This is used specifically for traffic jam situations where the system can take control of the car and maintains automatic user preselected security distances between vehicles by braking and accelerating [8].

When concerned with the speed of vehicles, attention here is focused on its reduction and adaptation to traffic and road conditions. With this in mind, Intelligent Speed Adaptation (ISA) systems are already available on the market. These systems inform the driver of the permitted or recommended maximum speed for each particular type of road. The most common data source related to road speed is provided from GPS data and radio beacons. This information is contained within a database in which the appropriate speed limits are stored, as a result the driving speed of the vehicle may be corrected at any time to be within the legal limits. The main disadvantages of this type of system is that it is only effective on highways as it is very difficult to store every road sign located within urban environments. Another disadvantage when using these systems is that they depend on a database which cannot be updated each time a new road sign or temporary road signs are being used (i.e. accidents or road works).

Studies which have been carried out on reducing fatalities resulting from road traffic accidents when using these types of systems has been presented in [118], where it has been estimated that this reduction ranges from 18-25% when advisory systems are used (i.e. the driver is informed of the speed limit and when it is exceeded).

It is also important to emphasize that these initiatives are supported by the end user i.e. the driver. The main goal of the SAFETY-TECHNOPRO project [32] has been to elaborate and test a training system for professionals on car safety technologies available in Europe. A lot of emphasis has been placed on issues related to Real Time Traffic Information Systems (RTTI)



(a) (b)

(c) (d)

(e) (f)

Figure 1.4: ADAS: a) Blaupunkt Navigator. b) Blind spot detection system developed by Mercedes. c) Parking assistant from Bosch. d) Drowsiness detection device by Nissan. e) Lane departure system by Iteris. f) Multi-function display with the information of the distance between vehicles of the adaptive cruise control of Mercedes.

where the driver is informed on en-route traffic conditions. The public acceptance for this technology is shown to be very high:

- 73% consider safety the most important factor when buying a new car and 72% consider RTTI a useful system.
- More than 80% of those surveyed prefer to have safety systems in spite of the fact that the safety alerts may be annoying.
- Most of the people surveyed prefer to manoeuvre their own vehicle rather than trusting an automatic steering control system.
- 84% agree that using assistance systems is very comfortable and it increases the ease of driving.
- 60% consider that the best way to operate these system is when their activation and deactivation is controlled by the driver.
- 83% would like to have an RTTI on-board system.

Summarizing, it may be concluded that:

- Road traffic accidents cause a large amount of fatalities, injuries and socioeconomic problems.
- The primary cause of accidents is directly related to vehicle speed.
- Implementation of different technologies is vital for decreasing the amount and consequences of accidents.
- Advisory systems are reliable and a publicly accepted method to reduce the amount of accidents.

Proposal

The main objectives behind the research work presented in this thesis has been the development of a computer vision system for real time road sign recognition, where this system must work for both highways and urban environments. It will be shown throughout this thesis that the conditions vary for the two environments.

A photograph of this system is shown in Figure 1.5. This system will be used to supervise vehicle velocity but it will be shown that it is an ideal platform to develop other systems, i.e. road sign inventory systems. The information is obtained from two different sensors, a color camera and an inertial sensor. The camera operates continuously and acquires images of the road, these are used as the input of the recognition thread that detects, classifies and tracks road signs. The second device, the inertial sensor, will provide the system with important data required by the warning system (speed) and the tracking stage (time, angle and speed).

The system presented in this thesis recognizes danger, prohibition, obligation, indication and yield road signs where the main focus of this work has been on road signs which provide information on the vehicle's speed, i.e. a yield road sign does not include any additional information on vehicle speed but the rules of the road indicate that the speed of a vehicle should be reduced when approaching this road sign. In this, and similar cases, a threshold speed limit will be set manually by the user. Failure to comply with the requirements of the threshold level or the information provided from road signs while driving will activate a warning message to the driver.

The system that has been developed will filter the messages avoiding unnecessary disturbance and also prevents incorrect vehicle manoeuvres. The characteristics of this filter will depend on the vehicles compliance to the speed rules which are indicated on the recognized road signs. The system will only broadcast a warning message when the speed of the vehicle is greater than that indicated by the road sign.

There are two different driver warning mechanisms, these are an audible acoustic signal and a graphical display system. Both of these are based on a PDA application which provides a pop-up interface for graphical displays and uses the PDA's loudspeaker for the acoustics signals. This configuration is similar in appearance to commonly used navigation systems, the reason for choosing PDA technology is based on the current well established user familiarity with this device, another inherent advantage is that the speed control application may be used along with standard navigation system applications thus removing the requirement of a separate speed control device.

As part of the driver warning system, if the situation arises where the speed of the vehicle does not comply with the preselected threshold levels set by driving rules or the user preset thresholds values, a picture of the recognized road sign will be shown on the screen as well as an audible acoustic signal which indicates the correct manoeuvre to be carried out by the driver.

The system that is presented in this thesis deals with the lack of upcoming and commercially available products. Those, which are already available on the market, will be discussed in detail in the state of the art section of this thesis. The problems associated with road sign illumination, such as poor illumination conditions, shadows and backlight will be dealt with for both urban and non-urban environments. A database with a large range of road signs is also used for the recognition process along with a message filter, this avoids any unnecessary disturbances to the driver indicating only relevant information and warnings.

Document Organization

This document contains 6 chapters which are summarized as follows:

- The second chapter covers a review of the areas related to the state of the art of this research work. The different techniques, methods and algorithms are all presented here along with a general view of the evolution of systems which are similar to that developed for this thesis.
- The third chapter gives a detailed review on the detection method which is based on the color and shape of objects. Both the color spaces and the models used to detect the road signs will be presented here. Experiments which were carried out to optimize the results are also described in this chapter.
- The Fourth chapter presents the research related to recognition, this is based on the two most commonly used techniques: template matching and neural networks NN.
- The Fifth chapter describes the tracking system which has been developed as part of this thesis, this system is based on the information from the inertial sensor and a camera motion model.
- Finally, in the Sixth chapter, the conclusions to this work and indications to future areas of investigation based on this research will be presented.

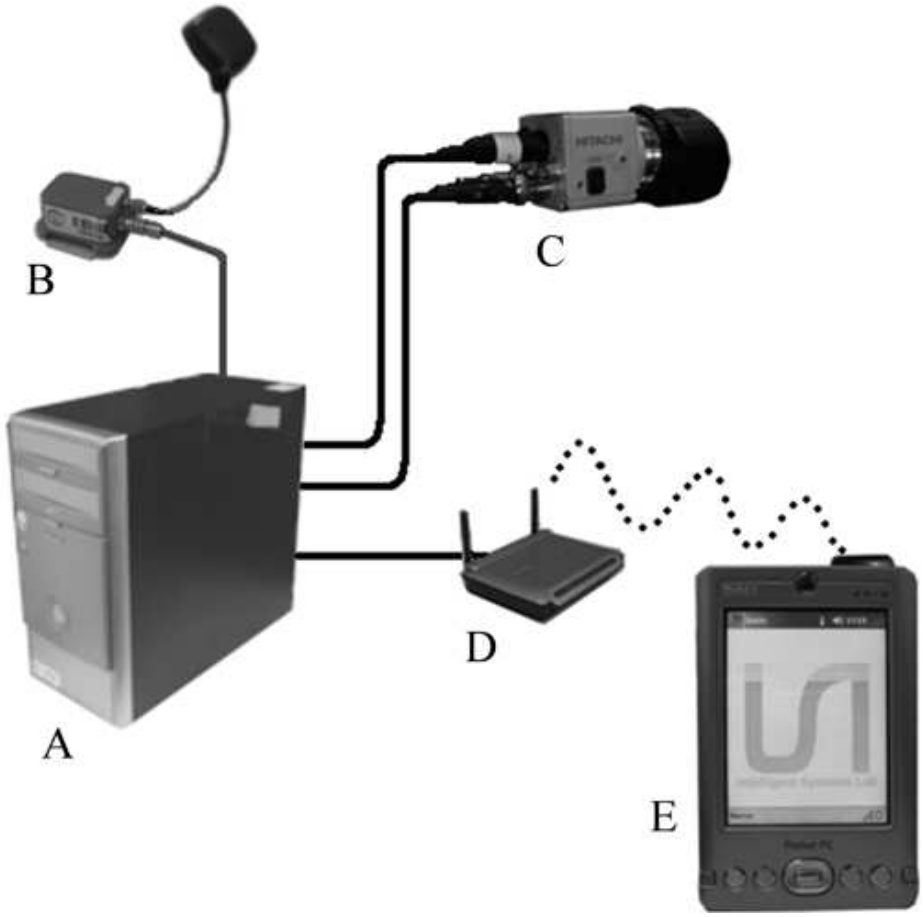


Figure 1.5: System proposed. A) Computer where the information of the inertial sensor B) and the camera C) are integrated. When necessary the system broadcasts a message through the wireless router D) to the PDA E) where the information of the road sign will appear and an acoustic message will warn the driver when required.

Chapter 2

State of the Art

2.1 Overview

There is a constant increase in the amount of research being carried out on the recognition of road traffic signs. This increase does not only account for the different types of detections schemes but also on the focus of different objectives. Initial work began with the detection of road signs as a possible extension to research on robot navigation. As it will be shown later, investigation in this field of research has become important because of the wide range of possible applications, one of the many examples is the Program for European Traffic with Highest Efficiency and Unprecedented Safety (PROMETHEUS). The ultimate goal of this project was to design an on-board computer system to monitor vehicle operation, provide the driver with information, and assist with driving tasks.

One of the tasks laid out in this thesis has been to design a road sign detection stage which still remains an active area of investigation, this is due to the large number of problems associated with outdoor detection such as changes in illumination, road sign rotation, aging and vandalism.

Once the road sign is detected, the next stage involves the recognition of its content, that is, the allocation of the road sign to the information it provides. The input to the recognition stage depends on the results from the output of the detection process. Different approaches to recognition of road signs have been made and, as it will be seen in this chapter, a good rate of reliability in the recognition process has been demonstrated. However, this rate seems is still not sufficient to fulfill all the conditions that an on-board

vehicle security system requires. For this reason, commercial detection and classification systems have very limited detection capability where only a small set of road signs may be recognized. (Several examples of commercial systems are presented later).

More recently, tracking of road signs, once they have been recognized, has been added to the list of tasks that a road sign recognition system must address. This area of recognition has been implemented as part of the research work presented in this thesis. The three main stages of the system designed and presented in this work are summarized as follows:

- Detection: The objective of this stage is to separate road signs from their background environment, avoiding false detections due to characterization of the pixels as belonging to a road sign, or lose the information from pixels that are incorrectly classified and that result in a lower probability of detection. To standardize the criteria, it is considered that the detection process is terminated when the system is certain that a road sign has been detected. This process is carried out even though there is no knowledge of the type of road that has been detected.

- Recognition: This stage is commonly referred to as classification, where different algorithms are used on the information from the detection stage to sort the detected road signs. This step is vital for the safety of the driver, especially in the case where danger or prohibition signs are recognized.

- Tracking: This is the final stage of the system. The incorporation of a tracking stage allows the system to be certain that a road sign has been detected. This is done using continuous positive detection of the road sign in different image frames. To perform this task the system is generally required to predict the position of the road sign on the screen. This accelerates the detection of a particular road sign as the search area may be limited to a much smaller area.

There are many different approaches and viewpoints on each of the stages presented above. Hence, it is difficult to establish a homogeneous state of the art for this technique. A large variety of methods may be used at different stages and sub-stages, where these methods can be used for very different tasks within each stage. The state of the art approach taken here follows the system requirements.

2.2 General Problems

– One of the unresolved problems within the detection stage is related to changes in ambient lighting conditions and its influence on the color of the road sign (Fig. 3.44a). Although road signs have a characteristic color, their colors are observed to change during the day and also when exposed to different weather conditions, such as: clouds, fog, shadows caused by trees or buildings or other objects, these factors lead to a reduction in the saturation of the road sign color. Another important cause of the change in road sign color is due to back light, this occurs when light from behind the road sign directly enters the camera, i.e. sun light. As a result of these problems, most of the documents listed here tackle the recognition process under the hypothesis of daylight and lighting conditions within this period. Generally this problem is faced using color spaces which are not as variable under illumination changes, another method used is to define experimental thresholds which depend on each special case. An exception to these techniques has been presented in [43] where gamma correction is developed to reduce the dependence on the color of the light source.

  Very little investigation work has extended the goals within this field, this includes road sign recognition during night time conditions [36] [167], which is due to the difficulty that this this task presents even during the best weather conditions. An interesting experiment on changes in the perception of road sign color depending on the time of the day can be found in [39] where it has been shown that a straightforward comparison of the RGB components taken two by two is sufficient to segment road signs.

– Occlusions of road signs due to the presence of objects such as trees are another important factor which must be considered for efficient road sign detection. When part of the road sign is covered the detection and classification of the road sign becomes very unreliable and also increases the difficulty in the recognition of the information contained within the sign, this situation is shown in Fig. 3.44b [154].

  To avoid this problem [52] have applied deformable models, since the system uses a large number of different features to detect the road sign it is assumed that some of these may be lost during the detection process. The results observed are only slightly affected and maintain a

good level of reliability. For recognition a NN is used which erases part of the information, here the system considers that there is a possible loss in data, using this technique over 80% of correct classification results have been obtained.

In [100] the authors have provided a description of the recognition technique where every geometric shape has four independent parameters, this makes the classification process more robust and achieves over 98% of correct classifications, however the probability that a road sign is not detected is 47.62%.

In [145] and [146] geometric fragmentation has been used, this is effective when dealing with this problem for circular road signs. This method has achieved a rate of 87% in detection assuming that the occluded area is less than half of the road sign, and is part of the left half, the right half or the middle. When occlusions cover the complete left or right half of the road sign the contour can be extracted using both the left and right fragments in the same way as the case for no occlusion.

Finally in [22] the authors have achieved over 70% of correct classification results when including the Jones approach to image occlusions.

– In-plane Rotations (Fig. 3.44c) or out-of-plane rotations (Fig. 3.44d) and the angle of vision can also cause problems in detection and recognition schemes. As the information is not complete robust algorithms must be developed to reconstruct the majority of the data or to manage the little information available which is required for the detection and recognition process. Some solutions to this problem may be found in [24][153][101]. The latter relates all objects of interest to a reference position, the authors conclude that the system is highly robust when dealing with typical road sign rotations. A method developed to avoid effects from skewing has been presented in [161] where the road sign is reconstructed to obtain a face on view of it.

– Aging and vandalism also affect the color of road signs. In the first case the deterioration not only produces a loss of color saturation but also other damages such as holes or loss of part of the road sign (Fig. 2.1e). In the second case, the information is typically covered or misrepresented, this affects the recognition process as may be seen in Figure 2.1f.

An interesting laboratory experiment on the effects of aging and vandalism has been developed in [143] where the luminance and the reflectivity of the road sign is evaluated using different cameras, this is done to provide an automatic recognition of the deterioration of the reflective sheeting material used during road sign fabrication. The authors conclude that the luminance and the retro reflectivity remain constant with respect to the distance when the illumination of the sign is constant, this provides the basis for a reflectivity evaluation of vertical outdoor road signs.

Finally, in [86] a degeneration model of road signs is introduced, this is based on different parameters which are used to detect road signs, the authors have obtained correct classifications for over 95% of the road signs and with more than 12 pixels.

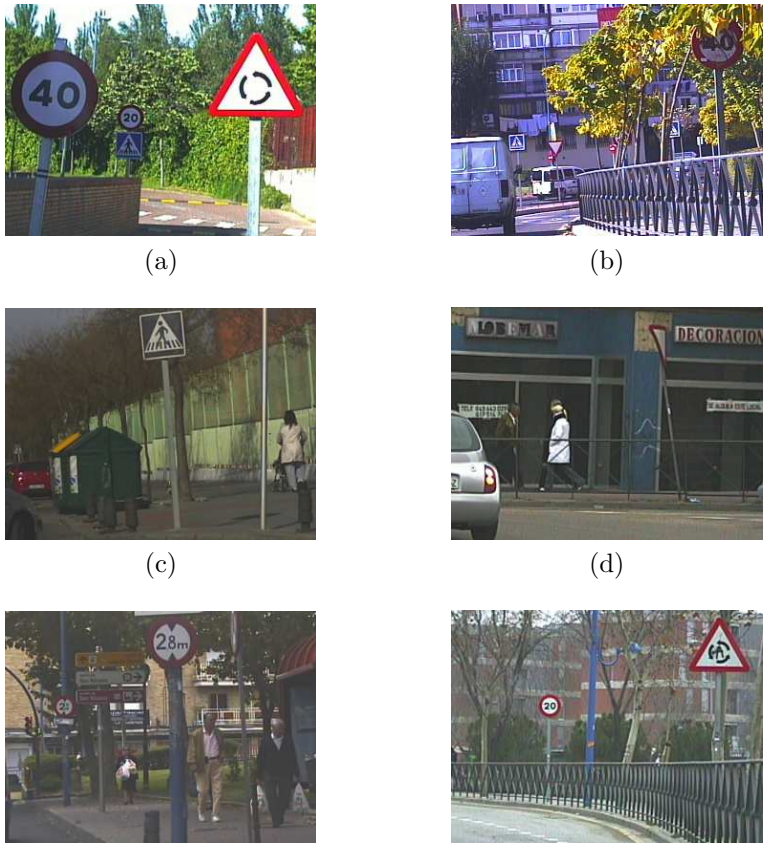
Several algorithms are useful to avoid these problems. Early work in this area, such as in [91] explicitly present a system that reduces the influence of noise, scale, rotations and translation in detection by applying geometrical transformations in the preprocessing stage and by using a NN for robust classification.

Almost all of the aforementioned problems have been faced in [125] by using different composite filters and a bank of nonlinear filters for scale, in-plane and out-of-plane rotations. In [49] the authors have presented a method which tolerates small translations, rotations, scale changes, lighting conditions, noise and small occlusions, this has been based on histogram matching. However, this system requires road signs sizes of over 100 pixels to achieve a recognition rate of 89% for prohibition road signs and 91% for obligation road signs.

2.3 Detection

In the detection stage of the system the main goal is to obtain good quality road sign segmentation. The extracted regions are then used to decide whether it contains a road sign or not. These regions are then used as the input for the recognition stage.

The characterization of road signs can be carried out using three different features: shape, color and the inner symbol. This allows each sign to be completely labeled and provides a complete solution to this problem. During



(a) (b) (c) (d) (e) (f)

Figure 2.1: DetectionProblems: a) Illumination problems: the red color of the road sign is observed to change dramatically for different levels of incident light, b) Occlusions: part of the road sign is lost, thus detection and recognition become increasingly difficult, c) Rotation in-plane and d) Rotation out-of-plane: this changes the normal perspective of the road sign which affects detection and recognition, e) Aging: this is commonly observed in cities, as a result the color changes and so too does the detection quality, f) Vandalism: the information of the road sign is misrepresented thus the classification process may be impossible or incorrect.

this stage, the shape and color of the road signs are the features that are considered to be of greatest relevance. Therefore, the research work, regardless of the method used, focuses on obtaining as much of the data as possible related to the first two features.

In the detection case the color is an important source of information. However, the main drawback to designing a road sign detection system based on color is that it varies according to changes in ambient lighting conditions. Currently, there is no well established technique to solve this problem.

Several authors have reduced the influence of the illumination by working with grayscale images, further details using this technique will be presented here. These authors have managed to obtain the same information from these grayscale images as that contained in color images in areas where the ambient lighting does not pose any problem. When the image is dark, it is considered that not all the color information is contained within it, thus it is assumed that information is not lost when compared to the color image case. Also it has been deduced that searching for edges, corners or contours within the image does not require color information. Some examples of this process can be found in [56][28][41][40][127][128][133][155][138]. grayscale images were most commonly used and studied in the 1990's, however this technique has been recently used in [146] ,[22] and [52] to detect the inner part of the road sign and in [44] an algorithm has been developed for end of ban road sign detection.

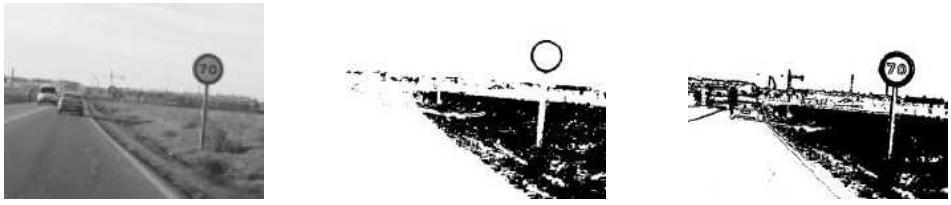
The most common strategy employed in the detection process is based on the use of different color spaces, these are not sensitive to changes in illumination. An example of this is presented in Figure 2.2 where it may be seen that the results obtained provide a grayscale image where only one of the components of the color space is represented. Little work has been performed on the use of two components, H and S [112]. Unfortunately, only a limited amount of comparative data exists for results obtained from different color spaces, thus it is only possible to provide a list of the most typical conversions. These results have not been recorded as it is evident that the quality of the enhancement using any of the so-called illumination-invariant color spaces is improved when compared with the straightforward RGB enhancement process. One of the exceptions to this has been presented in [63] where it has been stated that the saturation and intensity are not required for detection, however no results are presented to this effect. In [112] the authors have made use of the saturation along with the hue as this second component is also seen to change with intensity. Several experimental

procedures have demonstrated that, in the case of the red rim color, when the HSL space is used the results when considering the hue are only slightly improved when compared to those obtained from both the hue and saturation [45].

The different color spaces may be summarized as follows:

- RG-BY space from RGB through log transform has been developed in [95].
- HSV for pixel classification used in [130], this may also be found in [85] [34][128][121][153][109][160].
- HSL in [89][84], and a variation of it in [68] using the algorithm presented in [80] which introduces an improved version of HLS called IHLS, this provides independence between chromatic and achromatic components.
- HSI as in [94][26][162][27][63][112][52][141][100][166][101][99][108].
- Other color spaces which are not as common as the aforementioned are XYZ used in [156] and [167]; Lab [144];YUV [111]; YIQ [92]; uvh [91]; YCbCr [50] and HCL using CIECAM97[72] and [71] where a comparative analysis between HSI, HCL and CIECAM97 color spaces is performed. The results show that CIECAM provides better results when compared to the other three color spaces with 94%, 90%, and 85% almost an 8% improvement in the rate of accuracy for sunny, cloudy, and rain conditions.

Some authors do not consider the drawbacks due to illumination as part of their goals, research work is carried out directly on the RGB, examples of this



(a) (b) (c)

Figure 2.2: RGB to HSI conversion as in [101]: a) Original image, b) segmentation mask by red, c) segmentation mask by achromatic decomposition

are presented in [77][62][39] and [158][140][167][145]. In [24] the authors add an experimental threshold to obtain different dominant colors. Experimental thresholds are also commonly used to limit the regions of colors of interest as presented in [54] and [43]. Also In [147] and [109] a manual threshold has been set to transform the space into the colors red, white and black (RWB).

Another way to tackle the color problem is to make use of color normalization algorithms, these are applied over an image making it independent of both the color of the illuminant and lighting geometry. The goal here has been to group similar colors under different assumptions, although the final result generally is not in accordance with standard human perception of colors.

Some examples are found in [161] where the references of white and red and the illuminants are obtained using a simple value decomposition. No results are presented but the authors have stated that the algorithm achieves good qualitative results, one drawback to this technique is the time required to execute the process, for example a 768×576 image takes 1 minute.

In [104] the authors have used the color formation equations presented in [103] to predict the color responses of the camera to different road sign colors. The last example, [142] have performed a color standardization based on a NN mapping and have reduced the one million different colors to only 5: black, white, red, blue and yellow.

A special area of investigation has involved a study in the absence of color. Although intuitively this scheme seems straightforward, it is not possible to set the borders of the achromicity region for any of the spaces i.e. the zone where no color assignation can be carried out. In the RGB case, the representation corresponds to the main diagonal of a cube from the point of origin to the point where the maximum values of the components are reached (gray line) and its surroundings. The problem is to set the limits of the surroundings. This approach has been presented in [94][27][104] and [108]

Apart from the methods which directly manage the information from the RGB signal, other algorithms have been developed to extract the color, for example, using growing zones of color.

Reference seeds have been used in [161] from which pixels of a similar color are added. An original system based on a specific distance has been developed in [130] where the basics of this work may be found in [132][135]. The use of NN for the assignation of each pixel to a certain color has been presented in [165][156][63][137]. Fuzzy logic based methods have been presented in [85][49]

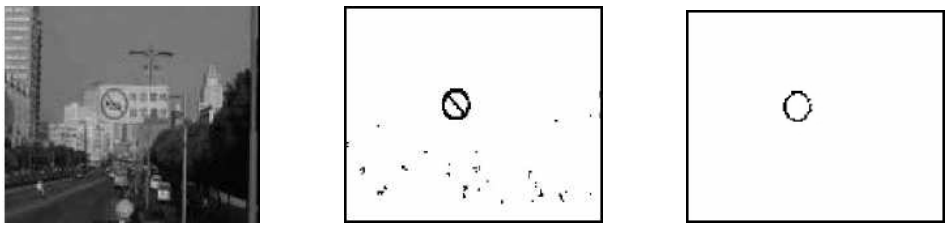
and a polynomial classifier in [69]. A Back propagation NN has been used for an eight color segmentation in [94].

After the image has been enhanced using a color space conversion, a grayscale image is obtained. This image usually contains the pixels belonging to the objects of interest and also the pixels from other different objects which appear in the same image. In many cases the information of these pixels, especially in the case of noise, perturbs the results, because of this several authors have applied filters to the image to remove or reduce the noise level (Fig. 2.3). There are various different methods such as noise cleaning masks [154] or median filters [93][89] which can be used to reduce sharp noises and increase picture stability. Morphological filtering criteria using a Hopfield network based on a set of rules has been presented in [77] where the pixels are activated or deactivated depending on whether they belong to an object or if they are surrounded by a certain amount of deactivated pixels.

Other methods used have been the eight neighborhood filter [25][158][109], here the final pixel value is the sum of the 8 neighboring pixels, where the result is twice the initial pixel value and the result is divided by 10.

Another way of to reduce image noise is to use binarization. This has been carried out by [144] where the authors have used a threshold along with the Otsu algorithm [119], this method is fast, simple and efficient. Morphologic transforms such as dilation and erosion which fill in some holes and smoothen out the shapes are presented in [56][93][91][95][166] and a combination of these, opening and closing, has been presented in [24].

In [92] a system has been developed which not only considers noise as the common Gaussian noise from an unclean camera, quality of the sensor or weather, but also from the change in color of the road sign, occlusions, distance, projection distortion and blurring due to motion.



(a) (b) (c)

Figure 2.3: Noise reduction: a) Original image, b) Color segmentation, c) Noise filtering. Figures taken from [166]

After the optional smoothing processing, several researchers add an extra stage to extract features:

– Corners or edges are the most typical features to be sought after when it is required to use the information pertaining to the shape of the road sign [63] [99].

The Hough transform is the most commonly used tool for this process [93][161][50]. An extension of the Hough transform, the fast radial symmetry transform, is commonly used to detect circular road signs, as in [37][126][67][144] and in [105] the authors have extended this technique to the recognition of other shapes, such as octagons, squares and triangles. The Canny (Fig. 2.4) has been used for edge extraction in [127] and a combination of both the Canny and Hough transform can be found in [73] where the authors have used the first method to perform a preliminary feature enhancement followed by the second for the final edge extraction.

A derivative filter has been used in [44] for line extraction from the end of ban road signs where this is followed by a Sobel operator which extracts the edge of the plate as in [93]. In the case of circular road signs the Hough transform is reduced to two dimensions which is less time consuming. However, other methods exist which are not commonly used, one example is fuzzy logic which has been used for corner detection in [90] to detect the shapes of down-triangles, up-triangles, circles, squares and octagons.

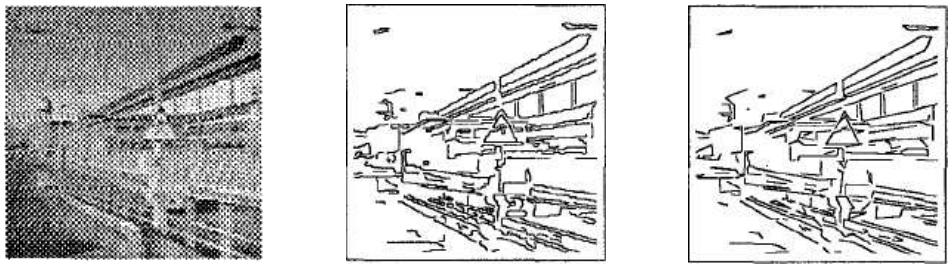
Different gradient methods have been used in [56]: the aforementioned dilation and erosion, Sobel, and differential Nagao Gradient. The first of these three is observed not to provide the best results especially when searching for round road signs which contain black and white information. The main disadvantage of the second method is that it is time consuming but provides good results, finally the third method is used to smooth out the image while enhancing the contrast.

An interesting comparative analysis between the different methods for recovering circular contours can be found in [128]: Hough transform; FEX [61] where after a three stage method, the segments are classified as straight-lines or circular arcs, where this classification is based on straightforward criterion; computation of the local curvature where the

segmentation of the edge data is carried out resorting to a rough calculation of the local curvature of edge chains and breaking the chain where the curvature significantly changes. An original method which allows partially corrupted circular contours to be recovered using a threshold which is based on the number of pixels of different annuli. The discussion has outlined that each method achieves, respectively, 87%, 98%, 95% and 96% in the case of "easy images" and 13%, 21%, 40% and 93% in the case of "difficult images".

Several authors make use of masks, these masks are matrices which are applied to images to enhance different features, such as: edges, contours and shapes. Some examples of masks may be found in [54] where the type of mask used depends on the type of corner related to the road sign, or in [84] where a 4x4 pixel pattern mask has been constructed to enhance features of horizontal, vertical, slope right, slope left, body, space and corner boundaries.

– Once the possible targets to be obtained have been limited, certain high-level features such as shape may be extracted posteriorly, as in [138] where a hierarchical grouping method has been completed to link low-level features of pixels, such as gradient orientations, to high-level features, such as triangles and ellipses. In [112] the authors have proposed a multi probability method in which a likelihood is associated to each feature to be extracted: first the color feature is used to provide a previous indication of the possible road sign this is then followed by segments of interest and shape probability functions which are used to



(a) (b) (c)

Figure 2.4: Edge extraction : a) Original image, b) Output of a Canny edge detector, c) Polygonal approximation of c). Figures taken from [127]

determine the detection, results as high as 96% for 16 different road signs have been obtained.

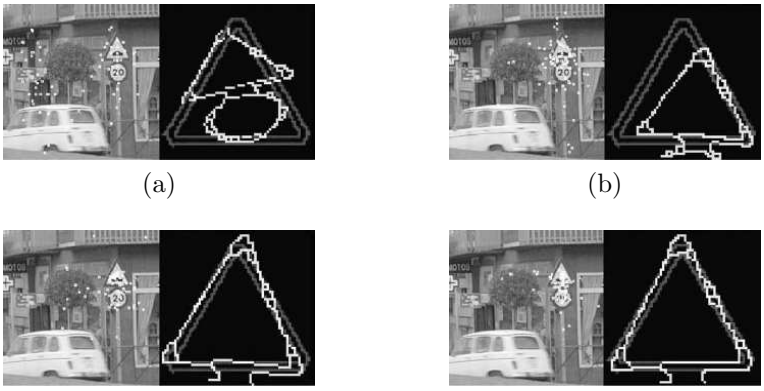
- Genetic Algorithms (GA) (Fig. 2.5) are also used to extract regions of interest, in [25] the mutation rate is observed to be high as when it is applied over one point for every calculation of the Fitness function. If the results are not improved using the mutation another point is tested. In [53] and [55] the initial models represent the road sign at a fixed distance and perpendicular to the optical axis. From this model and subject to an affine transformation an initial population of genes are created and varied so that they can set the contour of a real road sign for variations in scale, distance and rotation. One of the main points related to GA algorithms are the minimization functions used to stop the generation. This has been the case for simulation of annealing (which models the cooling of a material) in [52] or the Hausdorff distance in [55][145] which has been used to reduce the typically high computational cost of the GA.

- Gavrila has used in his research the distance transform (DT) concept [74][76][75] (Fig. 2.6) which is applied to the detection of road signs and other objects on the road. This method involves intermediate-level features which are extracted from different locations within the images. A DT converts a binary image into another image where each pixel indicates the distance to the nearest feature pixel.

Other algorithms which provide important features of the road signs, for example:

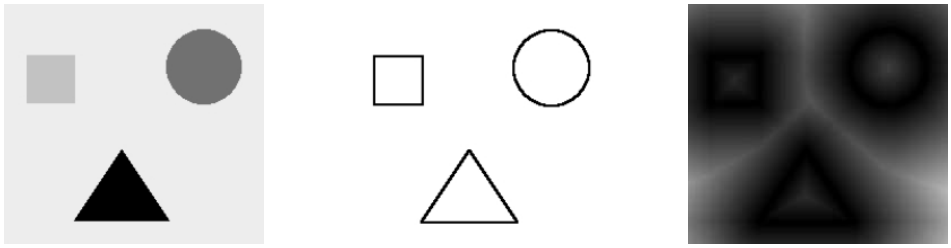
- Hierarchical Structure Code (HSC) developed by Hartmann in[82][83] has been used in [41][40]. This technique is based on the encoding of all the structural information related to an object for different levels of resolution. This multi-resolution description of the image allows objects to be described by a recursive process where this leads to the extraction of a chain code that defines the shape. The different fields of this structure are filled during the extraction process (e.g. bounding box or boundary length) or later by the shape recognition processes (shape class, surface, corners or coordinates).

- Some methods store and manage data that belong to the regions of interest (ROI) from the latter stage. For example a labeling process of



(a) (b) (c) (d)

Figure 2.5: Fitness function of the GA as in [55]: a) Best initial individual; (b, c) best intermediate results; (d) best final individual.



(a) (b) (c)

Figure 2.6: Example of distance transform as in [74]: a) original image, b) edge image, c) DT image.

the ROI is carried out in [137] for subsequent rejection of candidates. The road signs can be described by a set of primitive objects. This set contains geometric features which are related to the shape of the road sign such as circles, rectangles, triangles and special symbols such as arrows and textblocks. With these features and their geometric interrelation it is possible to model different road signs. To obtain knowledge of the contents of an image, an analysis of the symbolic description of all objects of interest is required. This is achieved using a frame, which is generated for every object in the image, and contains computed attributes (i.e. color, size, ...), computed relations with other image objects (e.g. has inclusion, has neighbors, ...) and references to procedures which compute an attribute or relation if required.

In [133] a perceptual grouping extraction has been introduced where the different pixels are first grouped, which is followed by contouring and segmentation in lines and arcs. It describes objects using their features which are invariant under the affine transformations. The objects in the image are described by sets of points of interest, these represent the corners, points of sharp concavities and convexities. It is possible to consider just three of these which are non-collinear with respect to the others and in this way define an affine basis. This method is also used in [42] where a vector is filled with parameters which describe an object subject to the transformation. The six-dimensional vector describes a rigid body motion and linear distortion of the object in an image with a certain position, rotation, contraction and skew. In [101] the authors have implemented two methods to extract feature vectors, one is based on the distance to the border vector and the second, on the signature of the objects. A similar method has been used in [95] where the relation between different distances is taken, in this case the borders of the bounding box and the contour of the road sign. It allows every road sign shape to be detected i.e. triangles, pennants, rectangles, pentagons, octagons, diamonds and conical shapes. The final results are satisfactory, producing only 28 failures for 400 samples (7%) for diamond and conic shapes and 7 failures out of 260 samples (3%) for the the rest of the cases. In [56] the authors have used the information from different moments (spatial, central, normalized) and also other properties such as compactness, perimeters and average gray levels to feed a decision making system for the detection. Two approaches have

been tested: first a decision system examines the values of the different aforementioned criteria and checks if they are within the given reliability interval. Results in this case have not been satisfactory as the system was not flexible for handling exceptions and when it does, it is two-fold time consuming. The second approach was an expert system which fulfilled all the requirements adequately. The candidates pass from one class to another if they satisfy the given model. Models are described by attributes which are associated to an error function which provides the final success result.

- In [146] the regions of interest in the image are divided in rings, this is followed by the extraction of a characteristic histogram for each ring. With this method it is assumed that the detection is robust for rotations. A similar method using artificial images to obtain edge orientation histograms is presented in [23]. In this case the road sign is divided into different subregions which change depending on the type of road sign. Thus, for each subregion the histogram is computed and stored and is used as an input for the classifier.

- Similarity factors have been applied to extract shapes in [153] and [160] where the regions of interest are rejected depending on the fulfillment of specific requirements such as size, aspect ratio and geometric constraints.

Once the feature extraction has been completed, different techniques to ensure detection are used:

- Template matching has been used to assure the shape in [147] [67] and [43][134]. Generally several models which represent the different shapes of the real road signs are proposed (Fig. 2.7). These models are then moved over the image and a correlation function is used to estimate the reliability of the detection. In [127] the authors have developed a double template matching technique: first to consider the best candidates and secondly, with a finer resolution for final extraction of the candidates. In the two stages the reliability function used was the normalized cross correlation function. A decision tree for shape classification and template matching has been used by Gavrila in [74][76][75] where the value of the Chamfer distance determines the final detection.

Figure 2.7: Circle models for template matching in [147]

- NN are commonly used in the final stage of detection after applying other algorithms for feature extraction. Examples of this are presented in [25] [153], a BPNN in [73] for shape extraction, and a Hamming-NN in [48]. A comparative analysis between different neural networks can be found in [50] however in this case only circular road signs have been detected.
- Given some features, probabilistic models like Dempster-Shafer rules have been used in [41] [40] to assign an object to a certain class, in this case the shape of the road sign.
- In [146] the image is divided in rings. A local color histogram of each ring is compared with a model. As the information is contained within the rings, this technique is robust for angular variations.
- The Viola-Jones cascade classifier [148][149] has been used in [22] with artificial images, also integrating classifiers to provide a combined prediction. A classifier that combines weak indicators has been used in [23].
- For shape recognition, [72] has used an extension of the Behavior Model of Vision for feature extraction. The tests include noisy images.
- Adaboost has been used in [36] to obtain the best wavelets which when rescaled can detect a set of different shapes.
- Support Vector Machines (SVM) are also used in this step, although they are more commonly used in the recognition stage. In [102] the authors have introduced two approaches for the vectors: signatures which contain information related to the road signs and edge distance.
- In [79] a method has been developed which compares the FFT signature of the candidate to a generic signature for each shape.

– In [109] the authors have used the region properties of the objects (e.g. area or center of the road sign) in a fuzzy shape descriptor.

– As part of [95] techniques have been carried out to perform hierarchical matching with a data base of the visual aspects of the road sign.

2.4 Recognition

Considering all the different technique described previously, the objective of the research work presented in this thesis has been to correctly classify the road sign. The input to the system can be either a candidate which is unknown if it is a road sign (so this recognition stage would have the dual function of classification and filtering) or a road sign whose class is required to be known. The main methods are:

- Pattern matching and correlation: Normalized cross correlation is one of the most commonly used techniques in template matching. In [111] the authors have used this method along with templates and images. This technique is well known for its robustness when faced with changes in lighting conditions and is also suitable for pattern matching in outdoor environments. As a result of these advantages, it is possible to recognize guidance road signs with 100% accuracy and speed limit road signs with less accuracy, 46.5%. In [128] the authors have presented a recognition system which makes use of the normalized cross correlation value to obtain the final positive recognition. The candidates are normalized to 50×50 pixels and tested over two possible templates, the first contains the circular road signs and the the second, triangualar road signs. The disadavantages of this technique is that it takes 0.5 s in a SparcStation ELC to perform this task, although considering that this experiment was performed in 1996 this time may not be considered as being of fundamental importance. This technique has achieved a recognition rate of 98% for a dataset with 107 road signs.

  Another example is presented in [42] where the authors have also made use of normalized cross correlation for recognition and include information on the complexity of the road sign (resolution) where this affects the results. The system is studied from the viewpoint of the features complexity, thus the results vary greatly depending on the complexity coefficient, results obtained have been in the range of 80%.

In [109] the authors have developed a template matching scheme with grayscale specimens of 16×16 pixels and have obtained an accuracy of 86% for danger and prohibition road signs. In [37] 8 different templates have been produced for each road sign and a normalized cross correlation is applied to subwindows of 5×5 for 3 different radii, thus for the final computation a $5 \times 5 \times 3 \times$s is used, where s stands for the number of road signs, the resulting accuracy is 90%. Template matching has been used in [167] for circular road signs during day/nocturnal external light conditions. Good results have been achieved, 80% during daytime, 55% during evening-time and 47% during nocturnal conditions with specimens of 128×128 pixels, it has been stated by the authors that the sytem provides the same results for cases of 64×64.

Template matching for circular shapes is presented in [144]. A simultaneous search for models with three different radii (10, 15 and 20 pixels) has been carried out and using single frame processing the system recognizes 43.92% of the road signs in a normal environment, i.e. daytime, 43.75% in the case of rainy weather and 6.92% during nocturnal conditions for images of 720×576 pixels. These values for detection and recognition have been increased to 81.2% from using a tracking stage.

It is possible to perform pattern matching not only with images but also with several important road sign properties or features. This has been considered in [104] where encoded data from images is correlated with a database of artificial road signs which includes deformations due to perspective and rotations, this increases the robustness of the system. Once the sub image which is extracted as a potential road sign it is normalized, each pixel is labeled using a 9-bit binary word, this provides the so-called image encoded template. The final correlation between the candidates and model is done using a straightforward word operations i.e. OR, IF, and AND and results of 95.6% have been achieved in detection and 90.4% in recognition.

- Neural Networks

– Radial Basis Functions (RBF): take consider the localized character of neurons and use locally supported radial basis functions as its activation function. In [99]a RBF and K-d tree for a two stage recognition system has been developed. The RBF neural network is used to identify the group the road sign belongs to in the first stage. This is then followed

by the corresponding K-d tree which is used to recognize the road sign. If the difference between the road sign and the recognized node is larger than a preset threshold this indicates that the road sign does not belong to the group. A correct recognition rate of 95.5% is achieved using this technique, however the set of images is small, allowing only 105 images.

In [75] and [69] the false positives detected in an image are rejected during a verification phase, where a RBF network is used as a pictograph classifier. The authors state that the performance of the system is adequate as there are generally only one or two false detections per frame and the system is capable of rejecting 95% of them, however the number of road signs that the system is able to distinguish is very limited, 10, which is not large enough when considering that the number of road signs is greater than 150.

In [165] the k-NN rule is used for road sign classification. It is intuitive that the k-NN rule does not consider that the different neighbors may provide different data on the neighboring road signs. This depends on the number of neighbors which are to be considered and the limits associated with the different decision regions. Thus, the aim of this technique has been to perform a weighted analysis of the evidence where this depends on the proximity of two objects, this is why a radial basis function is used with real images as training models.

Instead of the more common activation functions, an extension to wavelets has been developed, known as ridgelets, this has been used as the activation function in [159] as it demonstrates good properties for both localization in time and frequency and also the directional advantages which the visual cortex is assumed to have. Training of the network is done with noisy images of 32×32 and a formula is introduced to calculate the maximum number of neurons within the hidden layer where this depends on the features of the NN. Finally, results of the ridgelet network are compared to the common RBF and the BPNN, the recognition rate of the proposed system was almost 98% for 11 different road signs while for the BPNN was 72% and for the RBF 87%.

– Multi Layer Perceptron (MLP): currently this is the most popular neural network used for road sign applications. A calculation of a certain biased weighted sum of the neuron's input is passed as an activation level through a transfer function to produce an output that is used for

subsequent neurons. This is why the neurons are arranged in a layered feed forward topology (Fig. 2.8). In [88] the authors have used the NN trained with real images as models and have achieved 100% of recognition over the validation set with a reliability in all cases of over 80%.

The most typical training algorithm is back propagation, where the gradient vector of the error surface is calculated so that it converges to a better solution. NN's which are trained in this way are efficient for prediction and classification however, they are slow when compared to other learning algorithms. This training method is so widely used that only one other training algorithm example has been found from in the literary review for this technique, in [154], where the MLP is trained using the conjugate gradient descent algorithm. Real images are used, but artificial rotations and noise have been added for the training process.

The earliest work using BP may be found in [106], here the authors describe a system for a mobile robot which moves within an indoor environment and recognizes each sign in 4 seconds. The implementation includes a NN for three different distances: close, intermediate and far which correspond to 1, 2 and 3 meters, respectively. It has been stated that the images appear slightly different due the camera-sign distance. Only the results related to the error achieved per iteration during the training stage are presented, where the output of the correct neurons in layer 0 was greater than 99% and the output of all other neurons was below 1% after 1000 iterations, this signifies that the misclassification error is zero.

This example is regarded as a starting point for further research in this area: 5 different road signs has been the goal for recognition in [92]. System robustness has been studied under different noise conditions (centroid shift and occlusions). The results are satisfactory since the classification rate is demonstrated to be 98% when the centroid shift is below 4 pixels and 70% when the centroid shift is 8 pixels. It is also robust to occlusions, where 98% of recognition rates have been achieved for occluded areas below 10% of the complete sign and 50% when the occlusion covers 34% of the road sign.

5 road signs have also been recognized in [58] where multi-layer percep-

tron has been used, in this case at least 70% of the test vectors have been recognized. 82% were classified as "good" with the correct class neuron having the highest value. The probability of a false detection was below 4%, the probability of error was 2.5% and the system recognized the background 70% of the time with only one false detection.

One difference among the different approaches of NN is the use of synthetic or real models as training sets (Fig. 2.9). The advantages of using artificial models is that the training set can be easily designed, and also effects of rotation and noise may easily be added. This method has been used in [54] and [158] where noise and different angles have been artificially added to the models. In the first case, the signs have been separated into 5 circular road signs and 5 triangular road signs where the results obtained are very good, with an average output value of over 95% while the other outputs are close to 1%. In the second case, 210 road signs have been used for the experimental trials, this is followed by a comparative analysis of the three different methods: NN, k-NN and bayesian classifier. There are also comparisons made between different approaches using different inputs. For each network the raw input data is tested (i.e. the image) along with the image plus additional data from a singular value decomposition (SVD) which contains invariant features of the symbols. Finally, it has been found that the best performance is that of the NN with additional SVD information where the success rate is 78%, however it misses 17% of the road signs, this value is too high for the road sign recognition system.

Speed limit road signs are classified in [50] where a discussion is presented on different NN for the development of the system: MLP, RBF, LVQ and Hopfield. There is no comparative analysis presented in this discussion as each technique achieves satisfactory results within specific fields, the choice of technique will depend on the application. A summary of these results is presented in (Fig. 2.1).

In [147] a training set has been used which is made up of binarized real images for recognition of speed limit signs (6 outputs). The validation set is not very big: only 198 images from real traffic situations under different conditions were used in experimental trials. 115 of these contained a speed limit road sign and 83 contained other road signs or no road sign. The recognition rate was 87% while the rejection rate when no speed limit road sign was present in the image was 94%, thus 90.9%

Networks comparison	Networks			
	MLP	LVQ	RBF	Hopfield
Noise Robustness	60%	50%	10%	20%
Precision	99.4%	80%	60%	40%
Memory consuming	Rather high	Low	Low	Very high
Time consuming	77 ms	125 ms	0.5 s	<1s

Table 2.1: Networks performance comparison on non-feature extracted signs pattern recognition as in [50]

of road signs were correctly classified.

In [43] a NN is proposed for each road sign category (obligation, prohibition and yield) trained with synthetic images which are subject to different transformations such as rotations and translations.

In [73] the authors have proposed a system which not only classifies road signs but also detects whether the candidate is a road sign or not. The training package consists of characteristics which are subject to several perturbations, this system is intended to be invariant to rotation, translations and scale changes. The road sign trials have been performed by separating the road signs into two groups: speed limit road signs and warning road signs. In the first case 435 images have been used and the recognition rate 98.5% while for warning road signs 312 images were used and the recognition rate was 97.2%.

Another NN configuration is the 'NN cascade', this technique performs a coarse to fine recognition which is referred to as a tree decision system where the output of the NN is the input of a subsequent NN. This allows that different features may be shared by different routes which depend on the image, thus it is assumed that the recognition is improved. Using this technique a recognition rate of over 70% has been achieved by [142] where the neural networks have been trained with artificial models in the BP mode. The road sign recognition problem is then decomposed into two different steps: the first layer classifier confirms the class which

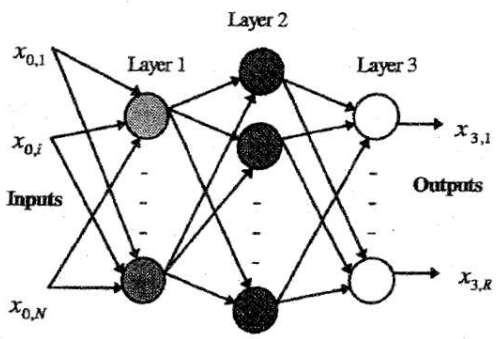
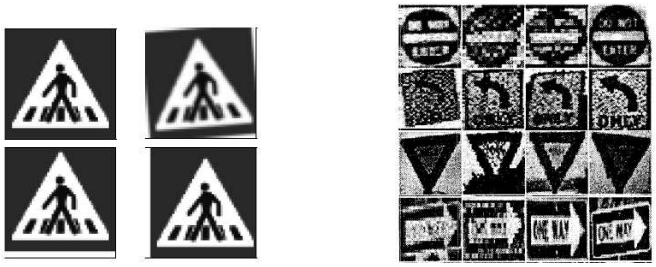


Figure 2.8: MLP scheme as in [154]. Three active layers (or 3 hidden layers, L=3)



(a) (b)

Figure 2.9: Training sets for a NN. a) Synthetic images in [43]. The original picture is transformed through artificial rotation and translations, b) samples of real images used for a RFNN in [88].

the road sign belongs to, and after, this object is sent to the second classifier which only contains objects corresponding to the road sign class recognized by the first layer classifier.

Fourier transform images have been used in [91]. Since a polar-exponential mapping has been used, the object must be located at the center of the image, to do this the object is first found and then translated to the center of the image. This ensures that the features associated with it are invariant to scales, translations and rotational changes. The method is compared to other methods such as the Shape Feature, Fourier-Mellin [46] and invariant moments, which are also assumed to be invariant to the aforementioned noise sources. The authors have proposed an experiment for recognition which depends on Gaussian noise sources

which are added to the images. The results obtained are exceptional as the system maintains a 100% success rate for Gaussian noise levels below 0.4 while the rest of the methods begin to fail at noise levels of 0.1. The decay stops at Gaussian noise = 1, where the success rate is 65%, the second best method is below 50%.

- Classifiers

– SVM: Currently, these are the most popular classifiers. Universidad de Alcalá de Henares commonly use this technique with a Gaussian kernel. Different Laplace kernel classifiers which depend on the class of road sign have been used in [120]. The input to the classifier is a feature vector constructed from the information from different moments (spatial, central, normalized) and compactness. This kernel is also used in [121] where the main goal has been to obtain a smooth function which determines the effectiveness of the classifier. Also a comparative analysis using this technique compared to other classifiers is performed (Laplace, Gauss, mixture of both Laplace and Gauss, ldc, qdc, k-NN) for different groups of road signs:

  – G1: triangular warnings. (red, white and black)
  – G2: circled. Close to all vehicles and One way. (red, white)
  – G3: circled. Prohibitions. (red, white, black)
  – G4: circled. No stopping. (red, blue)
  – G5: circled. Obligation. (blue, white)
  – G6: upside triangle. Yield. (white, red)
  – G7: octagon. Stop!. (red, white)
  – G8: diamond. Right of way. (yellow, black, white)
  – G9: square. Pedestrian crossing. (white,, black, blue)

  A summary of these are presented in the following table 2.2. However, it is not possible to select only one of these techniques among all the classifiers as the effectiveness of each one is dependent on the class of road sign. In general the Laplace classifier is the most suitable option as it obtains the best results in most of the groups.

  The group at Universidad de Alcalá de Henares [100][101][78] have also used SVM for classification, in the latter case the input to the SVM

is the absolute value of the FFT for the signature of the blobs that come from the previous stage. The blob signature is a representation in one dimension of the distance from the center of mass of the object to its edge, this is presented in function of the angle. The training set of the SVM is made from artificial models where different noise sources are applied (noise, rotations, translations and deformations). From the results obtained it is seen that this group always achieves robust classification, this may be seen in [108] where during the classification of over 4700 images only 11 false detections were found, the true positive rate was 92%.

Group	Classes	Samples	Value (%)	Methods
G1	17	1369	17.5 ± 0.4	Laplace
G2	3	720	2.4 ± 0.6	Laplace
G3	5	516	1.2 ± 0.3	Laplace
G4	2	222	0.3 ± 0.3	ldc & qdc
G5	9	627	5.1 ± 0.7	ldc
G6	2	557	0.5 ± 0.2	Gauss
G7	2	420	0.8 ± 0.5	Laplace
G8	2	216	4.0 ± 1.1	Laplace
G9	3	298	0.9 ± 0.4	ldc & knnc

Table 2.2: Mean error rates and standard deviations of mean estimates in percent

– Other classifiers: Other methods exist for the final recognition, in [131] the authors have proposed the use of different classifiers in cascade which are specialized in each step of the recognition process: The first module detects the position and direction of arrows. A second tool recognizes numbers and interprets them as reasonable speed limits. A third one is a general nearest neighbor classifier applied to three classes of ideograms (prohibition road sign ideograms, speed limits, arrows on obligation road signs). The basis of this technique is a structural analysis study as well as the learning components and nearest neighbor classifier which is trained iteratively to always choose the ideogram which cannot be identified. This learning method is very fast and requires only a small number of training patterns. The nearest neighbor classifier identifies more than 80% of all small road signs (n×n pixels

where n ranges between 24 and 35) and more than 90% for road signs of size n×n for n> 35.

Other possible methods are based on the use of a Bayesian generative modeling as has been carried out in [36]. The classification framework is based on the generative model, employing unimodal Gaussian probability densities. Prior to the probabilistic modeling, a feature transformation is performed using standard linear discriminant analysis (LDA). As a result of this procedure a 25 component feature vector of the road sign pattern comprises the first 25 most discriminative basis vectors of the LDA. Positive training samples are used per class which vary from 30 to 600. The detectors have been trained on five discrete scales, corresponding to a sign diameter of 14, 20, 28, 40, and 54 pixels and the test data set is comprised of approximately 1700 positives and 40000 negatives. The results are not related to hits but misses, where the rate of the false negative detection, the rate of false positives, the classification error rate, and the system recognition error rate are 1.4%, 0.03%, 6% and 15%, respectively.

A polynomial classifier for symbols has been presented in [98] where the objective here has been to maintain the false positive rate as low as possible, this system allows for a rather high fraction of missed relevant objects. Also binarized real images are used as the training set for the RBF for circular road sign classification and the results are compared to those obtained with a polynomial classifier. In the experiments the number of classifier classes change from 2 to 13 where the level of computation becomes more complex each time the classes increase, although better results are obtained, 5.5% missed road signs and 0.2% false positives in comparison to the RBF for the same value of false positives obtains 19% missed road signs.

In [97] a comparative analysis between the system proposed, which is a random forest (RF), with a support vector machine under different conditions along with a comparison of the bagging support vector machines with the AdaBoost naive Bayes approach has been made. The RF is an ensemble learning method that grows many identification trees. To identify an object from an input vector, the input vector is obtained from the trees in the forest. Each tree provides an identification and the forest selects the identification that has the most amount of votes. The set of grayscale images which were used for training and testing

has been 2500 with sizes of 30×30 including non-road-sign images to enhance the rejection capability of the system, the goal has been to recognize 15 different road signs. The first experiment was to compare SVM and random forest, where no noticeable difference was observed, the minimum recognition error for the SVM was 6.5% in comparison to 6.1% obtained from the RF. The use of PCA instead of directly using the information from the grayscale images provides slightly inferior results, 13.4% for SVM and 8.1% for RF, when using the HSI as the color space, 9.8% for SVM and 6.58% for RF, by applying PCS to the HSI 14.1% for SVM and 11.4% for RF. The comparison among the different methods was done for PCA and HSI color space conversion. 14.1% for SVM, 11.4% for RF, 13.8% for Bagging SVM, 21.1% for Naive Bayes and 17.2% for Adaboost Naive Bayes.

- Other methods

– HSC, was introduced in the detection subsection and was again used in [28] for road sign recognition and also it has been used for other objects on the road such as land marks. A study of the reliability depending on the distance to the road sign is carried out and the method achieves 87% for distances below 40m, but unfortunately the results for longer distances are as low as 45% for true recognitions.

– A ring partitioned method has been proposed in [146]. The core of this method is to divide the region of interest into a number of circles and to compute certain features of each one. A histogram is generated using these features which are then compared to a model-histogram. Since the histogram is circular, rotation problems are avoided. The final match will depend on a threshold which is based on the euclidean distance. Comparison among different approaches is presented using the results from the matching rate: histogram matching without partition (29.4%), fuzzy histogram matching without partition (46.7%), histogram matching with square partition (47.8%), fuzzy histogram matching with square partition (81.7%), histogram matching with ring partition (62.8%), and fuzzy histogram matching with ring partition (93.9%). The fuzzy histogram technique provides smoothened histograms. By smoothing the histogram imprecisions which result from the nature of the image or the digital image processing can be reduced.

As a result, in all cases the use of a fuzzy histograms provides better results. A similar algorithm has been used in [49] where histograms from a tensor of features have been compared (Fig. 2.10). The author has divided the problem into recognition of prohibition road signs and obligation road signs where the distance given by a Chi-square distribution is used for the final matching. The recognition rate is 88% for prohibition and 91% for obligation. Unfortunately the road signs must be bigger than 100 pixels which does not allow long distance recognition.

2.5 Tracking

In early research, tracking of road signs was carried out to assure the recognition process. Currently, the tracking concept has changed and now it is being increasingly associated with statistical methods instead of the past approach for temporal integration as there was no explicit tracking of objects. Later with the introduction of statistical approaches, based generally on the Kalman Filter, it was already possible to track targets to guarantee the recognition process and also to limit the search region. More recent tendencies have begun to use CONDENSATION or particle filters for multi object tracking.

– In [111] a model is proposed which only depends on the lapse between frames and considers that the vehicle is traveling at a constant speed. Unfortunately the results obtained are not effective, as only a few number of parameters are used to solve the problem. The successful tracking rate is as low as 23.5%.

  Temporal integration has been used in [69] where a minimum number of positive recognitions are required to verify the correct road sign. A similar proposal has been presented by the authors of [44] where two positive detections are required to consider a true recognition and to remove previous detected road signs from subsequent frames when two missed detections occur. In addition the bounding boxes of subsequent detections which overlap are also taken into account. The importance of improving the image stabilization to improve results in the case of road surface irregularities is also mentioned here. In [144] the authors have proposed a tracking method designed to minimize both the computational effort and tracking errors. Potential road signs are sent to a

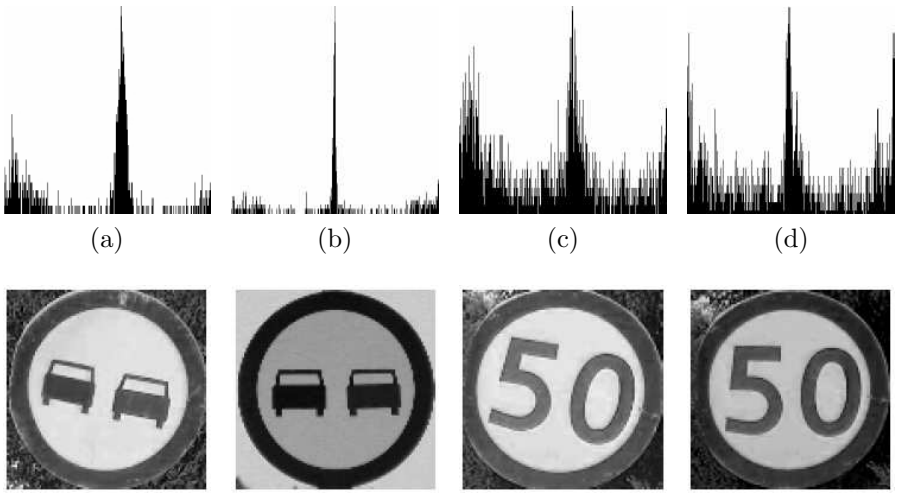


Figure 2.10: Histograms corresponding to each road sign as in [49]. The low variation of the histograms under rotations may be appreciated

module which then searches in the subsequent frames for the center of the road sign and symmetries in the neighboring region for radii bigger than those previously detected. To improve the results, three different radii are used (e.g. 10, 15, 20 pixels) instead of only a single model size.

– The general approach used is the Kalman filter. In [26] have introduced the Kalman filter which describes a model for the camera motion. The system has been tested in a circuit 200 m long and with a vehicle speed of 5km/h. It has been observed that it is the possible to predict the 3 coordinates of the road sign where an ellipsode uncertainty of 0.5 m after the 12th frame has been achieved.

  The Kalman-Bucy filtering technique has been proposed in [128], and a more detailed description can be found in [64]. It leads to the same set of equations as its probabilistic analogy, however this does not require arbitrary assumptions on the errors of the probability density functions and they can be taken as maximum uncertainties. Finally, it has been found that it is not necessary to address the recognition problem using Kalman methods, as a simple time weighted average of the road sign

image cross-correlation with the signs in the database provide very satisfactory results.

Constant speed and the size of the continuous detections have been used in [63] to track road sign candidates until they are recognized. After, prediction of both the position and size of the next detection is developed where it has been shown that this reduces the search time and space required for location of the same road sign in subsequent frames.

The authors of [73] have implemented a Kalman filter for circular and triangular road signs which use three parameters (x,y,r): for circular signs the center of the detected circumference (x,y), and for triangular signs the center of the circumscribed circumference (x,y), the component r is the radius of the circumference. N quantitative results are given in a series of images from the tracking procedure.

Two final examples are presented in [102] and [101] where a filter based on the features of the road signs assumed to be constant under certain conditions are presented. A special Kalman Filter is also added, $\alpha\beta\gamma$ filter, the results for different weather conditions are presented in Figure 2.11. It has been considered that the relative movement of the sign with respect to the driver is linear except in cases of curved road sections and for lateral displacements of the vehicle. This technique also provides results for constant acceleration and constant velocity. One of the main features of the model is that important data is kept as a signature and contains the positional coordinates, size, color, category type and the mean gray level of the region occupied by the object. This information is stored for each road sign that is tracked and when considering the complete set of information it is not possible to assign just one object to a tracking process, the euclidean distance between new objects and predicted positions for all tracked road signs is used to obtain the correct road sign. Using this tracking system the majority of false positives have been eliminated, this has been the case for 1 out of 18 in one experimental sequence and 0 out of 7 in a second.

– Later experimental work in this area has relied on different techniques, in [117] the authors have introduced a combination of tracking and detection using a particle filter. The filter must track the 3D object in the sequence and detect the model that is the most appropriate from

(a) (b) (c)

Figure 2.11: Road sign tracking under different lighting conditions using the Kalman filter as in [73]. a) Road sign tracking with rain drops on the car windscreen, b) Left bend warning sign and speed limit sign on a cloudy day, c) Yield sign on a sunny day.

the raw images. Generally this problem is solved separately, but in this particular case a method which allows joint detection and object tracking has been developed. A hybrid system is used which models the evolution of the shape hypothesis. This is followed by embedding the detection in the estimation procedure which is solved by the particle filtering. 5000 particles per hypothesis have been used and each iteration takes 0.8s. These application is able to recognize and track the road sign given 4 possibilities in no more than 3 frames.

In [157] the authors have used the CONDENSATION technique, where a series of low resolution images which contain road signs extracted using CONDENSATION from successive frames are stored. First, the regions for extraction which are assumed to contain a road sign are developed. Using this technique it is possible to detect road signs with sizes of 14×14 pixels, this corresponds to a distance of camera - target of 125 m. In subsequent frames the CONDENSATION method chooses different candidates before the tracking process begins. The fitness is calculated using an evaluation function that is based on the shape of the road sign and its characteristic color.

2.6 Other Important Issues

- Real Time: From the beginning of investigative work on an on-board vehicle recognition system, the requirements have always included a real time approach. There have been great advances in the computational time, from the 2.5 min recognition time performed by a system using a

486/25MHz PC in 1994 [94] to the detection and recognition using a NN system, the system proposed in [43], which was previously explained, is able to run faster than ten times per second using a Pentium 4 at 3 GHz where this is more than sufficient for the real time requirement.

There are a variety of other methods designed which have been implemented to face this problem: several libraries have been specifically designed as well as video computation [28] to reduce the execution time to between 1 and 2 seconds for detection, also a parallel configuration has been suggested to reduce the time even further. However, due to current computational power those times are now regarded as being obsolete.

It is important to emphasize that a good color space conversion provides better detection performance, this generally means less recognition time, some good examples of robust systems that consume less time due to their special features are the following:

Detection, recognition and tracking has been developed in [101] (explained before) at 3.26 fps with a 2.2 GHz Pentium 4-M using images of 400×300 pixels.

10 or more fps are achieved in [49] where a system has been developed for detection and recognition of road signs of more than 100×100 pixels in images of 320×240pixels. This system has been described in previous subsections.

In [105] the authors have achieved 20Hz in detection, time has been saved here by using an algorithm to finding regular polygons via the fast radial symmetry transform. The full classification was implemented in C++ to evaluate real-time performance and, for a 320×240 image, the radial symmetry detection and classification was capable of running at 20Hz, with classification times of only 1 ms.

More efficient structures have also been used in [60] where three processed images are obtained per second using a compact processing structure called raster scan video processing. This processing structure constructs incrementally and auto associatively the relations which allow the different objects to be segmented via relational properties.

In [147] the system only recognizes speed limit road signs and achieves an average image processing time of 130ms, this corresponds to 8 images per second. This has been measured on a PC with a 1.4GHz

Athlon AMD processor (512Mbyte RAM). It is stated here that the image filtering is the most demanding computational process where the average processing time is 90ms, while template matching and the number recognition both take close to 40ms.

The use of look up tables (LUT) has also been employed to reduce the execution time, it allows data to be stored in a vector before the application is run, this avoids the requirement of later computations. It is mainly used in color space conversions as these typically involve time consuming calculations. Some examples of pixel classification may be found in [69][52] and [108] or for the calculation of the activation function in a NN in [58].

Using a parallel search [23] speeds of 7.5 fps have been observed in detection and for 5 different road signs: yield, priority road, stop, no entry and speed limitation.

Another method to reduce the computational time is by using more powerful computers. This has been the case for the system presented in [22] where a multicore processor has been used to achieve real time processing, also a scan of only 120×120 pixels is performed in a region and threads are used to search for different models. Using this technique, the authors have demonstrated computational times of 25 fps which is more than satisfactory for real time systems. Vitabile has focused on hardware solutions to provide a real time system. In [151] the authors have proposed a change in the code from C to Handel C to improve the algorithms developed in [152] and [153]. The design of the implementation has been carried out on a Celoxica RC100 board, incorporating a Xilinx® Spartan II 200K gate FPGA, times close to 10 ms have been obtained for pixel color enhancement, image processing operations and frame acquisition time. Also Vitabile [150] have implemented a NN in a FPGA in which the sinusoidal activation function decreases hardware resources, recognition process times of 0.028 ms have been obtained. The implementation is based on virtual neurons which makes the mapping of a neural network efficient for the hardware devices thus decreasing the concurrent memory access.

- Geometric Restrictions: Another problem that authors have tried to solve is that of geometric restrictions (Fig. 2.12). The basis behind this development has been to accelerate the searching process, this is

possible as the complete image is not used, this technique also avoids false detections from parts of the image where there is no possibility of detecting a road sign. Tests have also been done on defining smaller regions to look for road signs. Nevertheless this approach is not very realistic and needs extra data or sensors to provide a robust limitation of the search region, this is mainly due to the flat-world approach and the vibration of the camera or its casing.

Examples of this type of system are presented in [27], this system has been designed for Japanese roads and it only searches for road signs on the left side of the image, once the sign is detected the region changes slightly to a more narrower area on the top left of the screen. A similar procedure has been developed in [22] but in this case the search is performed within a reduced region on the right hand side of the image, i.e. the region where the highest probability of locating a road sign is.

In [127][128] the aim has been to avoid false positives and improve the system execution time: the vanishing point is used to limit the search region. It is assumed that the triangle formed from the two bottom vertices and the vanishing point do not contain any road sign, but as long as it is not possible to define a triangular box the final search-in region is a rectangle, which is smaller than the complete image.

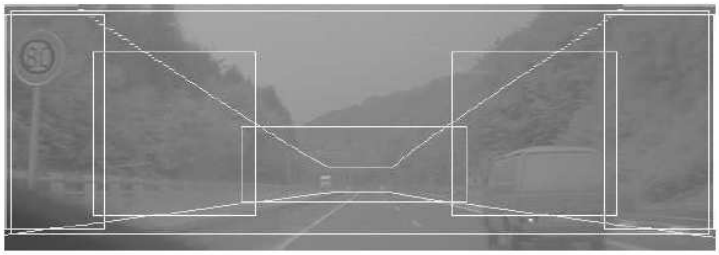


Figure 2.12: Geometric restrictions as in [113]. ROIs for the signs that are located to the left and right hand side of the road.

- Applications: recognition systems are currently being used for many applications. One of the most common applications is its use for road sign inventory. This idea arose in the 90's [94] where the convenience of this type of system to automatically store road signs has been considered. Also [26] have presented a proposal as part of the AUTOCAT project framework for off-line recognition and storage of road signs.

Other applications are that of an inertial navigation system for positioning of road signs. In [139] a system with a camera and a PDA allows an operator to insert road signs in a database by pointing directly at the road sign. The inventory application is explicit in [108] where the state of the road signs and their storage are the main goals.

In [125] the authors have proposed a road sign inventory system to perform a comparison of new data with previously stored information. Using this system a more robust recognition has been achieved.

The system proposed in this thesis may also be used for road sign inventory purposes, where it has been used as a basic platform to develop the DIASTIC project for automatic road sign inventory systems designed by GEOCISA S.A. and the Intelligent Systems Lab. Its inherent features include good recognition rate, GPS information, off-line image processing and positioning of road signs.

2.7 Other Similar Purpose Systems

In [38] the authors have proposed a system to reduce the number of traffic accidents. A signal is transmitted from the road sign to the car, and the receiving antennas determines the relevance of the incoming signal. If the information is important, the signal is broadcast to the driver, secondly the driver is notified and prompted to perform specified driving maneuvers, and finally the driver is prompted and forced to change the driving maneuver. The disadvantage of this system is that, as is stated, incorrect vehicles may receive information as the signal is sent not only to the interested vehicles but to their proximity. Another disadvantage to this system is the expense of the infrastructure.

Several companies are currently developing commercial road sign recognition devices. In recent years, this announcement has been continuous from different vehicle manufacturers however currently almost no devices have been made available, making it impossible for the author to test and compare such systems. For example, in the 3rd International Conference on Semantic and Digital Media Technologies (SAMT2008) in Koblenz, Germany, Daimler made a presentation of their work on a system which is not yet on the market, and which has also appeared in

an article in the Neue-Ulmer Zeitung [17] where the system is described. The camera acquires images every 30ms and is able to recognize speed limits and No Overtaking road signs, currently laboratory based systems also achieve recognition of indication and end of ban road signs, all of which are related to speed information.

Continental (Fig. 2.13) proposes a road sign recognition system (Speed Limit Monitoring, SLM [16]) which informs the driver of the appropriate speed limit by providing this information on screen, the system use a multipurpose camera for both a lane departure system and the road sign recognition system.

Figure 2.13: Speed limit monitoring system of Continental [16] and two examples of right recognition, speed limit in the center and end of ban on the right

Mobileye in collaboration with Volvo are currently developing a system to recognize various road signs under good illumination conditions. This system also uses additional information provided by the navigation system which is in turn broadcast to the driver [6]. Hella (see Figure 2.14) is working with Opel on a system to recognize speed limit and no-overtaking road signs which will also integrate the information from other devices to obtain a more comprehensive safety system. Opel has already put on the market, in their model Opel Insignia [19], a camera named "Opel eye" [18] which operates at 30 fps and is used as a road sign recognition system. It is able to recognize speed limit road signs and prohibition of overtaking, as well as their end of ban road signs at a distance of 100m, all of these recognitions are performed when good illumination conditions are present.

Figure 2.14: Image of the proposal of Hella in Opel vehicles.

2.8 Summary

- The large variety of methods and algorithms used in detection, recognition and tracking of road signs have been presented.
- Also shown has been a summary of the main techniques used for each stage where it has been highlighted that a number of them may be used for different purposes, e.g. NN and Template Matching for detection and recognition.
- It has also been highlighted that there is currently no well established agreement on the use of color or grayscale images, even in the case of color there is no unanimity in the discussion of the best choice for the color space, this is mainly due to the difficulties in finding a color space which provides constant results for different illumination conditions.
- The most relevant and commonly used algorithms for preprocessing of images and detection have been presented: color conversions, noise removal, feature enhancement and shape classification.
- The main recognition methods have been shown: NN, pattern and template matching and more recently classifiers are the most widespread

tools used to develop this stage.

- Tracking or temporal integration allows the recognition to be ensured and provides reduced computational time for both the detection and recognition processes. Features based on techniques from early works, and probabilistic models such as the Kalman Filter or CONDENSATION and Particle Filter are currently the most commonly used tools.

Chapter 3

Detection

Two different objectives for the detection process of road signs have been presented in the state of the art of this thesis. The first of these is to extract all possible road signs from the image regardless of the group they belong to. The second obtains objects which are most likely to be road sign candidates and these objects undergo a subsequent classification stage. The work presented in this thesis is based on the latter objective. Once the objects have been extracted from the image they are then processed by the recognition module that, among other tasks, filters false detections (objects miss-classified as road signs, in technical terms this is referred to as a false positive). Detection will be carried out using the template matching technique and the normalized correlation function which determines the similarity between the image and one of the models.

The overall principle behind the detection task is to avoid missing any road sign detections that are contained in the image, this is carried out knowing that there is a high possibility of introducing a greater number of false detections. The system has been designed to detect road signs which have red rims and those which have a blue background, the road sign classes contained within the system database caters for all different classes of road signs: Prohibition, Danger, Yield, Indication, and Obligation. The scope of this thesis covers all daytime illumination conditions where the detection must take place regardless of the ambient lighting.

It has been stated in the introduction that the detection system is required to work in both urban and non-urban environments. The road signs located on highways are generally bigger, are not subject to occlusion, have only slight rotations, and are constantly inspected to maintain high levels of

signaling quality. However this is not the case for road signs located within the urban environment as they are subject to all types of damage (aging, vandalism), these signs have smaller dimensions, have a higher probability of being subject to rotations and skewing as there are no fixed rules for their placement within the city limits.

In the state of the art the different problems associated with detection have been discussed where this has included occlusions, in-plane and out-of-plane rotations, and above all changes in ambient lighting. This latter problem is of great importance as it impedes the perception of color, apart from the shape and information contained within a road sign the color is one of its characteristic parameters. Considering this fact, a lot of emphasis has been placed on the correction of problems associated with road sign lighting, further studies have also been carried out on the detection of road signs which are located beneath a shadow and also road signs with ambient back light, i.e. the road sign is located between the vehicle and a light source, typically the sun. It has been observed that changes in illumination is one of the main factors which causes false negatives (the road sign is in the image but has not been detected), the study of this parameter is vital for efficient functioning of the road sign detection system.

The framework which has been developed for this thesis has involved the following set of tasks:

- Image preprocessing: this task is necessary to establish an image processing algorithm which enhances the pixels which belong to road signs while at the same time removing the rest of the pixels. This first task decreases the probability of a false positive occurring during the detection process as objects which are similar to those of road signs are removed. Considering this, use will be made of different color spaces, RGB, HSL and HCL, which are explained in detail in [51][4][3], and other normalized color spaces which are best suited to resolve the problem.

- Once the color filter process has been applied to the image, a method must be devised which efficiently and accurately detects possible road sign candidates. To do this, each area of the image which may contain a road sign is assigned a probability factor which reflects the similarities between the possible detected road sign and the typical shape of the road signs which are to be detected. This will be performed using the normalized correlation function, here a number of models are created

which represent the main shape characteristics of the road signs which are to be detected and take into account the image obtained from the preprocessing stage.

Considering the method used, detection success depends on both the preprocessing of the image where the template is matched and the different models used. Since they are closely related, the compromise between these two factors is very important. The image will determine the type of model, but also the type of model will determine the type of image to be obtained. The limitations and the features of each of these requires a compromise to be made for detection reliability.

This chapter is organized as follows.

- A study of different color spaces is presented, the main task here is to obtain the best enhancement of the colors of interest which belong to the road signs, where these colors are red and blue.
- Color normalization algorithms are presented as a possible approach to image preprocessing, also possible applications using these algorithms will be included.
- After defining the color space and model, a detailed study of the components of the color space will be carried out to obtain a clear understanding of the information provided from each one, from this analysis the possibility of assigning one general solution to the two different colors will be explored.
- Once the color information has been extracted from the image, the next step is to provide the base for the normalized correlation method, this is carried out by selecting adequate models for the correlation process. Here the different models proposed and results using each of them will be presented.
- When all the necessary information has been obtained, it is then possible to perform experimental procedures that optimize the results obtained from the detection process, this is part of the application which has been installed in the IVVI platform. For further information about IVVI see Appendix D and reference [5].
- A discussion on the advantages and disadvantages of the method which has been proposed is also provided.

- Finally, this chapter concludes with a brief summary of the most important aspects which have been developed throughout the chapter.

3.1 Image Preprocessing

3.1.1 Road Sign Color Enhancement. The Component Model

The image on which the template matching is performed must be processed adequately so that the result of the correlation is optimized. The main aim of this task is to obtain a grayscale image which highlights the most significant color components of the road signs amongst the rest of the image colors. These are red for prohibition, danger and yield road signs, and blue for indication and obligation road signs. Part of this task has involved experimental trials which have been conducted on different color spaces:

- To compare the two different approaches related to color enhancement: the first is based on the lighting dependent color space, RGB, and the second is not as dependent of the lighting, HSL and HCL. All of these approaches have been evaluated so that the most adequate is implemented in the system.

- To find the optimum color space which allows the colors from the road signs to be distinguished from the rest of the colors in the image.

- To minimize the failure of road sign detection, where this process is independent of the illumination in daylight conditions.

The study of the color properties and their results has been performed using a database of road sign images which has been obtained from the IVVI platform. The images are recorded for different daytime lighting conditions and for different seasons. Also particular emphasis has been placed on obtaining images from urban environments, as in these areas the largest changes in color occur. With respect to signs located on highways, these are bigger and well preserved, also there are no occlusions or rotations as the road signs are positioned according to the Spanish standard of the *Norma 8.11C de señalización vertical* [110], in this document the location rules for road signs may be found along with the dimensions and main road sign characteristics.

In urban environments, there are no strict placement rules for road signs. Contrary to highway road signs, urban road signs are generally subject to damage caused from vandalism and aging.

The main proposal here is to obtain as much information as possible on the range of variations of the red and blue colors required for the detection process. Considering this task, three groups of road sign images have been constructed where they depend on the daytime lighting conditions that the road signs are generally subject to, these are: sunlight, shadowing and half-light. The latter of these also includes doubtful cases which involve sunlight, shadowing and back light conditions. The main sources of false positives for both red and blue road signs has also been incorporated into this study. A common source of false positives is due to red bricks where these demonstrate a similar coloring to the rim of red road signs, as a consequence windows surrounded by red bricks may be falsely detected due to their white background. In the case of blue road signs, the main false positive sources are due to the blue sky and pavement.

In the sky and asphalt case and as a result of their location in the image, it is commonly considered that they may be avoided using geometrical restrictions, thus decreasing the number of false positives. However, these simple algorithms may not be applied in urban environments as there is a wide range of possibilities for the location of road signs. Examples of this situation are presented in Fig. 3.1. Also, other factors exist which make it impossible to solve this particular situation without the use of special algorithms or electronic sensors, the slope of the road may easily cause the road sign to be missed during the detection process.

One of the main tasks here is to include the typical colors which correspond to bricks, pavement and the sky and to verify the possibility of filtering these colors by taking into account the values of their components in the different color spaces. A study is performed on the different color spaces where an attempt to isolate the colors which represent road signs from those which represent other objects located in the image is made. A component model is obtained for each of the objects under investigation which when applied to the image will enhance the particular color features. This experiment, which is outlined here, is based on a number of tasks which are described as follows:

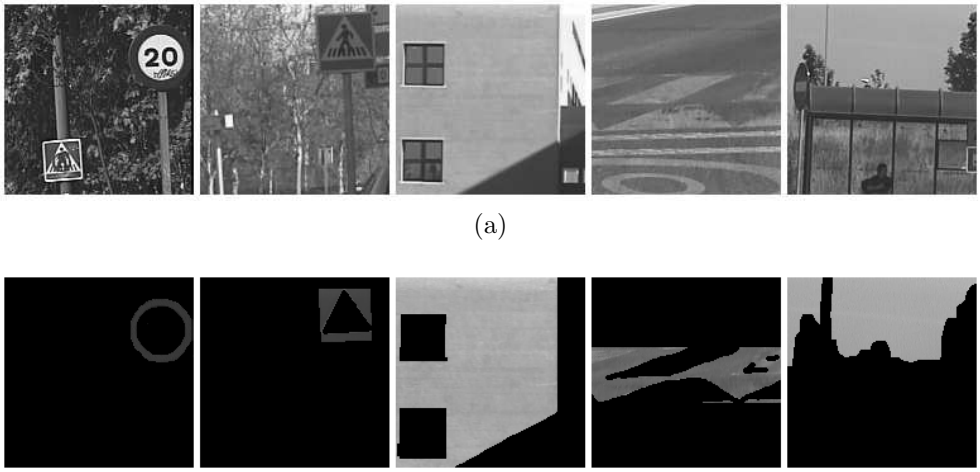
– Natural road sign images are divided into different groups as may be seen in Figure 3.2a depending on the lighting conditions or the type of false positive it represents: sunlight, shadowing and half-light for

(a) (b)

Figure 3.1: Samples of original images which demonstrate the difficulties in applying geometrical restrictions to avoid false detections. In the first case the road sign is seen to be surrounded by the sky. In the second case the road sign is located in the lower portion of the image.

the lighting conditions, bricks, sky and pavement as the most common sources for miss-classification of pixels. This is followed by a manual setting to reduce the levels of pixels in the image which do not belong to road signs, see Figure 3.2b. The main objective of this procedure is to obtain the pixels that only belong to the road sign and which are required for further processing.

– The histogram for each of the road signs is computed over the appropriate color space and corresponding color. Each component is represented and normalized to 255 except for the L component as its value varies over the complete grayscale range without a pattern, thus the information provided is not useful for the detection stage. The histogram representation (Fig. 3.3a) provides the probability that a pixel belongs to a specific component hue, saturation or chroma value of the road sign color. As every possible color belonging to the road sign must have the same weight, all are normalized to their single maximum and represented together, the complete set of histograms represents each one of the samples used in this study, as a result, it is possible to identify the variation range of the spectra which corresponds to each color for each color space.



(a)

(b)

Figure 3.2: Samples from different sets which describe the specimens under study from left to right: red road sign, blue road sign, bricks, pavement, sky. Row a) Original cropped images. b) Assessment of samples from row a) background to low level

– By using the data from the complete set of histograms, a final histogram is generated where each gray level corresponds to the maximum value from each histogram, this procedure is demonstrated in the following Figure 3.3b. Using this new histogram the appropriate profile for the color variation is obtained for each representative color (red or blue) of the road sign, and for each of the components of the chosen color space (RGB, HSL, HCL). It is important to point out that this procedure is not intended to obtain information on the most likely profile, but to obtain the full range of variation for each color regardless of the lighting. This procedure assures that the likelihood of any color that belongs to the road which has not been enriched is low.

– The representation of the histogram of maximums becomes a model for the computation of the LUT for each part of the color space chosen. This representation has been constructed using the procedure which is described as follows: a straight line is used to join the first maximum to the last maximum, always when these maximums belong to the same range. This same process is carried out on the minimums. For the

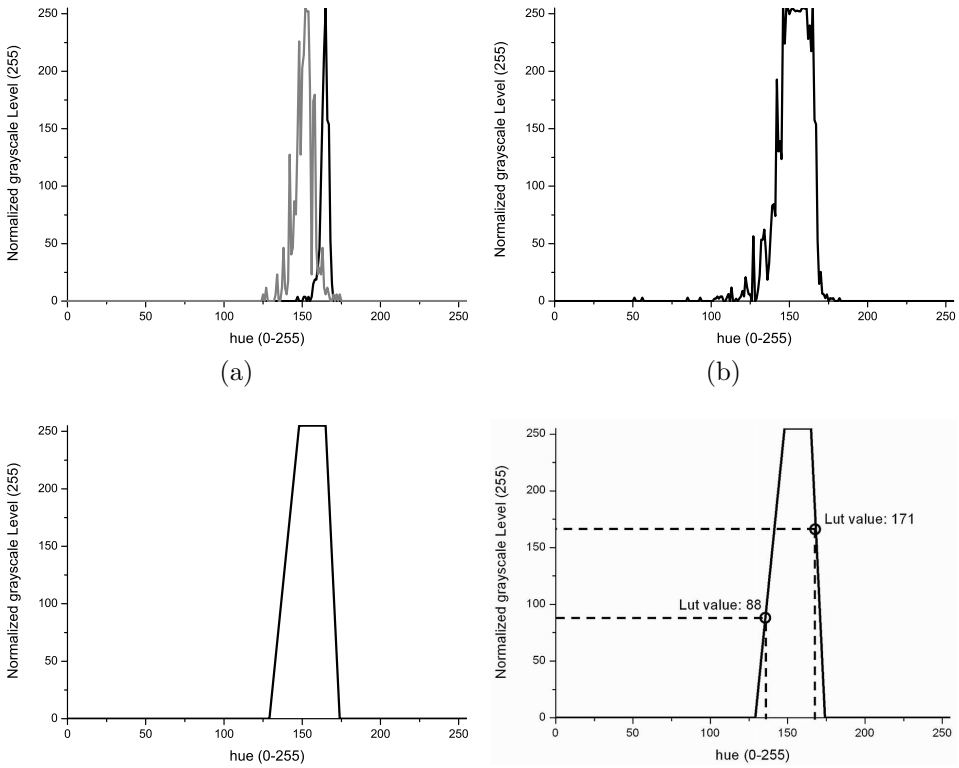


Figure 3.3: Example of the process used to obtain a color model which is used to fill a LUT in the HSL space, for the H component of the blue road signs. a) histograms corresponding to two samples of rim road signs. The graph contains only two histograms for simplicity. b) histogram of maximums which represents the variation range of the component values. c) Linearized model of the histogram in b). d) Representation of a LUT value assignation following the model in c).

maximum to minimum (and *vice versa*) the slope is taken from the first minimum to the first maximum, and from the last maximum to the last minimum, this process is depicted in the following Figure 3.3c

– Finally, the model is applied for each R, G, and B value thus obtaining the corresponding value for the component in the chosen color space. The final value of the grayscale level in the enhanced image is the one provided by the height of the curve at that point as depicted in Figure 3.3d. The LUT is calculated off-line to avoid unnecessary time consumption while the application is running, otherwise the conversion could not be done in real time using the system developed for this research.

3.1.2 RGB, HSL and HCL Experiments

Experimental work has been performed considering RGB, HSL and HCL color spaces and red and blue road signs. Part of the this work has involved a comparison of the results obtained with a color space that is light dependent(RGB) to the results obtained when using color spaces that have less light dependence, both the HSL and HCL. For each component under investigation a grayscale image is obtained from the results of the specific conversion.

Enhancement of the R and B components in the RGB color space is obtained by applying the commonly used formula:

$$GSV_{red} = 2R - G - B$$

and

$$GSV_{blue} = 2B - R - G$$

where GSV is the grayscale value and the subindex indicates the RGB component.

As each single R, G, B component of the RGB color space is composed of information on the color properties such as hue, lightness and saturation, it is not possible to identify each specific property. This is the reason why, in the RGB case, just one data curve is presented which represents all possible objects of interest such as road signs under different lighting conditions (sunlight, shadows, half-light) and also the main source of false positives (bricks, sky, pavement).

With respect to the HSL and HCL case, as their different components represent different color properties, it is possible to separate the results and obtain more detailed information on each component. The experiment has been developed to account for the components due to hue (H), saturation (S) and chroma (C). Also studied here has been the joint information from H and S for HSL and H and C for HCL. The results may be summarized as follows:

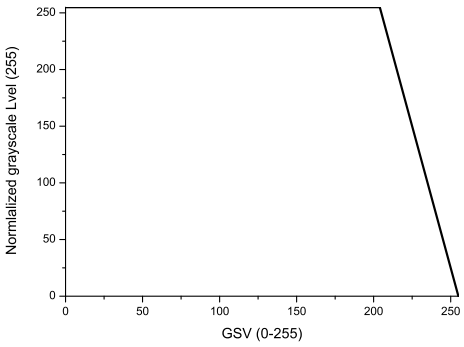
Red Rim

- RGB red enhancement (Fig. 3.4). Logically, in this case and considering the enhancement formula, the peaks due to the red road signs will cover the complete range of possible values. The pixels cannot be grouped as no specific range of values can be limited by the object of interest (road sign or typical false positives) or the different light conditions as was expected. Two images are shown which present this curve generation process: Figure 3.4b and 3.4c it is seen that every reddish pixel is enhanced and as a result noise is introduced, this increases the complexity of the detection process.

- HSL color space, component H (Fig. 3.5a). It can be appreciated that for high hue values there is an overlap among the three curves which represent the different lighting conditions. For sunlight conditions the main peak is located at the high H values while, for shadows and half-light the peak is seen to shift to the blue zone.

  Fortunately a large amount of information from bricks does not overlap the road sign area for low values of hue, while there is a considerable amount of extra information at high H values, where this coincides with information from half-light conditions and as a result it is not possible to completely separate brick information from the road sign information.

- HSL color space, component S (Fig. 3.5b). Contrary to the H case, S from the different objects under study overlap covering almost the complete spectra. It may be seen that only a small amount of the information from bricks is outside the range of the road sign information, this result shows that the saturation may not be considered as a fundamental component for the brick color filtration process.

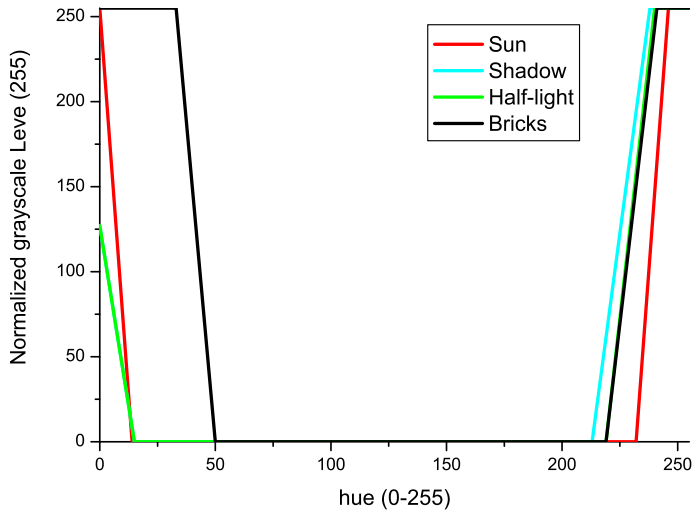


(a)

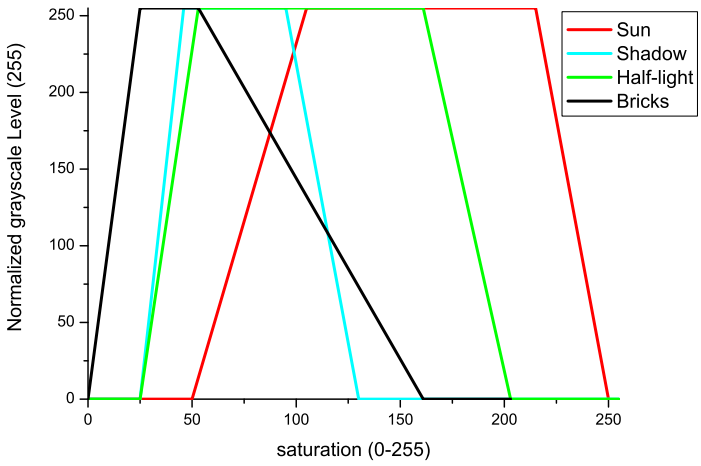
(b)

(c)

Figure 3.4: RGB enhancement through the conversion $GSV_{red} = 2R - G - B$. For a better understanding, an example of the red enhancement of b) is in c.



(a)



(b)

Figure 3.5: Red models of LUT for H and S components in the HSL color space for the cases under investigation and depicted in the label of the graph.

- HSL color space, components HS. (Fig. 3.6). The surfaces which belong to sunlight, half-light and shadows partially overlap for high hue values and completely overlap for low hue values, half-light and shadow cases. It may also be seen that the surface which belongs to the bricks almost covers the data from the road signs for low hue values, and only a small amount of information is obtained for high hue values.

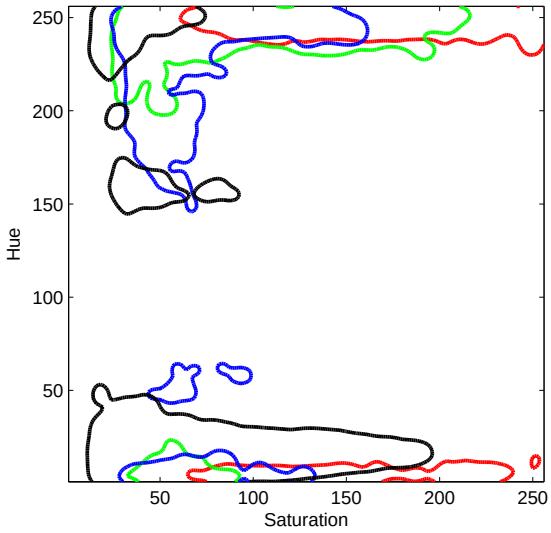


Figure 3.6: HS components in HSL space for red road signs. Each surface represents the information of H and S together for each of the objects under investigation. Red: sunlight; green: half-light; blue: shadow; black: bricks.

- HCL color space, component H (Fig. 3.7a). In this case, the peak due to sunlight shifts to the orange region with respect to the HSL case. The data obtained from bricks and the light conditions completely overlap. Unfortunately this occurs at high H values, the half-light case covers the information from bricks and as a result it is not possible to disregard the signals obtained from bricks without loosing information from other components.

- HCL color space, component C (Fig. 3.7b). For shadows the chroma is very low and confined to a small region, but the information form the half-light case overlaps the information from shadows and bricks

and also part of the sunlight. As a result it is only possible to filter information with low chroma values.

- HC component, HCL color space (Fig. 3.8). In this case there is also an overlap between the influencing areas which correspond to sunlight, shadows, half-light and bricks. For high hue values only a small part of the road sign information that is under shadow and half-light will overlap. Meanwhile, for low hue values the red brick region covers a larger part of the information from both the half-light and shadows, thus in this case if we consider the brick region for enhancement the number of false positives will increase and in the case that this region is not taken into consideration some of the road signs will not be detected.

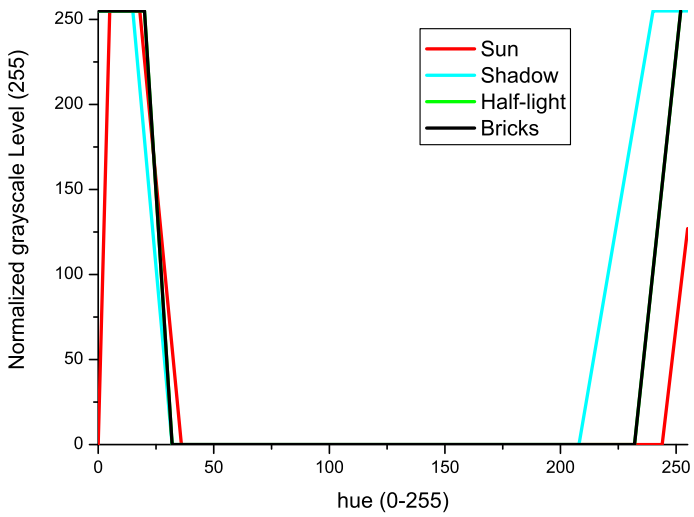
Blue Plate

- RGB blue enhancement (Fig. 3.9). It may be seen from the blue enhancement case that the grayscale values are distributed all over the axis, this means that no grouping can be expected and that only post-processing of the image is carried out after enhancement.

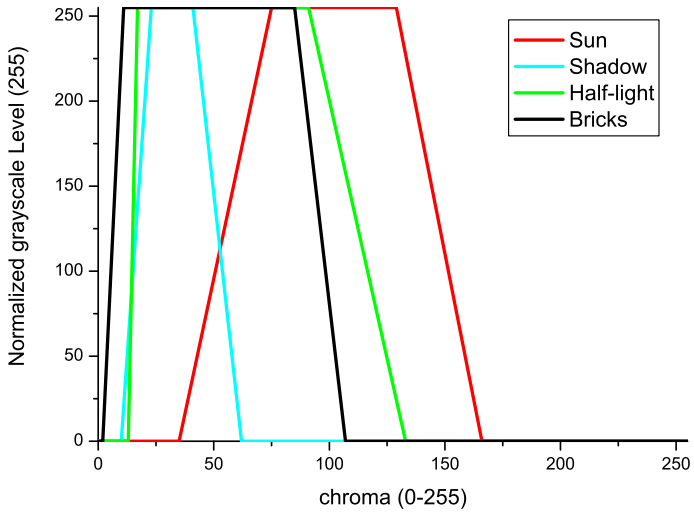
- HSL color space, component H (Fig. 3.10a). For the three lighting conditions the information is found over the same range, although this range increases for the half-light case, no differentiation may be established from this data. Also the sky may have exactly the same color as the road sign when it is not fully saturated. The bluish pavement, due mainly to crosswalks, produces the majority of the information around the zone of the road sign, however for H values over 175 it is possible to filter part of the data.

- HSL color space, component S (Fig. 3.10b). Though the pavement and sky typically have low saturation values, they are not low enough to be considered as a parameter for filtering. This is because part of the half-light case falls within the same zone. Shadows are located close to the center of the spectra and are almost covered by the sunlight curve.

- HSL color space, components HS (Fig. 3.11). It may be observed that the complete information from road signs is almost completely covered by the information from false positive sources. Only a small portion of the pavement and sky surface is isolated. Fortunately this does not



(a)



(b)

Figure 3.7: Red models of LUT for H and C components in the HCL color space for the cases under study depicted in the label of the graph.

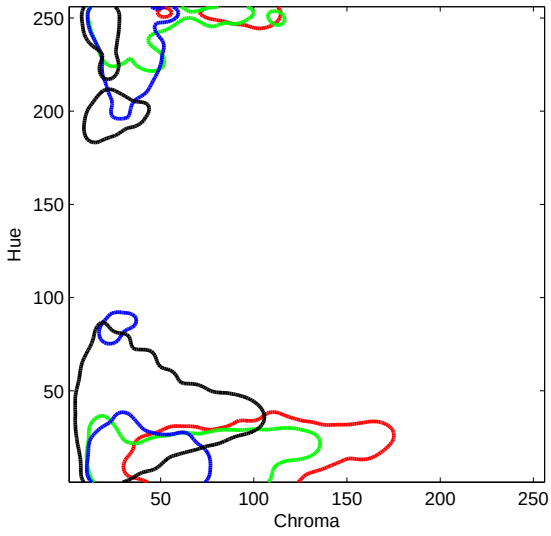
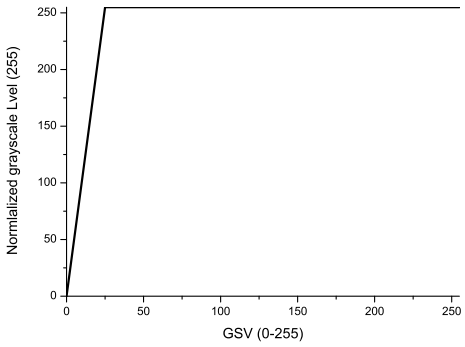
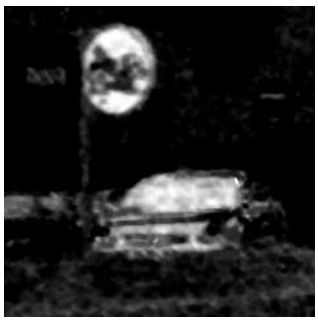


Figure 3.8: HC components in HCL space for red road signs. Each surface represents the combined information from H and C for each of the objects under investigation. Red: sunlight; green: half-light; blue: shadow; black: bricks.



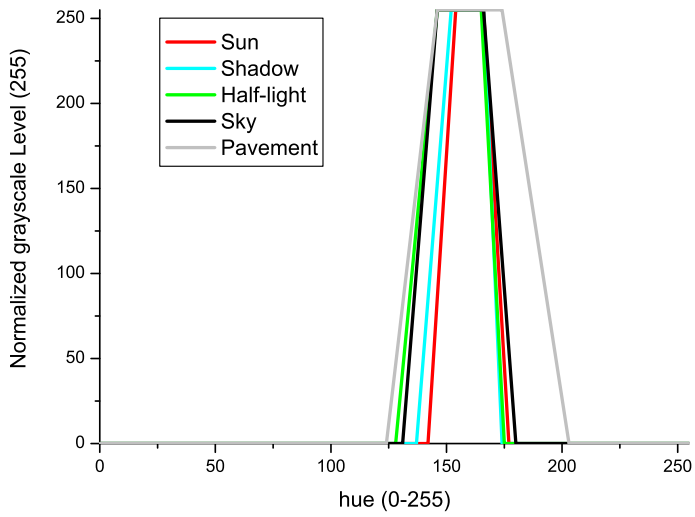
(a)

(b)

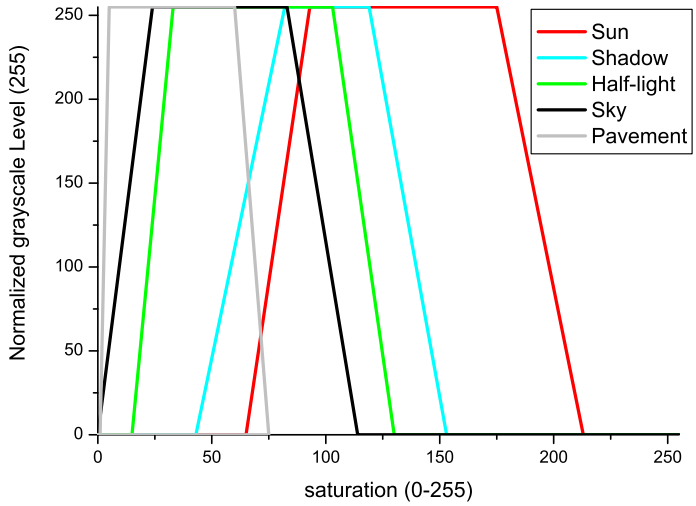


(c)

Figure 3.9: RGB enhancement through the conversion $GSV_{blue} = 2B - R - G$. For better understanding, an example of the blue enhancement of b) is in c.



(a)



(b)

Figure 3.10: Blue models of LUT for components H and S in the HSL color space for the cases under study depicted in the label of the graph.

mean that the colors from the sky and asphalt are always enhanced, in fact the sky and pavement are almost always on the grayline in the white zone and in the gray, respectively.

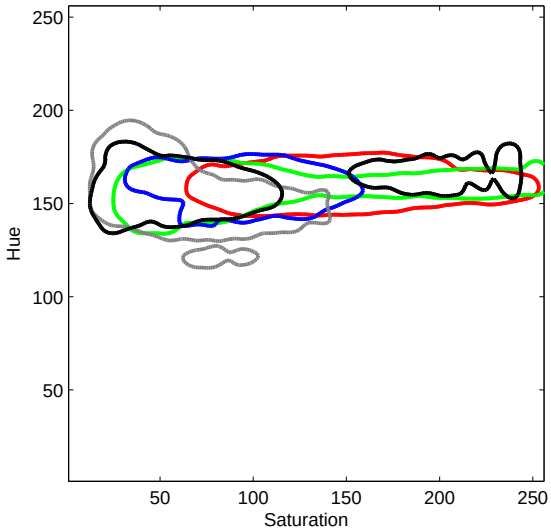
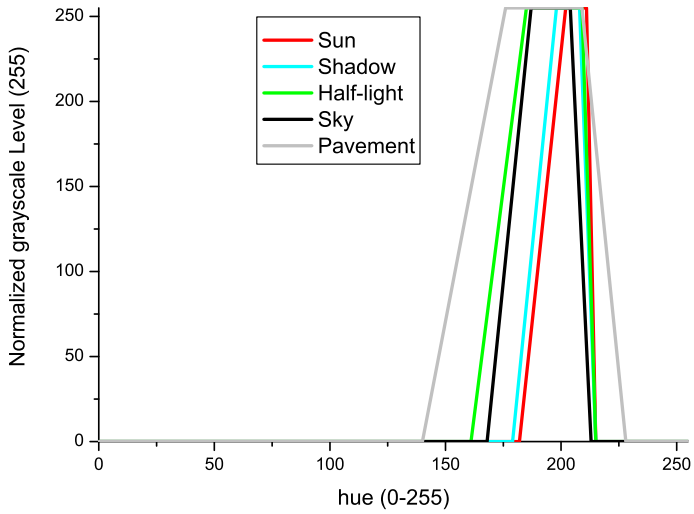
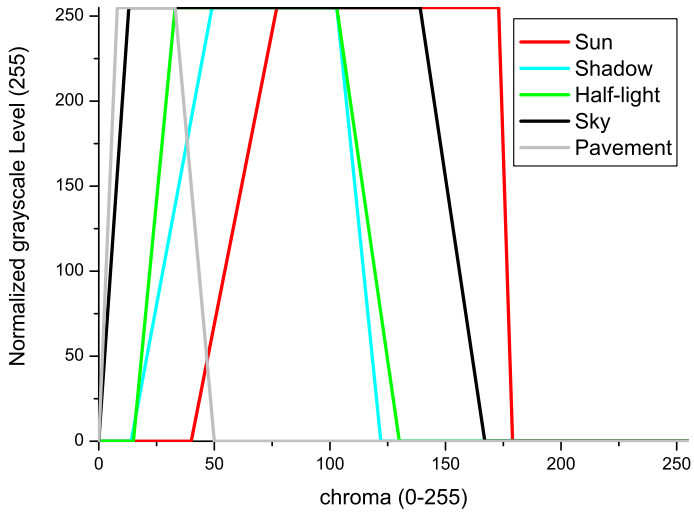


Figure 3.11: HS components in HSL space for red road signs. Each surface represents the information from both H and S together and for each of the objects under investigation. Red: sunlight; green: half-light; blue: shadow; black: bricks.

- HCL color space, component H (Fig. 3.12a). The behavior here is very similar to the HSL case, the only difference is that at the lower part of the spectra the pavement partially influences the other components. Unfortunately in this case the majority of the pavement values are located in the road sign region. The sky, as in HSL, may have values very similar to those belonging to the road signs.
- HCL color space, component C (Fig. 3.12b). The most significant data here is that the chroma almost covers the complete amount of information provided by shadows, half-light and sunlight, while pavements have very low chroma values.
- HCL color space, components HC (Fig. 3.13). In this case the regions belonging to the pavement and sky are completely overlapped by the



(a)



(b)

Figure 3.12: Blue models of LUT for components H and C in the HCL color space for the cases under investigation and depicted in the label of the graph.

road sign information. However, it is interesting to note that the regions due to false positives are small and are in areas of low saturation.

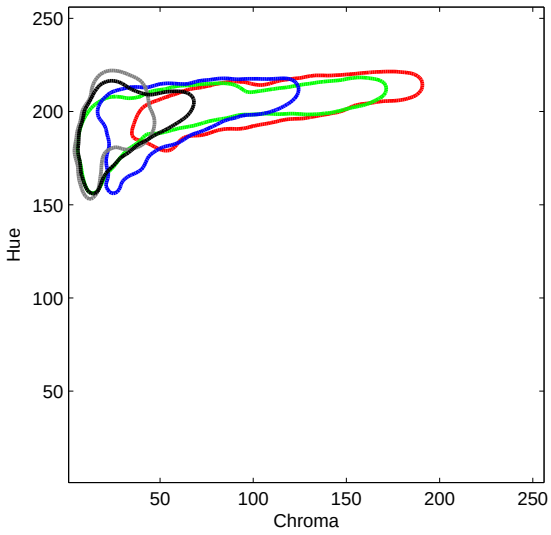


Figure 3.13: HC for blue

3.1.3 Experiments on Normalization of Color.

As has already been discussed in the state of the art at the beginning of this document, the main problems associated with object detection when using color as the principle feature is that, due to the illumination and the lighting geometry, the same object provides different color information. The main objectives laid out for this research work has involved the use of color to provide a detection reference feature, if this information is not useful due to illumination then an important part of the data available is lost. To reduce the illumination effects a wide variety of color normalization or color constancy algorithms may be used, these make an estimated calculation based on color information which is independent of illumination conditions. The main assumption considered is that of the *gray world*, meaning that the color in each sensor channel averages to gray over a given image, thus a variety of colors appear on the image with a similar probability.

The experimental work presented in this section was carried out at the Koblenz Universität where the color normalization algorithms were developed

in the AGAS following the literature [123][47][122][124]. The main goal here was to obtain preliminary indications on the possibility of using algorithms to achieve better results for red and blue road sign segmentation than those results obtained from HSL and HCL color spaces.

To obtain the histograms using different approaches, the method explained in section 3.1.1 has been used. To improve the comparison using the different algorithms an artificial mask has been built for each image which is then used on the images resulting from the conversion and which sets all the pixels which do not belong to the road sign to a low level. By doing this it is assured that the comparison is made with exactly the same pixels in every image and the time required for manual extraction of the rim or plate decreases considerably.

The algorithms for color normalization are summarized as follows:

- **Chromaticity normalization**: is based on the removal of intensity from the color pixels by normalizing the magnitude of the color pixels according to equation 3.1:

$$(r, g, b) \rightarrow (sr, sg, sb) \qquad \text{where} \qquad s = \frac{1}{r + g + b} \tag{3.1}$$

- **Color rotation in RGB** [123][47]: First a computation of the center of the clusters for all pixels $\mathbf{f}$ is carried out on the image obtaining a vector pointing to the center of gravity $\mathbf{m} = E[\mathbf{f}]$. Where $\mathbf{C}$ is the covariance matrix given by $\mathbf{C} = E\left[(\mathbf{f} - \mathbf{m})(\mathbf{f} - \mathbf{m})^T\right]$ with eigenvalues λ_i. To fulfill the requirements of the *gray world* assumption the cluster must lie along the main diagonal of the RGB space (D_{RGB}) to do this, a rotation of the cluster is required. The largest eigenvalue of $\mathbf{C}$ gives the eigenvector $\mathbf{v}$. $\mathbf{v}$ and D_{RGB} provide a plane whose normal vector is the rotation axis $\mathbf{n}$. The cluster is brought to the origin, the angle is rotated ϕ given by D_{RGB} and $\mathbf{m}$ and translated over the middle of D_{RGB}. A picture depicting the magnitudes is presented in Figure 3.14

- **Color rotation in other color spaces**: This approach follows [129] where the principal components of the clusters are computed by means of a neural network. The color spacings used are (RG, BY, WB). Following the same method as above, the cluster is rotated and lies over the BW axis. This algorithm may be used for any color spacing which has a linear dependence on RGB as XYZ.

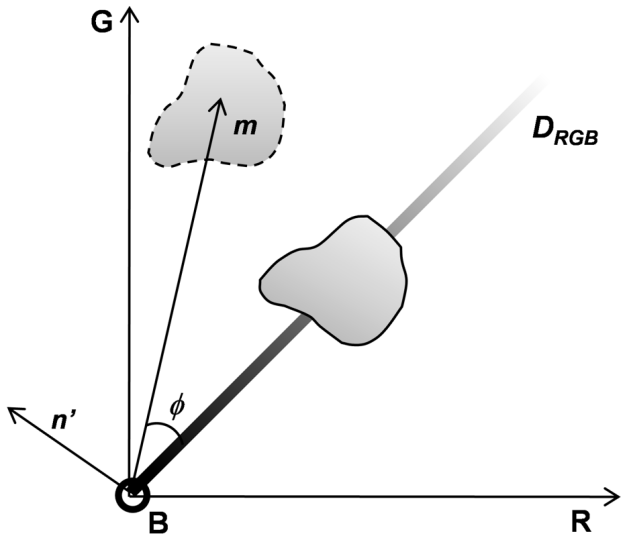


Figure 3.14: Color rotation model.

- **Whitening transform** [70]: Following the previous algorithms a first computation of the eigenvectors of the cluster is carried out. These eigenvectors are scaled by their own eigenvalues using $\frac{\mathbf{v}_i}{\lambda_i}$. At this point in [47] the algorithm is modified so that it can be applied to color normalization. Only the largest eigenvalues are considered λ_{max} as in the Color Rotation and a new cluster is built using $\frac{255}{\sqrt{\lambda_{max}}}\mathbf{V}^T(\mathbf{f}-\mathbf{m})$. The cluster is rotated 45° along the axis R^+, then 45° along B^+ and finally the result is shifted to the center of D_{RGB}.

- **Comprehensive Color Normalization (CCN)** [66]: this is a two step iterative method which starts with the normalization of the RGB pixels from the image followed by the normalization of each R, G, and B channel. The process continues until stable results are obtained. A small modification to this process is done in [47] where the result of an equalization of the intensity of the image is also taken into account, the final result is dependent on both the CCN and normalized intensity of the image.

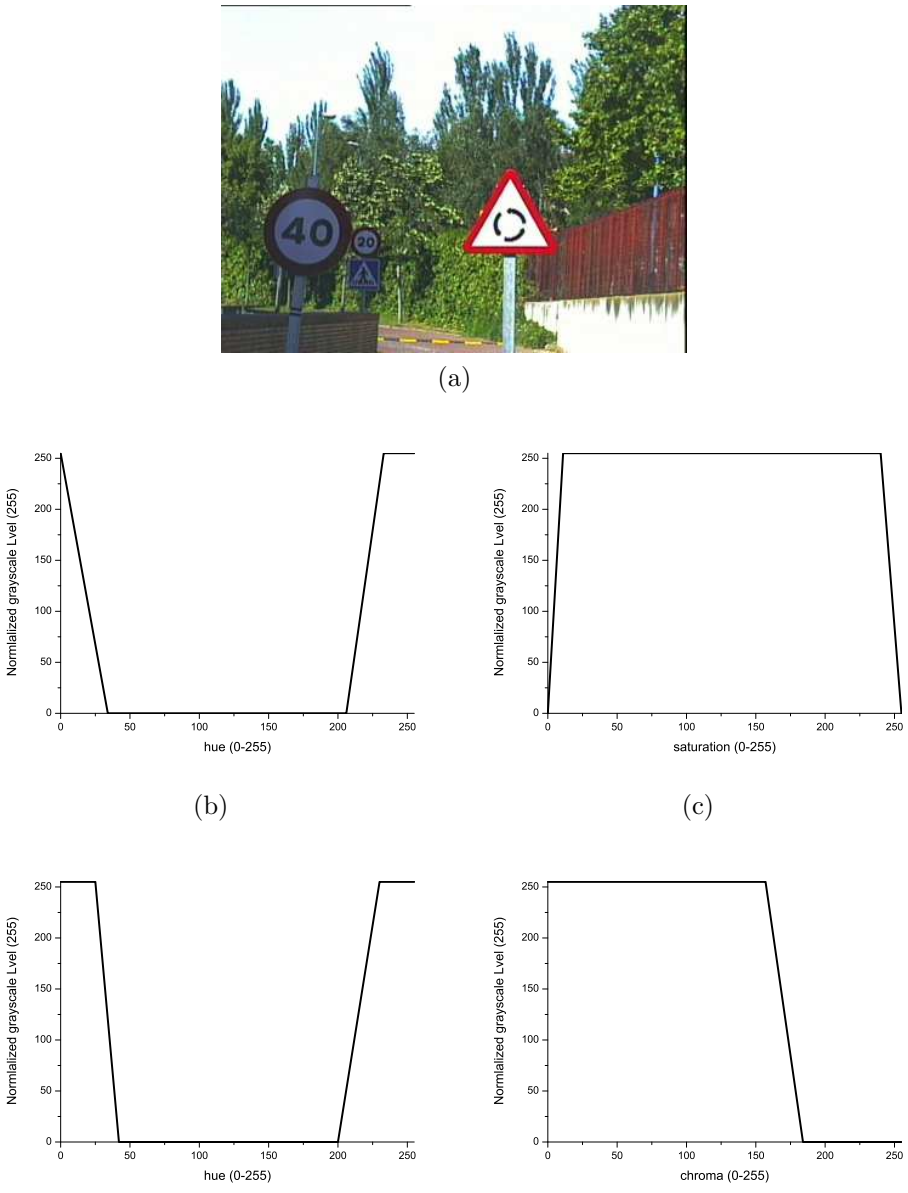
An image testing database has been used to obtain qualitative results from these conversions. This set has been chosen so that difficult cases may also be added, those which have sharp back lighting and shadowing. In order

to perform a comparative analysis of the different algorithms, once the image has been converted to each normalized color space, they are converted to HSL and HCL color spaces for the comparison. Examples of the application of these color algorithms may be seen in Figure 3.15 to Figure 3.21 for red road signs and Figure 3.22 to Figure 3.28 for blue road signs. Also added here are graphs which contain information on the hue, saturation and chroma of the HSL and HCL color spaces respectively and for each of the conversions.

It may be seen that the chromaticity normalization (fig. 3.16 and fig. 3.23), rotation (fig. 3.17 and fig. 3.24) provide similar results to those from the HSL and HCL, this indicates that it may be possible to only use the normalization from the RGB space to obtain a similar profile to that obtained in the HSL or HCL color space.

The results from the rotation in the XYZ color space are also of interest (fig. 3.19 and fig. 3.26). The two peaks which appear are for both the HSL and HCL for the red road signs and correspond to sunlight conditions and some of the half-light samples which were not as straightforward to classify as those from the sunlight or shadowing cases. The other peak belongs to the rest of the half-light and shadowed road signs. This color space is able to "decide" whether a road sign belongs to sunlight conditions or not and it is also possible to classify the uncertain half-light cases. Unfortunately when applying this conversion to the RGB image the result obtained provided an enhancement of the road sign colors but also all the gray pixels in the image, this means that this space converts the achromatic pixels to a specific region of the spectra. This is not a good result for the objectives laid out in this research work but is an interesting result as it may be used to detect end of limits road signs. The space identifies the "kindness" of red colors and places them within a specific range while taking the "unpleasant" to the D_{RGB}. An example of the enhancement applying this algorithm is presented in Figure 3.29.

The CNN algorithm (fig. 3.20 and fig. 3.27) presents a similar behavior to that of the rotation of XYZ, since the histograms for the red road signs and good illumination conditions are always located in the areas where they are assumed to be, while the rest of the information is distributed along the graph. But the case for blue road signs is completely different since the information is compressed for the HSL and HCL color space. The saturation and chroma show the same behavior for both red and blue colors as the HSL and HCL color space, the saturation level can take on any value depending on the conditions while the chroma is limited to low and medium value regions.



(a)

(b)

(c)

(d)

(e)

Figure 3.15: Original RGB image sample and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

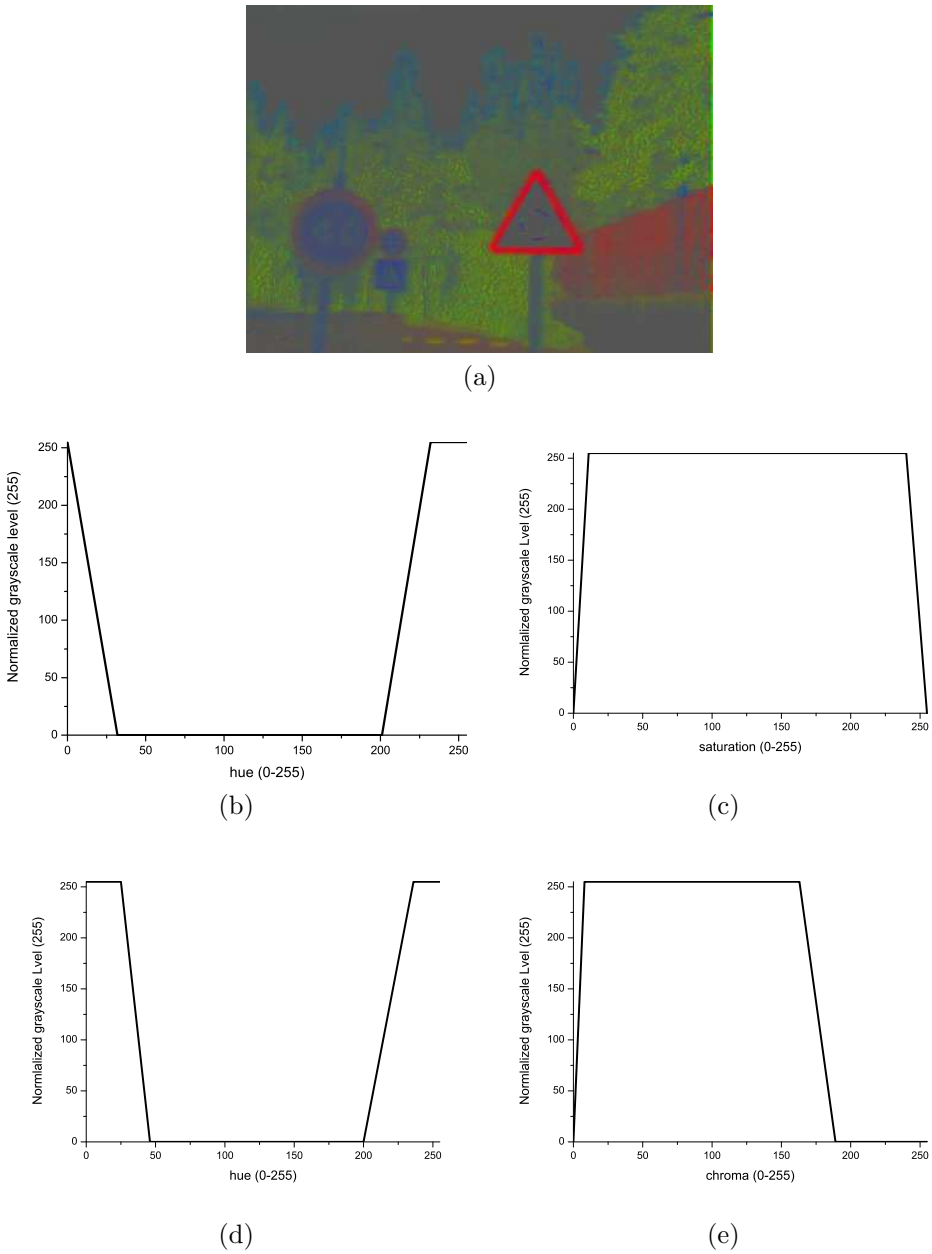
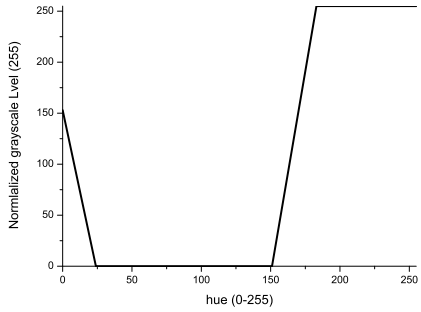
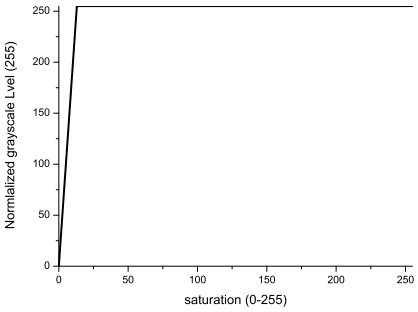


Figure 3.16: Image sample obtained using the Chroma normalization algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

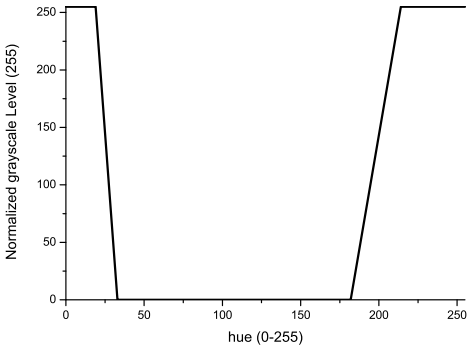
(a)



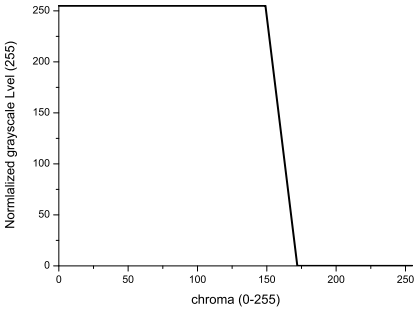
(b)



(c)



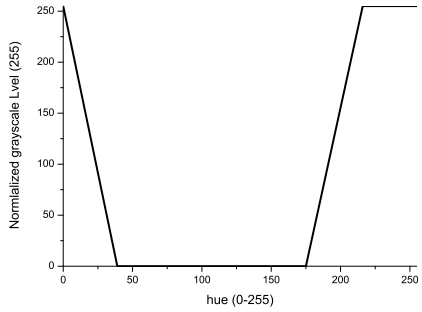
(d)



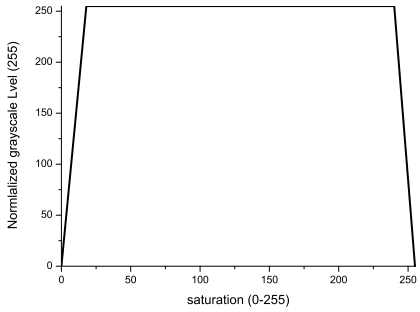
(e)

Figure 3.17: Image sample obtained using the application of a simple rotation algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

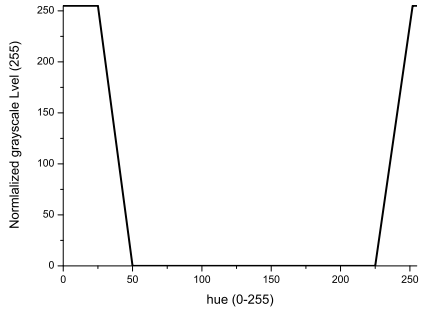
(a)



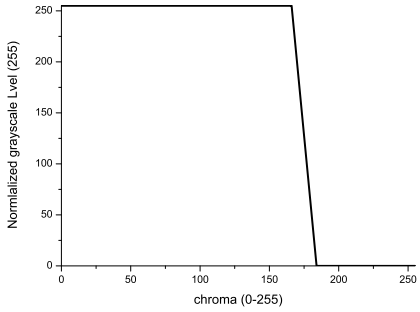
(b)



(c)

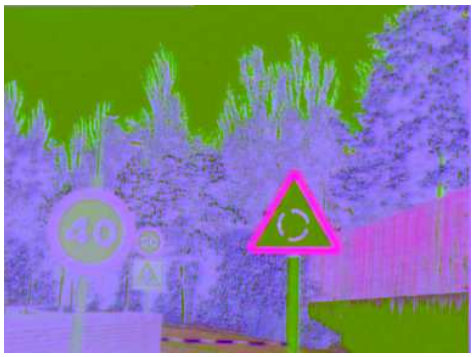


(d)

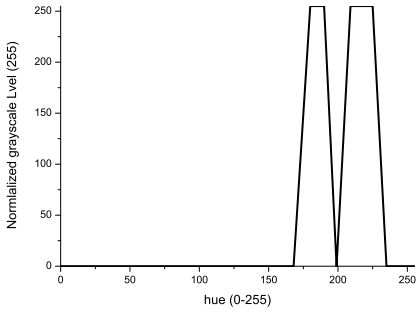


(e)

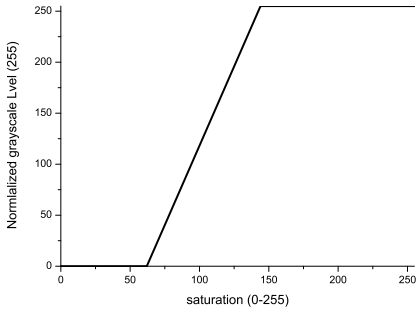
Figure 3.18: Image sample obtained using the rotation in the RG-BW color space and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.



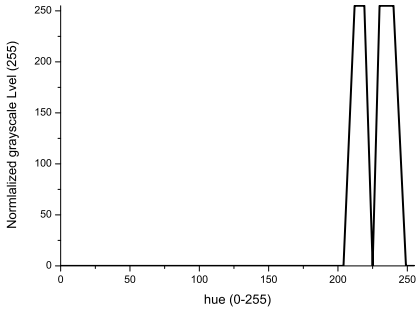
(a)



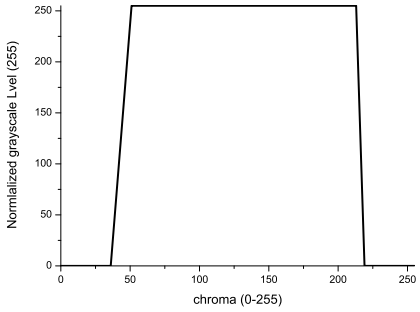
(b)



(c)



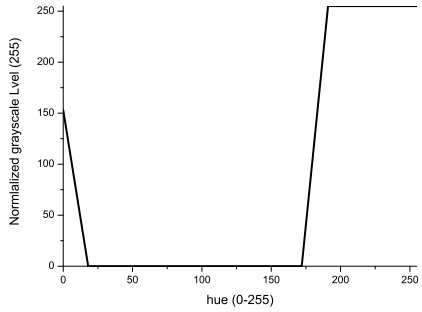
(d)



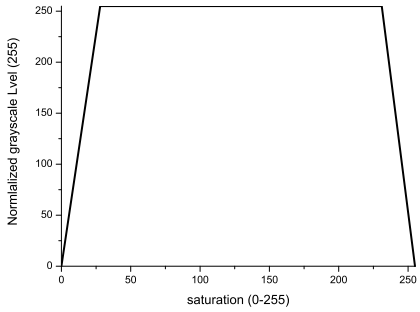
(e)

Figure 3.19: Image sample obtained using rotation in the XYZ color space and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

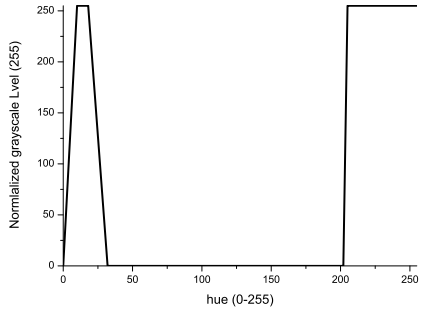
(a)



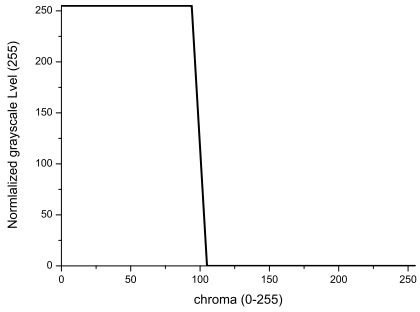
(b)



(c)

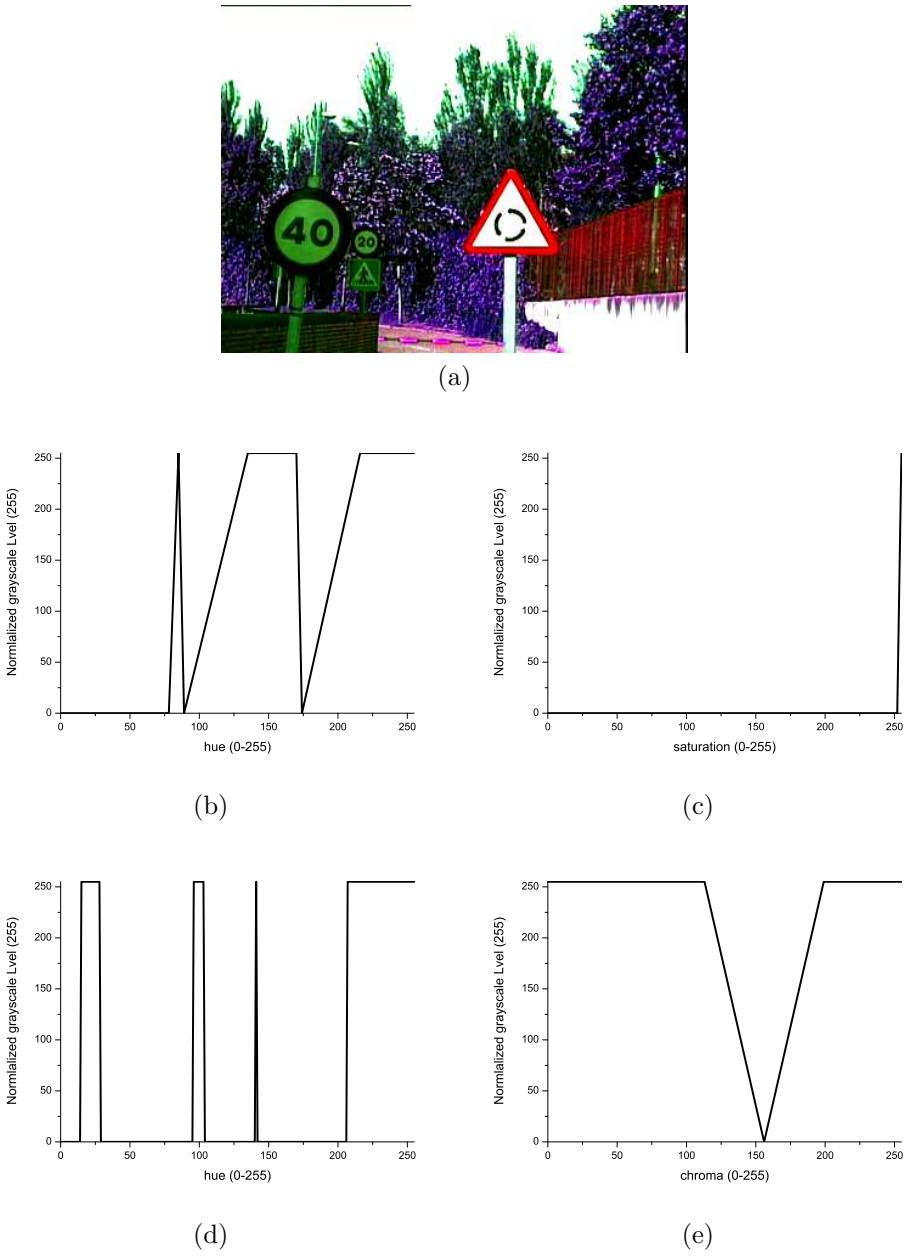


(d)



(e)

Figure 3.20: Image sample obtained using the application of the Finlayson modified algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.



(a)

(b)

(c)

(d)

(e)

Figure 3.21: Image sample obtained using the application of the whitening algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

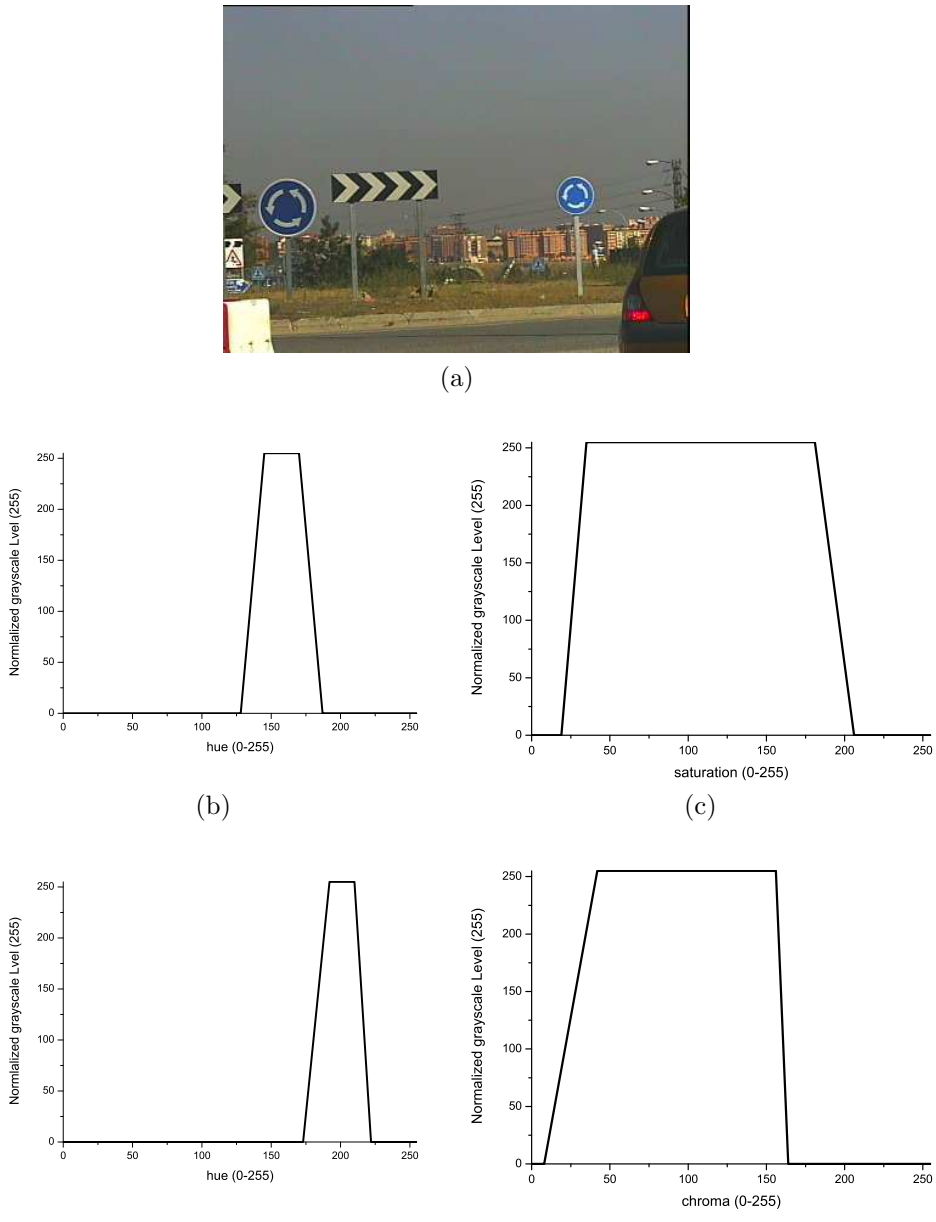


Figure 3.22: Original RGB image sample and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

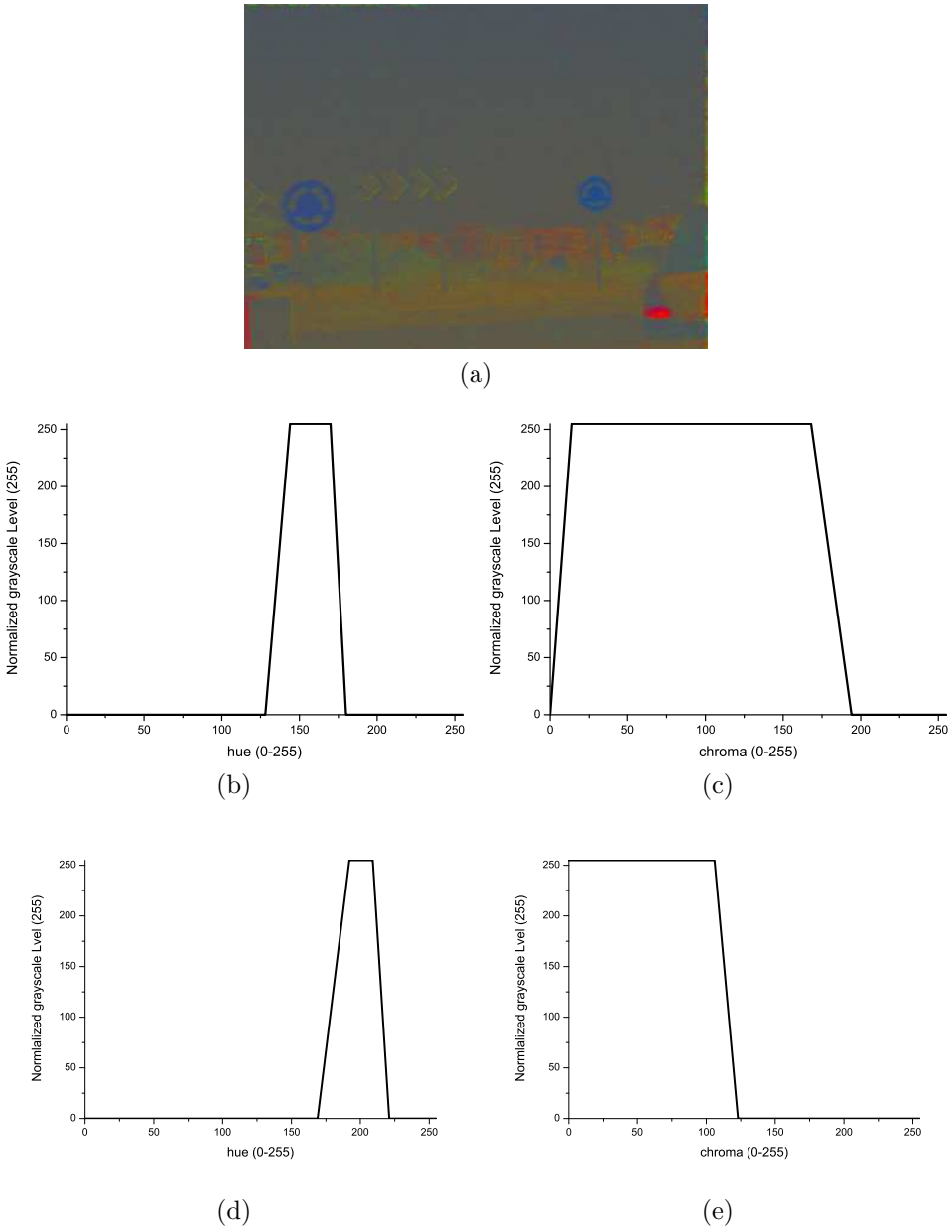
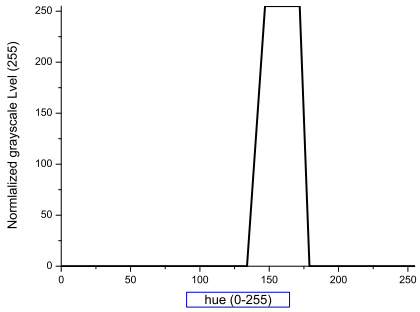
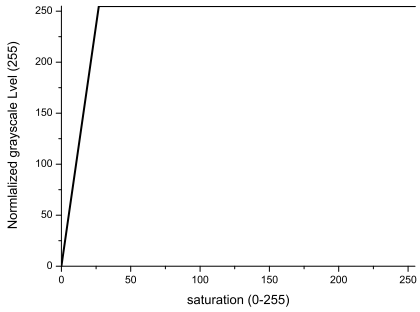


Figure 3.23: Image sample obtained using the Chroma normalization algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

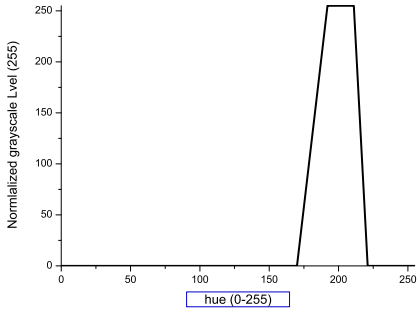
(a)



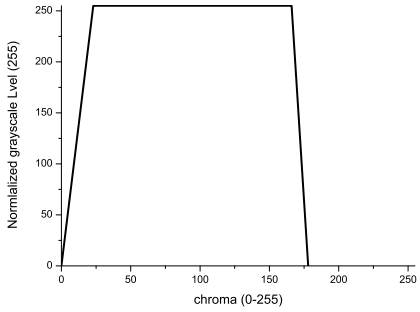
(b)



(c)

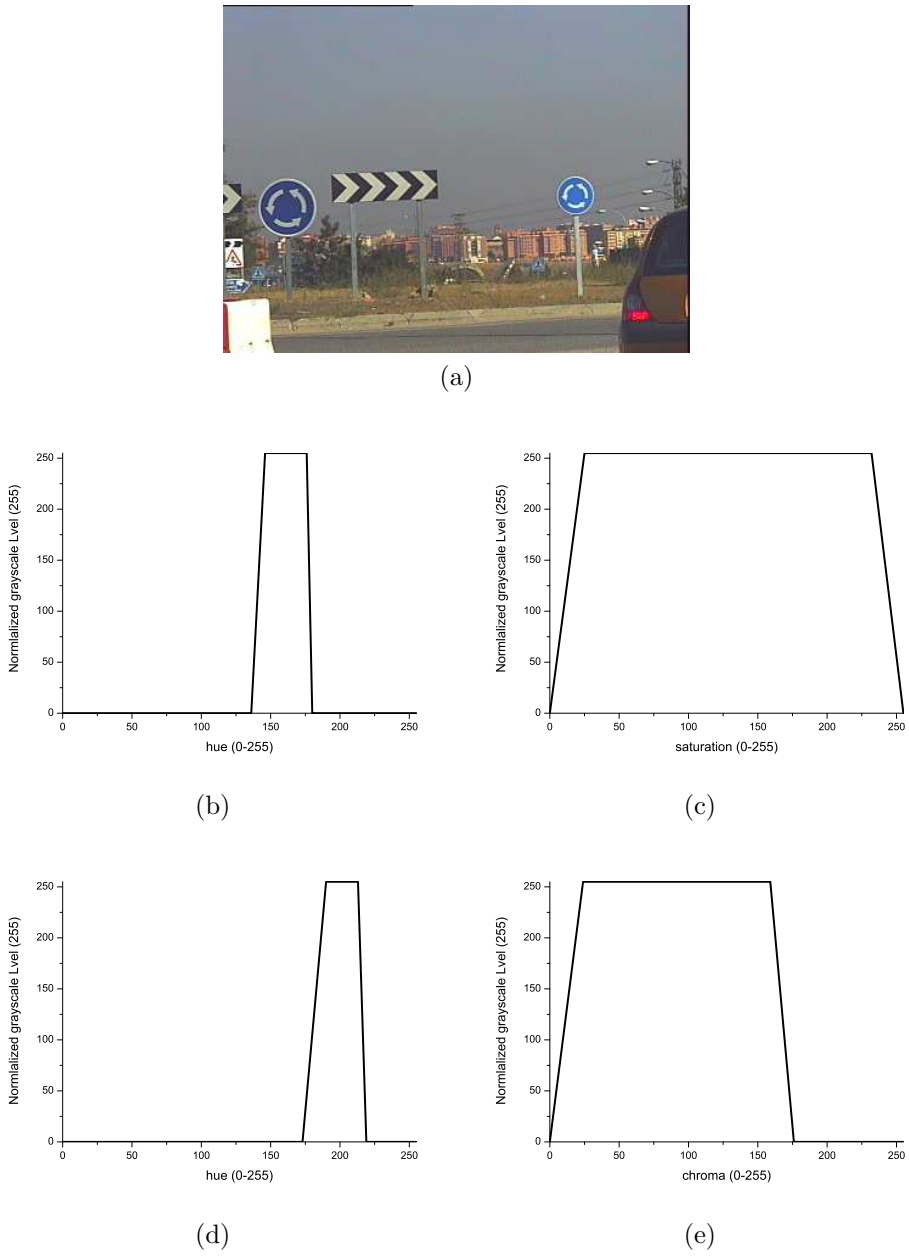


(d)



(e)

Figure 3.24: Image sample obtained using the application of simple rotation algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.



(a)

(b)

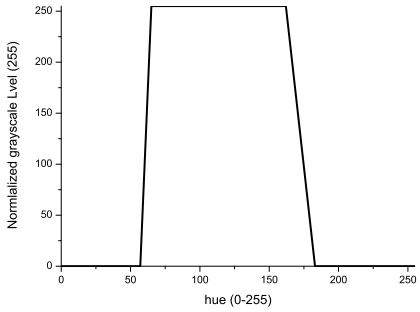
(c)

(d)

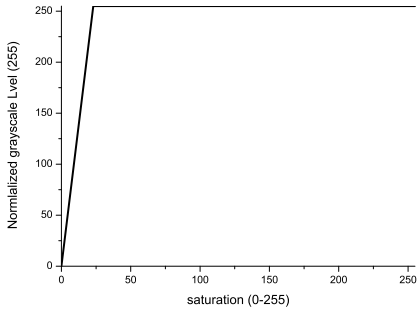
(e)

Figure 3.25: Image sample obtained using the rotation in the RG-BW color space and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

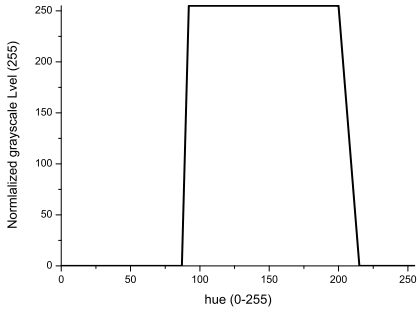
(a)



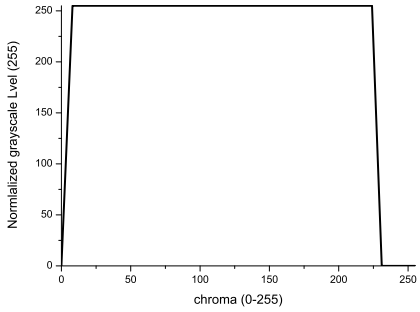
(b)



(c)

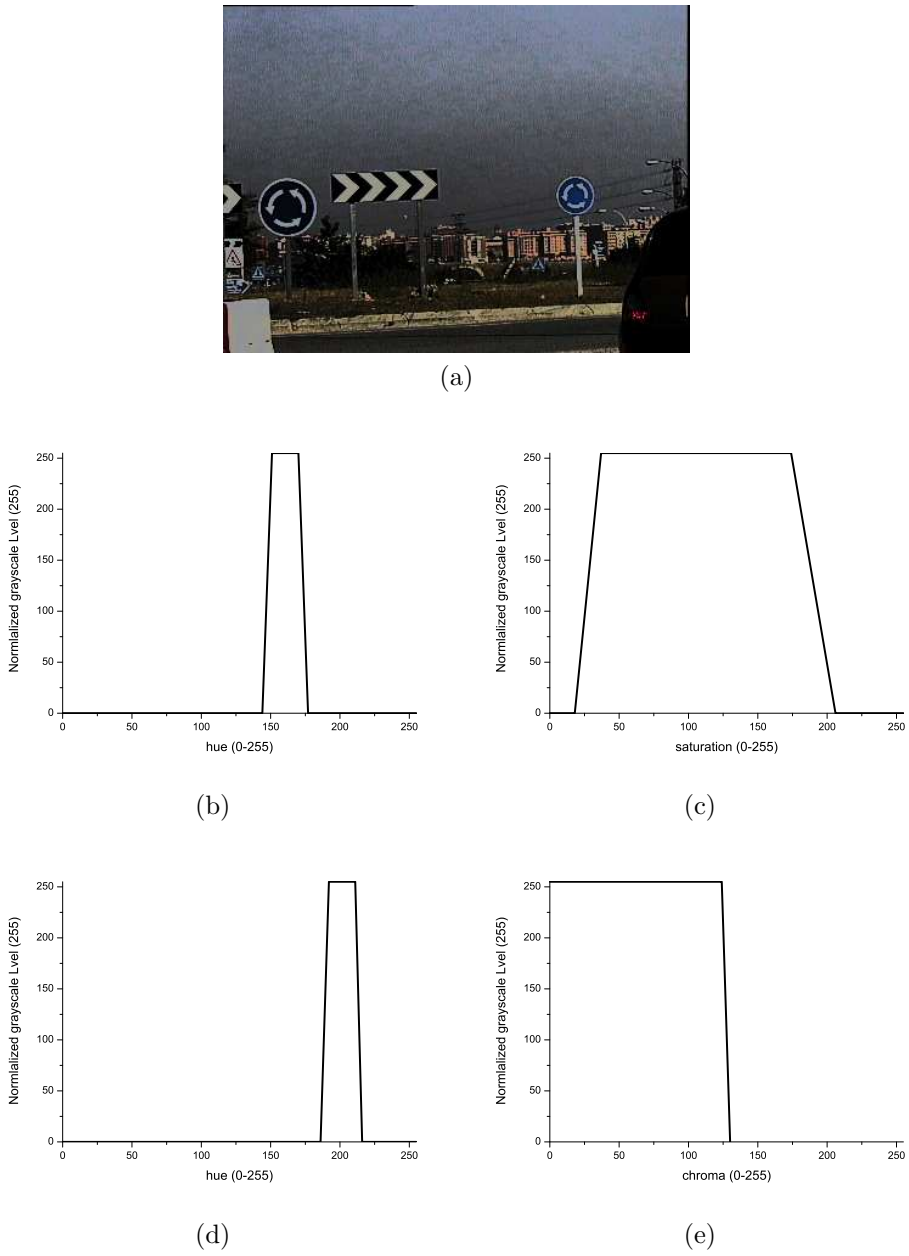


(d)



(e)

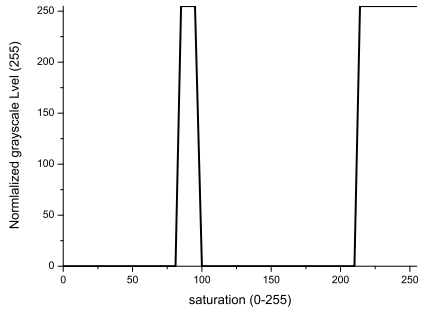
Figure 3.26: Image sample obtained after rotation in the XYZ color space and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.



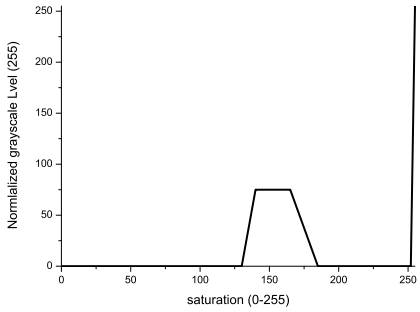
(a) (b) (c) (d) (e)

Figure 3.27: Image sample obtained after application of the Finlayson modified algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.

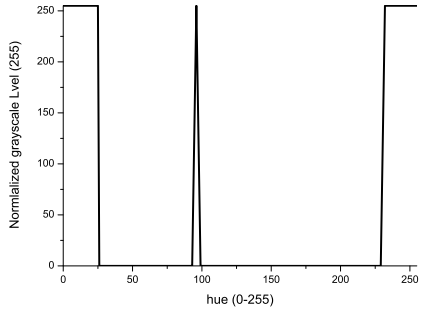
(a)



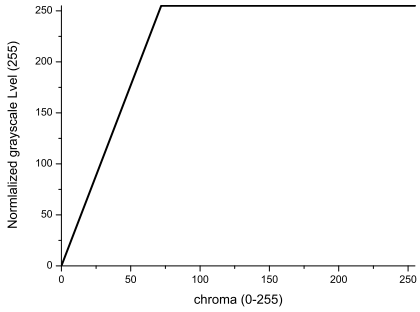
(b)



(c)

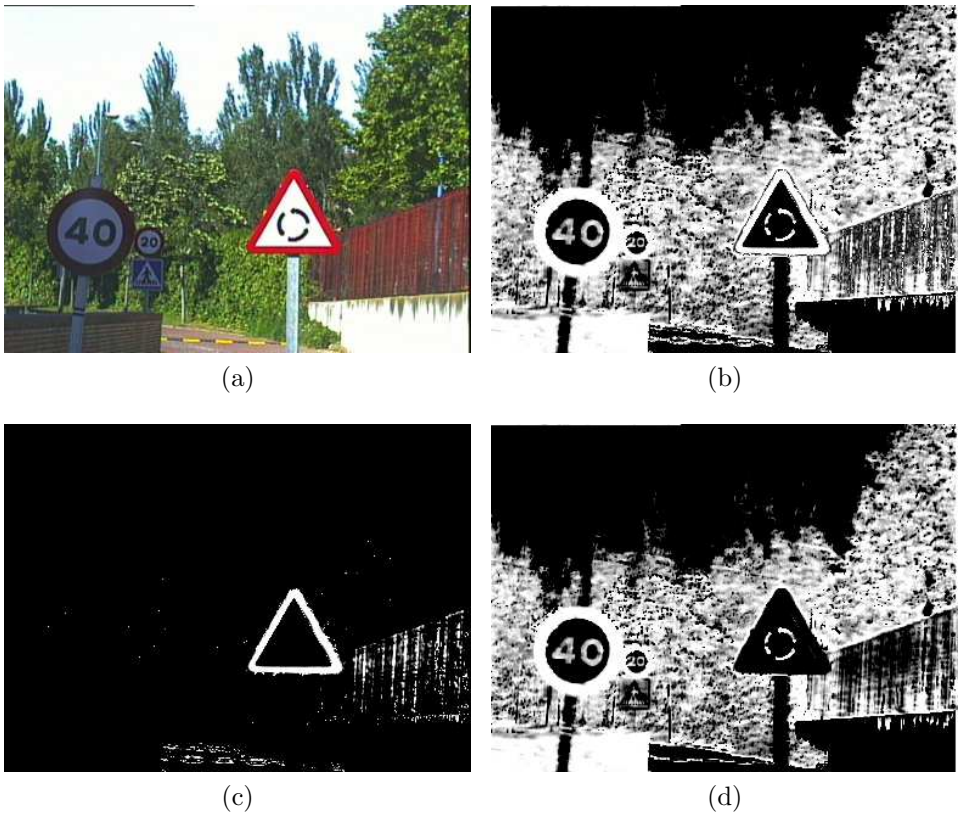


(d)



(e)

Figure 3.28: Image sample obtained using the application of the whitening algorithm and the model of its components in HSL and HCL. b) hue in HSL, c) hue in HCL, d) saturation in HSL, e) chroma in HCL.



(a) (b)

(c) (d)

Figure 3.29: Example of enhancement of color components by rotation of the XYZ color space. a) Original image. b) Full conversion through the enhancement due to the two peaks. c) Enhancement of sun condition pixels. d) Enhancement of the rest of conditions, shadow and half-light pixels.

Finally, it is also worth mentioning that the whitening algorithm (fig. 3.21 and fig. 3.28) spreads the peaks of H all over the graph for red road signs in the HSL space while in the case of the blue road signs it concentrates the histograms in two different regions, but no assignation can be done to any of the groups of interest. Another effect of this conversion is that for red road signs the color is saturated while in the HCL color space the chroma may take on any value.

From the beginning of the research work carried out in this area, emphasis has always been placed on the importance of the execution time for each task. In this case a measure of the time taken for the conversion from RGB to each normalized color space has been taken and these times are presented in the Table 3.1. Here it can be seen that the conversion times are quite high considering that the detection system is based on a real time application. All of them, except chromaticity normalization, HSL and HCL depend on the information of the current frame so no previous estimation can be calculated from data and hence it is not possible to make use of the LUT which decreases considerably the time required for the conversion process. The only way to reduce the time taken to perform these tasks is to consider the following approaches:

- Convert an initial image and use the results obtained for subsequent frames. To do this it should be assured that the lighting conditions will not vary by a large amount during the journey.

- Use only a few pixels for the first conversion and extrapolate the results over the complete image. In this case it should be guaranteed beforehand that the pixels chosen are gray upon averaging.

If these approaches may be implemented it is then possible to reduce the computation time considerably and the real time requirements of the system would be satisfied.

This study has shown satisfactory results for the algorithms which have been applied, also when a *gray world* has been assumed and when the execution time is not of great importance, as in this case there have been no simplifications to the algorithms. It is important to highlight that the results obtained for the separation of the red road signs depending on the type of illumination and also the separation of the achromatic area which is still a problem that currently has no solution.

Algorithm	Time (s)
HSL	0.017
HCL	0.017
Chromaticity normalization	0.017
Color Rotation (RGB)	0.279
Color Rotation (RG, BY, WB)	0.238
Color Rotation (XYZ)	0.267
Whitening transform	0.249
CNN	0.303

Table 3.1: Time required for the conversion of the HSL and HCL color spaces and the different normalized color spaces. HSL, HCL and chromaticity normalization were tested using an AMD 64, 2 GHz, 1 GB and the rest of algorithms using an Intel Centrino Duo, 1.6 Ghz, 4GB.

3.1.4 Election of the Color Model Component

It has been demonstrated in previous sections that the different linearized models may describe the color properties of the road signs as well as the typical objects which lead to false positive detections. The aim of this section is to determine which part of the color components model best fits the lighting conditions of its own group and also those from other groups. By doing this it is envisaged that a suitable choice for the selection of the model is found which is to recover most of the road signs independently of the group from which they belong.

An outline of the experiment is as follows: Three different sets of images have been built for the red road signs. One for road signs in sunlight, the second for road signs with shadows, and the third for half-light road signs. Each specific model representing each group has been tested on its own group and on the rest of the groups and takes into account the hue component of HSL and HCL separately, and also the H and S components together for the HSL and the H and C for the HCL. Finally another model has been added which covers the information obtained from all of them, this is called a "group" model. The results of the experimental procedure will be the rate of detection for each lighting condition.

3.1.5 Results

The results presented in Table 3.2 show that for the HSL the group model provides the best performance. It may also be seen that the results are not satisfactory in the majority of cases for two color components. This is primarily due to two different factors: First, the results obtained when using two components is the outcome of the $GSV_H \cdot GSV_S$ thus the resulting grayscale value of the pixel is at times lower than that obtained when only one component is used. Secondly, the restriction for enhancing is harsher when considering two components instead of one. It is logical to assume that the best performance is that for the group case since it recovers the hue information from all of the objects which are of interest. As a result, the saturation, which was identified as a parameter filter for non-grouped cases, now has no effect when two components are taken together, as the saturation covers almost the complete range of values.

condition	model	sun	half-light	shadow	TOTAL
Sun	H	90.0	90.0	43.7	74.6
	HS	81.3	89.4	39.8	71.2
Half-light	H	90.0	87.2	47.7	74.9
	HS	80.9	85.6	46.4	70.9
Shadow	H	56.4	63.9	64.6	61.6
	HS	43.2	43.0	62.5	49.5
Group	H	84.4	81.9	76.6	80.9
	HS	83.9	80.0	76.5	80.1

Table 3.2: Results from the combination of color models (%). Group stands for a model which contains the information from the combined lighting conditions

In the case of the HCL color space, Table 3.3, the results display a slight improvement over the HSL color space with an increase of almost 5%, and the best performance is demonstrated again for a group model but for two components instead of one. The results obtained are similar, thus at this moment it is not possible to outline the final conclusions to the problem of finding the component or components which obtain the best results for the current problem.

Nevertheless, it seems clear that the grouping of models achieves the best

performance and is independent of the color model, from here onwards the group model will be used for the remaining experiments.

condition	model	sun	half-light	shadow	TOTAL
Sun	H	93.0	86.2	62.7	80.6
	HC	90.3	87.4	65.8	81.2
Half-light	H	87.2	88.3	50.4	75.3
	HC	87.1	86.7	48.9	74.2
Shadow	H	56.2	67.4	73.6	65.7
	HC	42.6	45.7	72.5	53.6
Group	H	90.4	86.6	81.2	86.1
	HC	90.4	84.7	82.5	86.8

Table 3.3: Results from the combination of color models (%). Here a group stands for a model which contains the combined information form the three different lighting conditions

3.2 Election of the Models

3.2.1 Introduction

As previously stated, the method used to search for road signs in images is based on template matching, where the normalized correlation function presented in Eq. 3.2 is used to evaluate the similarity between the candidates and road sign models. In this equation f and g represent the value of the intensity of the image and the model to be correlated in a certain pixel; $\bar{f}$ and $\bar{h}$ are the mean values of the intensities of the image and model; x and y are the coordinates of the images where i and j are the common indexes which move over the images.

$$r = \frac{\sum_i \sum_j \left(f\left(x+i, y+j\right) - \bar{f}\right)\left(h\left(i,j\right) - \bar{h}\right)}{\sqrt{\sum_i \sum_j \left(f\left(x+i, y+j\right) - \bar{f}\right)^2 \sum_i \sum_j \left(h\left(i,j\right) - \bar{h}\right)^2}} \tag{3.2}$$

On performing a study of the images obtained from different conversions (RGB, HSL and HCL) it has been found that the rim and plates of the road

signs usually present a transition between the rim and inner part of the road sign (in-background) and between the border of the plate and the rest of the image (out-background). For the ideal case, the border of the road sign should be perfectly sharp so that the transition between the rim and the plate and the in-background and out-background are a perfect step. The problem is that the scintillation of the atmosphere, distance between the camera and the road sign, color uncertainties, limitations of the camera and movement of the vehicle produce the effects of border smoothness on the plate and rim. An example of this may be seen in Fig. 3.30.

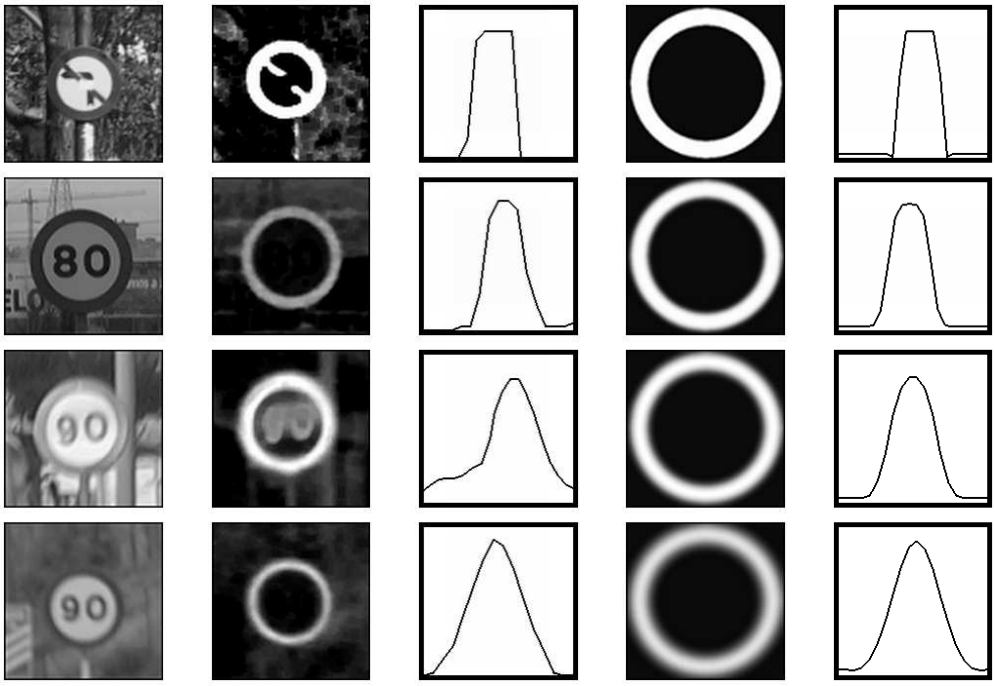
To deal with this problem, no *a priori* assumption has been established and different models corresponding to the most typical profiles have been built and tested to be used for detection. Some examples of the process are presented in Fig. 3.31.

(a) (b) (c)

Figure 3.30: Example of a real image in which a RGB enhancement and HSL enhancement has been applied. It may be seen that there is a blurring which depends on the color space. a) Original grayscale image. b) RGB enhancement. c) HCL enhancement.

3.2.2 Building the Models

This study has been carried out by first creating a set of models, some of which are depicted in Fig. 3.32, 3.33 and 3.34. The main differences between these models is the slope width corresponding to the in-background and out-background, and the consideration of pixels known as do-not-care pixels, where this means that they will not be taken into account in the final computation obtained from the the normalized correlation. The purpose for this



(a) (b) (c) (d) (e)

Figure 3.31: Examples of models for detection and the creation process. From left to right. column a) Original images. b) RGB enhancement. c) Profiles of the rims from the images of column b). d) Proposed models to fit column b). e) Profiles of the rims from the images of column d). It may be seen that the blurred models are a better fit in the cases where the profiles are fuzzy.

setting is to avoid, if possible, situations where the in and out-background are also enhanced, and have a lower weight in the global similarity calculation.

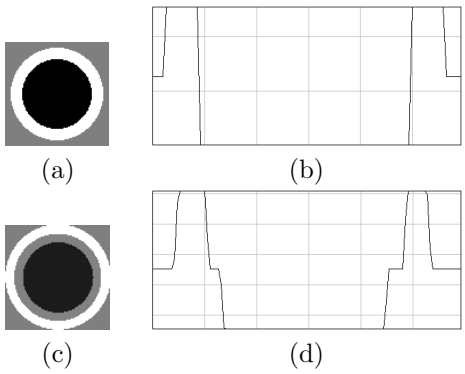


Figure 3.32: Sample of blurred models set 1, model and profile. In this case the two models are built using a step function which is only set to the low value for the in-background case.

Using this technique the rest of the models used for detection can be created using the same procedure. Each of them are presented in 3.35a, also with examples of the corresponding class in 3.35b.

3.2.3 Experiments and Results

The models have been tested using a set of different trial images, this is done to determine the best model for the detection procedure. Test images were chosen manually so that blurring is almost always present. This experiment was carried out for circular red road signs in the RGB color space, since RGB presents the most amount of blurring behavior and the red color is one of the two colors which has a larger spectra and as a result more blurring.

Tables 3.4, 3.5, 3.6 and 3.7 contain the statistics related to the best models which are represented in Figure 3.36. The parameters in these tables are as follows: $\bar{x}$ is the average of the results given by the normalized correlation function for each model over the detection image set. σ is the standard deviation of the data results. $\uparrow_{40\%}$ is a percentage value for the results which are over 40% of the score, where this has been set manually after several trial experiments and which avoids a false positive majority without the lack of true detections. Performance is the % of times that each model achieved the best results among all models. It is important to note that 0 performance

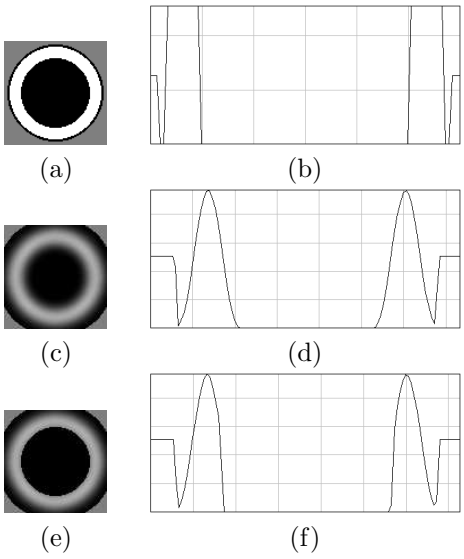


Figure 3.33: Blurred models set 2, model and profile. a) depicts the source model which in this case forces a clear difference between the background and sets part of the out-background to do-not-care values. c) blurred in and out backgrounds, while in d) the in-background model is set to a low value thus preserving the same size as if it were a binarized model.

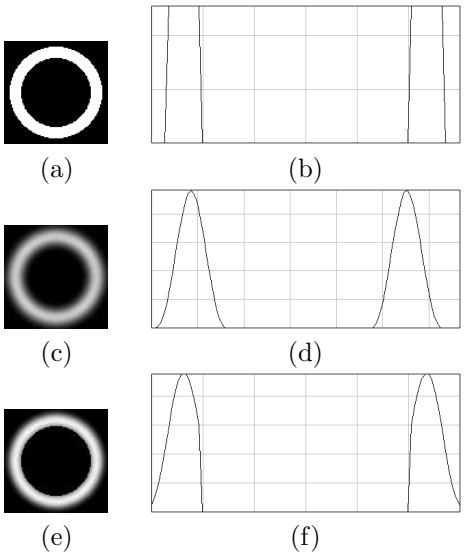


Figure 3.34: Blurred models set 3, model and profile. a) step function between rim and in and out-background. In c) in and out-background are blurred, while in e) in-background is set a low value.

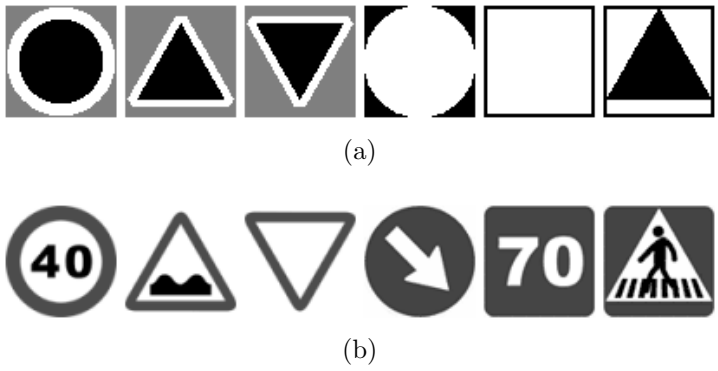
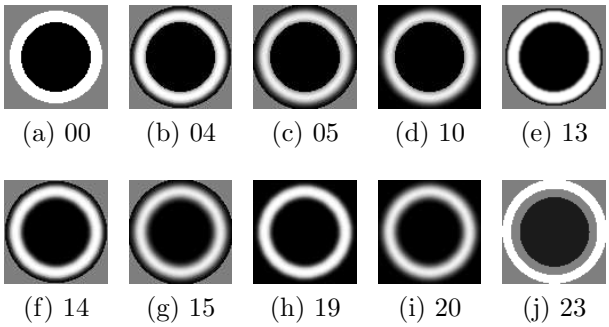


Figure 3.35: Binarized models used for detection and samples of each class. row a) From left to right: Prohibition, Danger, Yield, Obligation, Indication, Crosswalk. row b) Maximum speed 40, Uneven road, Yield, Obligation to drive on the right hand lane, Advised maximum speed 70, Crosswalk.

does not mean that the model is inadequate, but that it never achieved the best rate in an experiment, though the mean value in every experiment may be high.



(a) 00 (b) 04 (c) 05 (d) 10 (e) 13

(f) 14 (g) 15 (h) 19 (i) 20 (j) 23

Figure 3.36: Best models for detection and their code.

Model	$\bar{S}$ (%)	σ(%)	$\uparrow_{40\%}$ (%)
23	83.3	17.2	84.9

Table 3.4: Statistics of model 23. $\bar{S}$ is the mean value of the total results rate for the complete set of detections. σ is the standard deviation of the data. $\uparrow_{40\%}$ is the percentage of the results which are over 40% of the total detection results

Model \ Range(px)	30-35	35-40	40-45	45-50	50-55	55-60	60-65
23	73.7	87.8	87.7	89.1	93.6	87.5	73.3

Table 3.5: Results in % of the model 23 depending on the size of the road sign detected in the image.

From the results presented in Table 3.4 and Table 3.5 it can be seen that model 23, which corresponds to the profile of 3.32d, achieves satisfactory results. An improvement in the results is seen when this model is compared to the results obtained from the the rest of the models, this is due to the increased flexibility achieved by setting the two annular parts of the road sign to do-not-care values. It is obvious that in this case the same model can

deal with different road sign rim widths for detection as the model locates road signs with dimensions within the limits of the annulus. Unfortunately due to the specific characteristics, its flexibility renders a large amount of false positives which results in increased time consumption during the detection procedure which is 2.4-times greater than that obtained when using the binary model (Model 00). The extra detection time is due to:

- Checking the overlapped detections due to different model occurrences at the same point.
- Normalization of the candidate (resizing, preprocessing) in the Recognition stage.
- Classification of the candidate.

One way to deal with the overlapping is to optimize the number of models thus avoiding the overlap of the annulus from different models. Even though detection would be improved this brings about two further problems, which are:

- In the recognition stage knowledge of the exact model size is required, this is due to the process which involves background masking. Using this model, the limits of the road signs are unknown as the same model can be correlated with different sizes, so the noise due to the background cannot be erased.
- The accuracy in the size of the detected road sign is very important during the tracking stage, the information provided from the dimensions of the model is a key part of successful predictions.

Although this model achieves the best detection results, it does not have the best detection rate when compared to that of Model 00 which follows the profile of Fig. 3.32a. Also there are other satisfactory results which have been obtained from intermediate blurring levels, in fact only one of the best models, number 13, belongs to low blurring and none of the high blurred models are among the best performers.

To assure high detection rates and according to results from experimental trials, the best choice is a binary model, however it is important to note that in 62% of cases, when the binary model did not reach 40% of the detection

Model	$\bar{S}$ (%)	σ(%)	$\uparrow_{40\%}$ (%)	Performance (%)
00	72.6	11.2	90.6	63.8
04	62.6	6.9	44.2	0
05	62.3	7.2	50.7	0
10	62.9	8.2	60.1	0
13	63.4	8.3	62.3	0
14	64.8	7.6	75.4	1.4
15	63.7	7.9	72.5	3.6
19	64.8	9.2	72.5	6.5
20	64.1	9.5	68.1	9.4

Table 3.6: Statistics of the best blurred models. $\bar{S}$ is the mean value of the total results rate for the complete set of detections. σ is the standard deviation of the data. $\uparrow_{40\%}$ is the percentage of scores which are over 40% of the total number of detections. Performance is the % of times that the model obtained the best fit over all the detections. It is important to note that 0 means that the model did not obtain the best results from an experiment although it may have achieved high detection rates.

Range(px) / Model	30-35	35-40	40-45	45-50	50-55	55-60	60-65
00	69.3	68.1	71.8	75.1	76.4	75.8	79.1
04	<40.0	58.2	61.0	61.6	62.6	65.0	64.9
05	55.1	56.8	59.2	63.9	63.4	63.9	68.7
10	57.0	58.9	63.6	61.4	64.2	68.4	69.3
13	<40.0	56.4	63.9	59.6	65.7	65.5	69.1
14	<40.0	63.2	63.0	62.5	66.7	63.3	70.0
15	<40.0	61.3	62.3	63.5	64.5	62.7	67.8
19	<40.0	61.1	61.7	64.6	67.1	67.4	68.3
20	55.4	61.6	60.5	62.8	64.5	67.8	69.2

Table 3.7: Scores over 40% of the best blurred models depending on the size of the road sign detected on the image.

result, there was a blurred model that achieved better results, thus they may have had success or improved results when the binary model failed.

The influence of distance on the choice of model has also been studied. The detection distance range in pixels [30- 64] has been divided in intervals where the results for each model are taken. Table 3.7 shows the results for the models which achieved the best results when concerned with the size of the road sign in the image. It can be seen that the best performance is presented by the binary model showing that distance from the camera to the road sign is not an important factor for the choice of model. But a more in depth review of the results shows that for the best performance of the blurred model occurs for an average size of 43 pixels, this being a more reduced size, also 50% of them showed their most efficient performance for sizes under 40 pixels, thus blurred models achieve the very good results for large camera-road sign distances.

In conclusion, for a real time application where normalized correlation is used as the filtering function, the best choice is to search for road signs using a binary model as they present the most adequate average results. The use of the blurred models could be considered for a hierarchical search, mainly for detection at large distances, however this introduces problems with time where the addition of an extra model involves a considerable increase in the processing time.

3.3 Final Color and Model Analysis

Once the best color model has been studied with respect to the color of the road sign and the model to be used, it is possible to perform experiments to determine the color space that is best suited to our system. As part of this objective, a set of real images taken from the IVVI platform were used along with the HSL and HCL color spaces. The experimental procedure is the same as that presented in previous sections but with one difference, only the positive detection rate will be used to indicate the reliability of the method and also the information from the receiver operating characteristic (ROC) curves.

A ROC curve is a plot of the fraction of true positives versus the fraction of false positives for a binary classifier in which its filtering threshold is varied. A fundamental reference for developing ROC curves is in [65]. By generating this graph two valuable pieces of information are obtained:

- Threshold setting. Given a detection result it is possible to obtain information on the rate of false positives and true positives that occur for a given cutoff point. Thus, the setting of the detection parameters of the system may be done considering this information to fulfill the requirements of a given application.
- Quality of detection. The area below the curve indicates the detection performance. The ideal case arises when the true positive rate is equal to 1 over the complete graph, thus no false positives are found leaving only true detections. As the rates of true positives and false positives are normalized to 1, the area varies between 0 and 1, where 1 is the best outcome.

An example of a ROC curve is shown in Figure 3.37. This graph was obtained from one of the experiments developed in this section, in particular the H component from the HCL color space for red road signs. Each point of change in curvature correspond to a threshold, thus it is possible to determine from the value given the corresponding rates for the false and true positives, this permits *a priori* setting of the parameters for the system. In the experiments developed for this research work the first threshold will be set for results of 40% as lower values are affected considerably by the false positives.

3.3.1 Red Road Signs

The results from the experiments for the detection of red road signs are presented in Tables 3.8 and Figure 3.38. Here it can be seen that the best performance corresponds to RGB and HCL (H and HC) color spaces, where the detection rates (true detections over the total amount) is 86%, 85% and 85%, respectively. The results from the HSL are also satisfactory but achieve approximately 10% lower results when compared to the previous two cases. It is important to remember that one of the intial requirements of the system is to avoid missing any road signs maintaining the execution time of the application within the appropiate limits required by a real time system. Keeping this in mind the HSL color space will be initially dismissed.

The main disadvantage of RGB enhancement is that it cannot deal with roadworks road signs. These are very important in the system as they always provide important information on speed and danger and in all cases they are circumstantial so no data base contains their information.

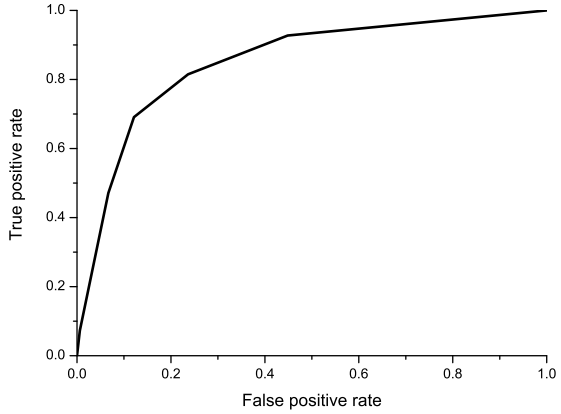


Figure 3.37: Example of a ROC curve corresponding to the experimental detection results from the red road signs and for the hue component in the HCL space.

Red				
Color space	component	TP_{rate}	FP_{rate}	area
RGB	R	0.86	0.41	0.93
HSL	H	0.78	0.13	0.76
	HS	0.78	0.13	0.74
HCL	H	0.85	0.16	0.85
	HC	0.85	0.16	0.85

Table 3.8: Final results regarding the color space and the components of color used for the red road signs. The detection rate (TP_{rate}) takes into account the detection the total amount of objects. The false positive rate (FP_{rate}) is the rate of false detections over the total amount of objects detected. The area stands for the area below the ROC curve which is normalized to 1.

The problem with this type of sign is that they have a yellow in-background. The whole plate is enhanced in the conversion so no model can be used for correlation. In the case of HSL and HCL just the rim is enhanced so all possible roadworks road signs may also be detected.

Other main differences between the results for RGB and HCL are the inconveniences presented by false positives (false positives over the total amount of detections) that are present when the RGB color space is used. It has been stated previously that this causes a 23% increase in the processing time of the application, where this is an important factor which must be taken into account when selecting the most efficient system configuration.

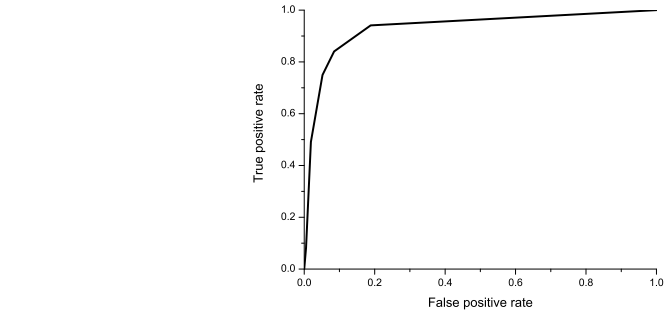
The results of the ROC curves are quite good for both spaces. The main difference is that when using the RGB the majority of false positives are present in the lower part of the graph while for HCL it is possible to find more FP with higher results. It is important to note, as stated in [65], that care must be taken when ROC curves are used to make conclusions about the effectiveness of classifiers. In this case, upon studying the intervals of FP and TP data it has been found that for RGB 81% of FP are in the interval from [40,50]% in this same interval there are only 6% of TP which cannot be lost thus the cut-off cannot be increased.

When considering all the experimental data taken, it has been found that the best choice for this application is to use the HCL color space. It is important to note that the behavior when using H or HC is the same. Hence the most appropriate and simple solution is to only use the hue component for color filtering.

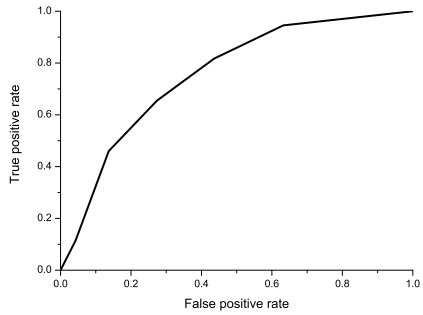
It is also important to emphasize that if only one ambient lighting condition is considered for detection i.e. sunlight, half-light or shadow, then it is much more straightforward to use the information from saturation and chroma to adequately filter the rest of the objects. This is especially true for the case of sunlight conditions where the detection results may have no errors.

3.3.2 Blue Road Signs

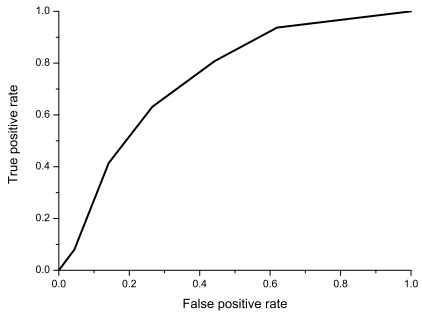
The results of the experiments for detection of blue road signs are presented in Table 3.9 and Figure 3.39 here it is seen that the best results are achieved when using the HCL color space when concerned with the TP_{rate}, these results are similar to those from the HSL, and much improved over the RGB color space: 88%, 87%, 40%, respectively. For this reason, implementation



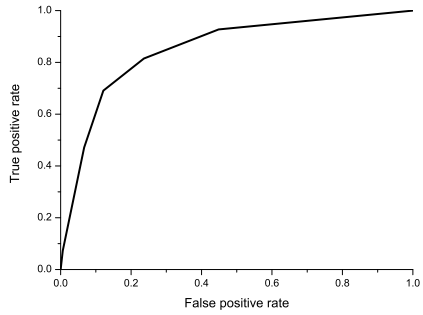
(a)



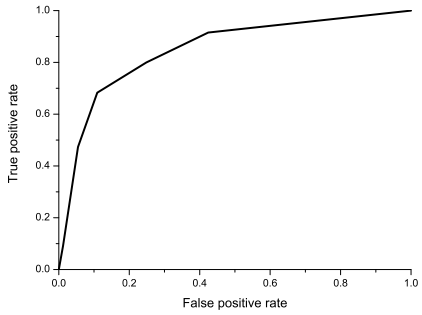
(b)



(c)



(d)



(e)

Figure 3.38: ROC Curves corresponding to the results of the study for red road signs with an experimental threshold of 40%. a) RGB. b) HSL component H. c) HSL components H and S. d) HCL component H. c) HCL components H and C.

of the RGB will not be considered.

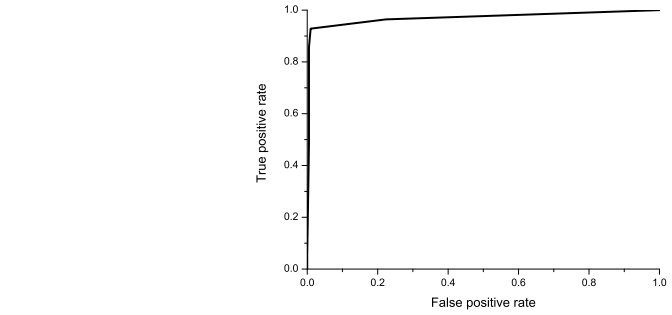
It may be noted that the results from the HSL and HCL using the information from the two components (HS or HC) are inferior to those when using only one (H); in fact nearly 16% less in HSL and 29% in HCL. This is primarily due to the restrictions that the saturation and chroma components impose on the colors allowed, since some of the pixels which fulfill the conditions for the hue component may not fullfill those for the saturation or chroma components.

Blue				
Color space	component	TP_{rate}	FP_{rate}	area
RGB	B	0.40	0.79	0.97
HSL	H	0.87	0.81	0.97
	HS	0.71	0.59	0.91
HCL	H	0.88	0.79	0.96
	HC	0.59	0.43	0.85

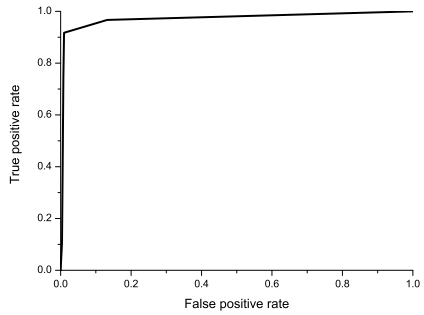
Table 3.9: Final results regarding the color space and the components of color used for blue road signs. The detection rate (TP_{rate}) takes into account the rate of detections over the total amount of objects located. The false positive rate(FP_{rate}) is the rate of false detections over the total amount of objects located. The area stands for the area below the ROC curve which is normalized to 1.

The FP_{rate} is outstandingly high when just one component is used from the HSL and HCL color spaces, which as stated before, has a bad influence on the requirements for a real time application. Fortunately, when considering the ROC curve it is easily seen that the majority of false positives are in the range of [40,50]% that is, the first line of the curve. Carrying out simple calculations it is possible to increases the detection threshold to reduce the FP rate without influencing the detection rate.

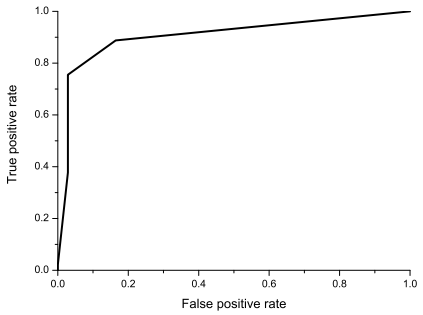
For the HCL case 85.0% of false positives are in the range of [40,50]% , while the true positives in the same interval are 4.9% of the total amount of detections. For the HSL color space in the same range there exists 86.7% of false positives and 3.3% of true positives. Thus by increasing the threshold to 50% the final result is slightly better for the HSL case as the D_{rate} falls to 0.84 and the FP_{rate} to 0.36, while for HCL the TP_{rate} falls to 0.83 and the FP_{rate} to 0.36.



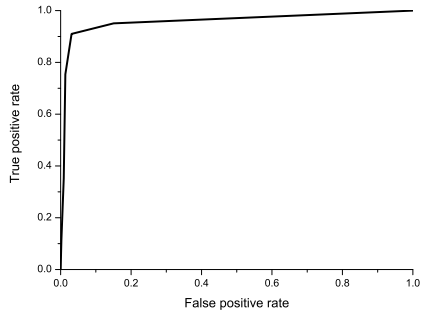
(a)



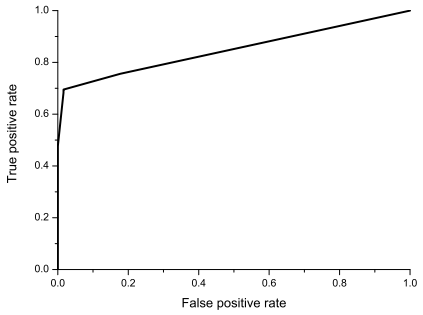
(b)



(c)



(d)



(e)

Figure 3.39: ROC curves corresponding to the results of the study for blue road signs with an experimental threshold of 40%. a) RGB. b) HSL component H. c) HSL components H and S. d) HCL component H. c) HCL components H and C.

In conclusion, HSL (H) and HCL (H) achieve acceptable results for blue road sign detection. Results obtained using these color spaces are very similar, thus it is not possible to reach any conclusions on which of these two spaces is more effective.

3.4 Discussion of the Method

A unique solution to a system which involves a certain amount of degrees of freedom is difficult to attain. Road sign detection is not an exception, as in any other case there must be a balance between the advantages and disadvantages of each method to finally reach a compromise which allows the best and most suitable results possible to be obtained.

An example of this has been shown in this section where use has been made of the road sign color. It is important to consider that road signs are fully defined by their color, shapes and symbols. Taking advantage of the color information is a logical and natural solution for the detection process, even though and has been seen throughout this section, there are a number of problems associated with the use of this property.

In the work performed and presented in this thesis, template matching along with the normalized correlation function has been used as part of the detection stage. The use of these tools have problems associated with them while at the same time other problems are solved due to the specific features of these methods. i.e. even though the normalized correlation function severely punishes the value of the grayscale difference level of the pixels, it is not necessary to have more than a certain number of occurrences between the models and images to obtain a true positive, where when using most other techniques the object in the image may easily have been missed.

Following is a series of comments on significant results which demonstrate both the advantages and disadvantages of the proposed detection method.

- **Time consuming**: one of the main problems when using the normalized correlation as a method to obtain results from template matching is the time it takes. Although an optimized hierarchical search has been used in the image, the amount of classes and hence models used dramatically increases the execution time. Table 3.10 presents the execution times in function of the number of models used for a typical search.

Total number of models	Pixel interval (px)	Time (ms)
34	1	574
17	2	270
11	3	179
8	4	136

Table 3.10: Search time differences for every class and size. There are 6 different classes: Prohibition, Danger, Yield, Obligation, Indication and Crosswalk. The pixel interval stands for the difference in pixels between consecutive models. The time taken represents that for a single frame search.

By plotting the data from the Table 3.10 it is possible to find the dependence of the time of execution on the number of models per class (N_{MC}). The resulting graph is presented in Figure 3.40. By adjusting the points of the graph to a line it is possible to obtain the relation between the ($Time_E$) and the number of models used, N_{MC}. The values are obtained from the following equation:

$$Time_E = -7.55 + 16.99 \cdot N_{MC}$$

Thus, it is possible to select the most suitable amount of models for the system when considering the importance of the time parameter.

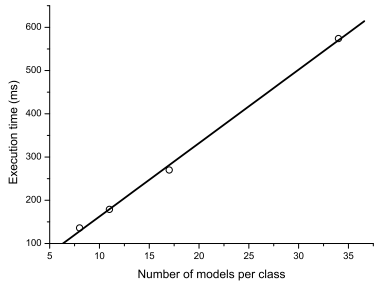
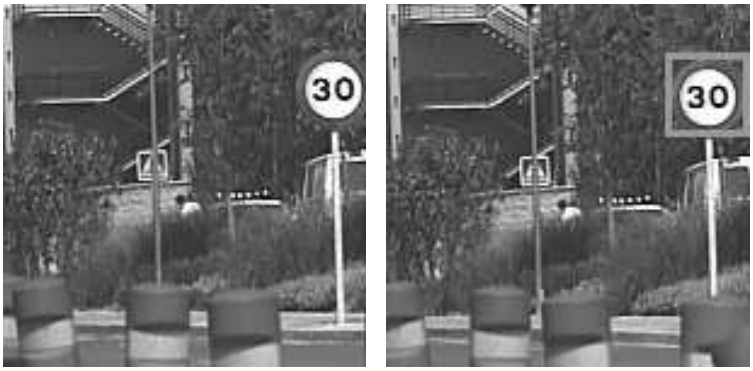


Figure 3.40: Linearization of the dependence of the execution time of the road sign search per frame and the number of models per class of the road sign used.

- **Occlusions**: a minor disadvantage is that special cases of occlusions may not be solved. This occurs when the road sign is partially out

of the image. Even in cases when just a few pixels are outside the image provokes a false negative. There are no models in the system set which fit the image and hence, the detection is impossible, this case is presented in Figure 3.41b.



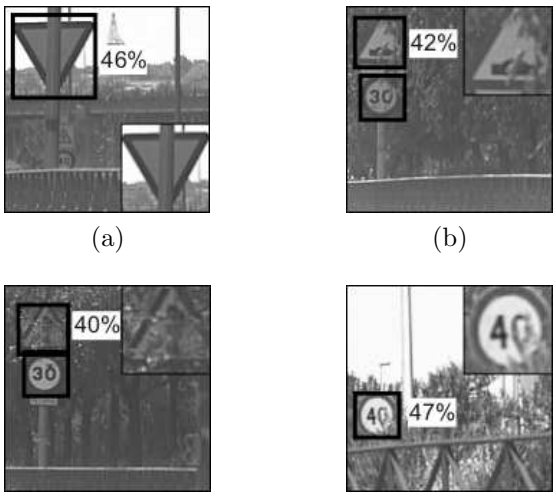
(a) (b)

Figure 3.41: Miss-detection of a partial road sign which is not completely located within the image. The two images are consecutive but as can be seen the system is not able to detect the road sign when just a few pixels are located outside the image as no model of the set can fit this case.

But other occlusions may be satisfactorily solved without using any special algorithms as the most important parameter for detection is that a certain percentage of the road sign is visible in the image. As a result road signs which are partially occluded may be solved and present a true positive. Examples of this are presented in Figure 3.42b where even though there is an occlusion the model can partially match the road sign and hence provide a true detection.

- **Rotations** The system is robust for low levels of perspective changes and small rotations as depicted in Figure 3.43. Due to the distance of detection the effects of small rotations and skewing only slightly affect the search results, thus adding to the system advantages. Even though the road signs are in general attached to straight poles, on curved trajectories the camera suffers a roll rotation due to the movement of the vehicle. More information is provided in Appendix C.0.4.

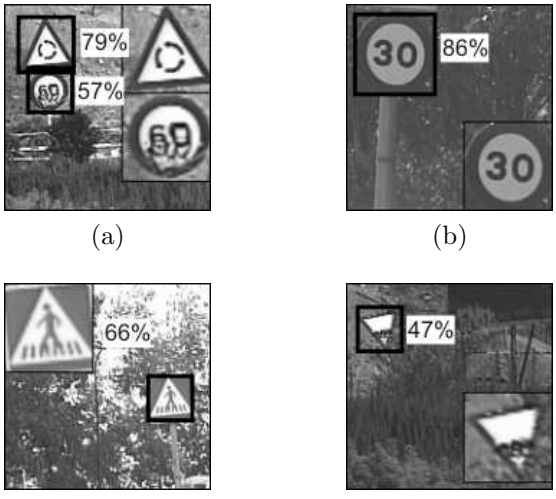
- **Polluted images**: sometimes the in-background is saturated with the



(a) (b)

(c) (d)

Figure 3.42: Examples of solving the problem presented by different occlusions due to the special features of the normalized correlation function. It only requires a certain number of occurrences to detect the object.



(a) (b)

(c) (d)

Figure 3.43: Examples of detection of rotated images and their scores. Example presented in d) where both occlusion and rotation occur in the same image.

surrounding colors so the symbol, especially at longer distances, is also enhanced even though it should be at a low level. Example of this are presented in Figure 3.44a and 3.44b where it is observed that although the symbols seem to be black and white they are reddish and bluish. As a result the normalized correlation value is seen to suffer and the road sign is not detected.

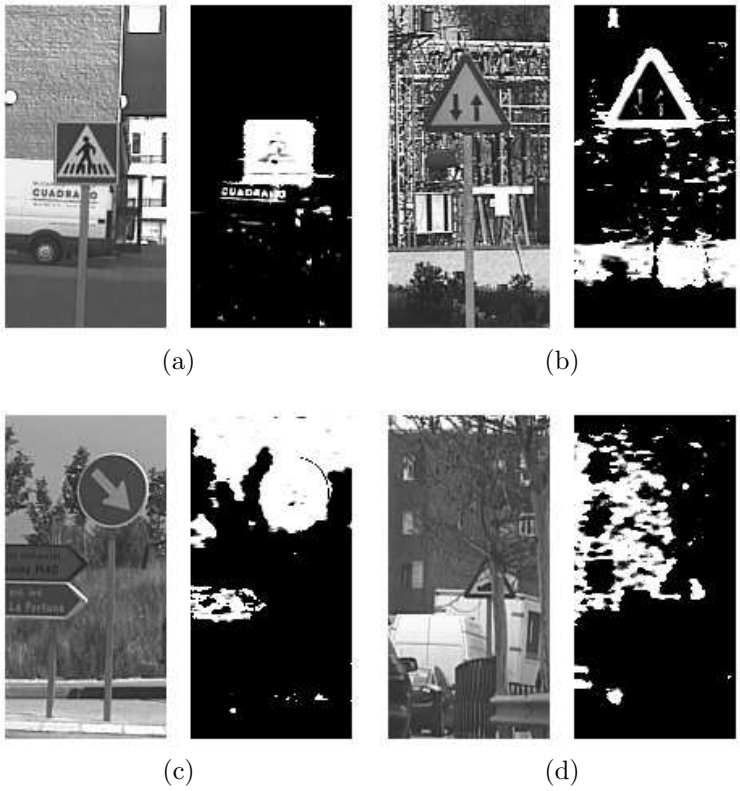


Figure 3.44: Samples of difficult detections when concerned with the enhancement of non desired objects. a) In this case the complete in-background is polluted by the surroundings. b) The symbol of the road sign is assumed to be black but it may be seen that it is also enhanced. c) The in-background is polluted and the out-background is fused almost completely with the road sign. d) Different objects also enhanced may cause false positives due to the similarities to the road signs shapes.

- **Model limitations**: The features that define the models are very im-

portant, the models must be constructed to be as similar as possible to the road sign they represent, however on occasions this causes additional problems. Especially in the case of obligation road signs since a great amount of the area in the plate is blue, sometimes an area with a color which is within the range of the enhancement may be taken as a possible obligation road sign where this leads to a false positive. When combining this effect with that of pollution, detections such as that presented in Figure 3.44a may occur in which a road sign is detected as an obligation road sign instead of a crosswalk.

- **Mixed road sign and out-background**: The road sign and the background may sometimes have the same color as in Fig 3.44c so it is not possible to distinguish the plate and the background, even though the color of the road sign was enhanced correctly the result obtained is then a miss-detection.

- **Similar objects**: Within natural environments there are a wide variety of objects which can be confused with road signs, as may be seen in Figure 3.44d, this depends on their color and the shape where they may be detected producing high scores in detection. This effect is mainly presented for small road signs and fortunately the amount of false detections decreases rapidly as the size of the road sign model used increases.

3.5 Summary

- A study of different color spaces, dependent (RGB) and less dependent (HSL, HCL) on lighting conditions, has been presented. RGB, HSL and HCL have been compared for the two characteristic color groups of road signs, both red and blue. The superiority of the HSL and HCL color spaces when compared to the RGB case to perform improved enhancement of the colors of interest has been shown.

- A study related to color constancy algorithms has been shown to provide useful information related to the color, specially the achromatic information and isolation of specimens which depend on illumination. Its application is mainly based on the task of road sign inventory systems where the execution time is not an important parameter.

- The different components of the ambient light conditions for each color have been studied, this has been carried out in order to find the one which achieves the most satisfactory results when applied not only to its own conditions but also to the conditions of the others. In addition, a special artificially built group condition has been added to investigate its behavior, the main features of this condition is that it contains all the information from all the other conditions. Finally it has been shown that the best option is the artificial grouping of components which has been developed to miss the least amount of road signs in the image as possible.

- A full study of model borders which are to be used along with the normalized correlation has been developed. It has been shown that though the best scores are achieved using models with flexible widths, the best performance is that presented by a binary model. It has also been proved that when used in an appropriate manner, it is possible to take advantage of the features of blurred models, especially in the case where there is a large distance between the camera and the road sign.

- With the aforementioned information a final experiment using the most effective model and the different color spaces has been developed. It has been shown that the best results are achieved when using HCL (H) for red road signs and HSL (H) and HCL (H) for blue road signs. ROC curves and statistical data allow the most efficient configuration to be found and also optimize the cut-off detection value.

Chapter 4

Recognition

In the previous chapter, a method has been presented for the detection of road signs based on two different principles, these were: color and the election of models which represent the shape of the road signs, i.e. color and shape.

In this chapter the information provided by the detection stage will be analyzed, the shape and size of the model required for correlation with the image is obtained. The process is carried out to provide the third important information given by the road sign, the symbol it contains.

The main objective addressed in this chapter is the classification of the candidates provided by the detection module, this will assign the road sign to its model allowing the information to be extracted.

It has already been stated in the state of the art that the performance of a robust recognition system has a double goal:

- The first, to act as a filter for the false detections which come from the previous detection stage, this is done because false detections should not be assigned as road signs.

- Secondly, to classify the road sign in order to use the information described by the road sign symbol. This second objective covers several different requirements, such as: the parameter required to indicate a warning message to the driver, provide autonomous driving and inventory of road signs.

In the work presented here, these two goals have been studied in detail: first of all, the candidates to be road signs will be rejected if they do not fulfill the necessary requirements. Secondly, after the initial filtering, the

classification of the information will be used to warn the driver, if necessary, or will be stored to compile the road sign inventory.

To determine the best solution for the recognition problem, the two most commonly used methods for recognition have been evaluated. The first one uses the same technique as in the detection stage: template matching based on the result of the normalized correlation function. The second, performs a recognition technique based on neural networks (NN). The experiments which have been developed will provide a comparative analysis of both techniques allowing the most satisfactory method for this application to be selected for the final system. Also this study has allowed a direct comparison under the same experimental conditions where, to the best of the authors knowledge, this study has not yet been published.

The chapter is organized as follows:

- First a variety of preprocessing methods for the regions that were stored in the detection stage will be presented. Those are the regions with the highest probability of containing a road sign, where from here on in they will be referred to as candidates. The different preprocesses will be applied to the same set of candidates where different sets of preprocessed candidates from the same basic samples will be obtained. This study will provide vital information on the effects of preprocessing on the final recognition results.

- This is followed by a detail explanation and results on the experimental recognition procedures using template matching.

- The next section will contain information on the NN approach for recognition and the results obtained.

- Finally this chapter ends with a summary on details of the conclusions from both experimental techniques.

4.1 Candidate Preprocessing

The first part of the recognition process is to obtain the candidates from the detection stage. As mentioned throughout this thesis, different models which correspond to different shapes were used to detect the road signs. It was also explained that the points where the normalized correlation function

reaches its maximum value coincides with the center of the models and hence the objects in the image. Knowing this, retrieval of this important data is straightforward. As long as the center of the candidate and the size of the model that is best suited to the correlation of the image are known, it is possible to extract the region in the image where the candidate to be a road sign is and store it. Cropping of the image is performed on a grayscale image as the color information is not required for the recognition process. As a result of using the grayscale image, the region of interest is less dependent on illumination. The conversion from RGB to grayscale is carried out using the well known equation 4.1:

$$GSV = 0.3 \cdot R + 0.6 \cdot G + 0.1 \cdot B \tag{4.1}$$

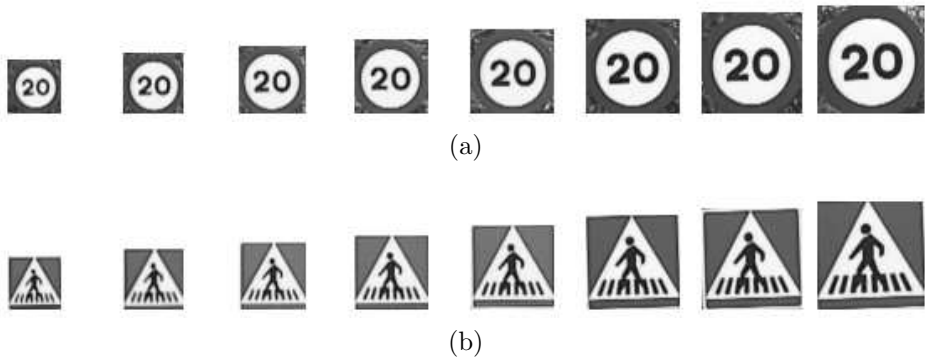
An example of road sign detection and the regions which have been extracted is presented in Figure 4.1.

(a) (b)

Figure 4.1: Example of real detection and cropping of the candidates. a) Grayscale image with the frames corresponding to the candidates detected. b) cropped regions which contains the candidates detected in a)

Here the frame which surrounds the road sign can be seen, also the dimensions of the model used for correlation have been taken into account. The candidates are easily cropped and stored in a separate image buffer for subsequent use (see Figure 4.2). This buffer always has the same dimensions

(64×64) pixels, this is done to simplify both the algorithms and data management. By using this method the initial group of candidates are obtained, shown also in Figure 4.3, where the images to be stored are raw grayscale images. Using these images the rest of the groups will be built using the different post-processing techniques. This is done to determine the models that best fit the template matching and NN.



(a)

(b)

Figure 4.2: Example of real candidates which all belong to the same detection sequence before being resized. These represent the typical models used for detection. The initial and final model set the limits of the range of dimensions used for the search: from 30×30 to 64×64 px. a) prohibition road signs and b) Crosswalk

Once the candidate has been rescaled, different image processing algorithms are applied. These are primarily based on histogram stretching to mitigate the effects of illumination on the image and threshold algorithms to avoid the effects of both illumination and image noise for both the in-background and out-background. The groups defined and their characteristics are as follows:

G1. The result of cropping the grayscale image will provide the first set of candidates, i.e. the raw grayscale images extracted from the detection stage as may be seen in Figure 4.3.

G2. To obtain the second set of candidates histogram stretching is performed for each single candidate histogram. The method is depicted in Figure 4.4 and follows this scheme:

- The histogram of the grayscale image is computed.

(a)

(b)

Figure 4.3: Example of candidates in Figure 4.2 resized to 64×64px. for a) Prohibition road sign and b) Crosswalk

- The histogram covers the lowest value until the amount of pixels for a grayscale level is over an experimental threshold of the total. These values have been chosen to be equal to 1% of the total amount of pixels in the image. The grayscale value GSV' corresponding to this point will be min_H and every grayscale value below it will be set to 0.
- The same process is done starting with the highest histogram value, where the point reached is max_H and the pixels with values greater than this are set to 255.
- Intermediate cases will take the value of the difference between min_H and GSV, these are then normalized to $255 \cdot (max_H - min_H)$.

The mathematical definition of this function is described by the following equation 4.2:

$$GSV' = \begin{cases} 0 & if \quad GSV \leq min_H \\ 255 \cdot \dfrac{GSV - min_H}{max_H - min_H} & if \quad min_H < GSV < max_H \\ 255 & if \quad GSV \geq max_H \end{cases} \tag{4.2}$$

The result of this processing for the previous images can be seen in 4.5.

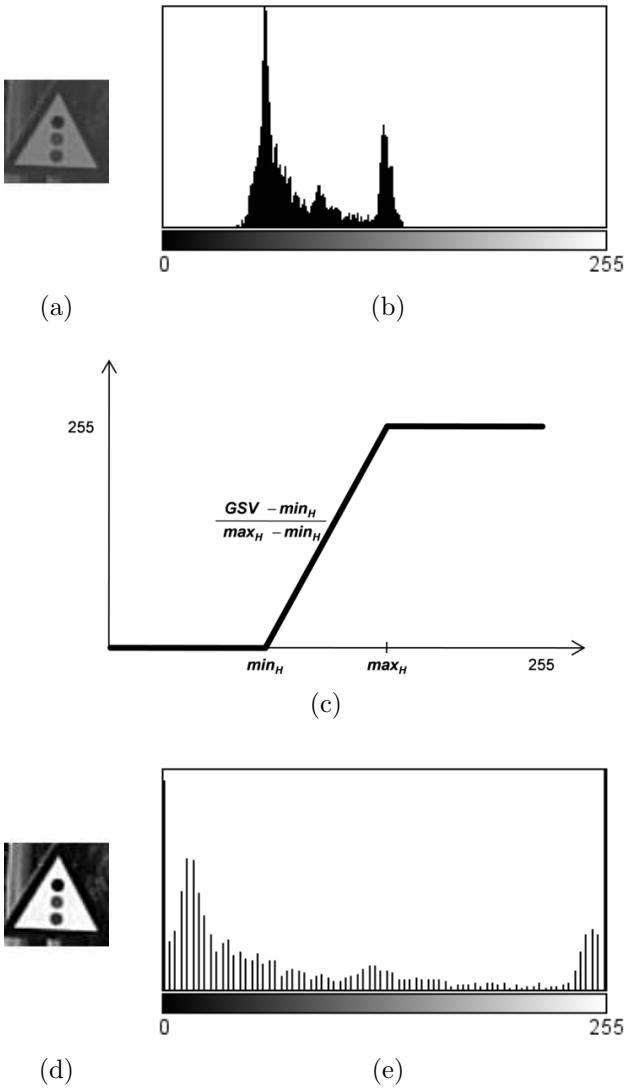
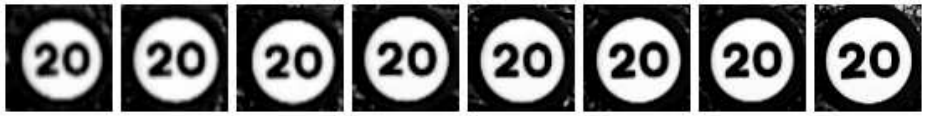


Figure 4.4: Example of the histogram stretching application. a) Grayscale candidate. b) Histogram of the image a). c) Function applied to a). d) Grayscale image resulting from histogram stretching. e) Histogram of image d).



(a)

(b)

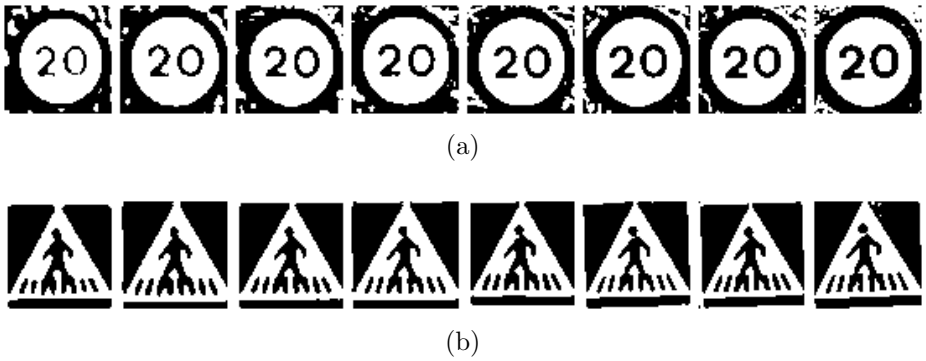
Figure 4.5: Example of candidates in Figure 4.2 subject to histogram stretching. for a) Prohibition road sign and b) Crosswalk

G3. A P-tile conversion [51] has been performed on this group of images. Pictures which represent the standard road signs have been downloaded from the *Instituto Superior de Formación y Recursos en Red para el Profesorado* [12]. For each class all artificial images have first been converted to grayscale images and after evaluated experimentally to set a specific binarization threshold to fit each road sign class. This threshold has been chosen by observing the difference obtained from different threshold levels. An example of this process is presented in Figure 4.6.

Figure 4.6: Examples of the process for optimal thresholding of the P-tile method. From left to right: Grayscale image, thresholding at $GSV = 254$, thresholding for $GSV = 1$, thresholding at $GSV = 90$. It is clearly seen that the threshold level has a noticeable influence on the final binarization results.

Finally there is a percentage range for gray pixels from each class as the symbols do not initially have the same number of black pixels. Fortunately there is not a large variation, this allows a value to be set for each class so that the images are binarized correctly.

The values obtained from the artificial images will be used in the preprocessing of the the real candidates and each threshold level will be applied to each corresponding candidate. It is important to note that from the detection stage, information on the shape of the road sign has already been obtained where this ensures that the threshold is applied to the correct candidate shape. An example of applying the threshold level may be seen in Figure 4.7.

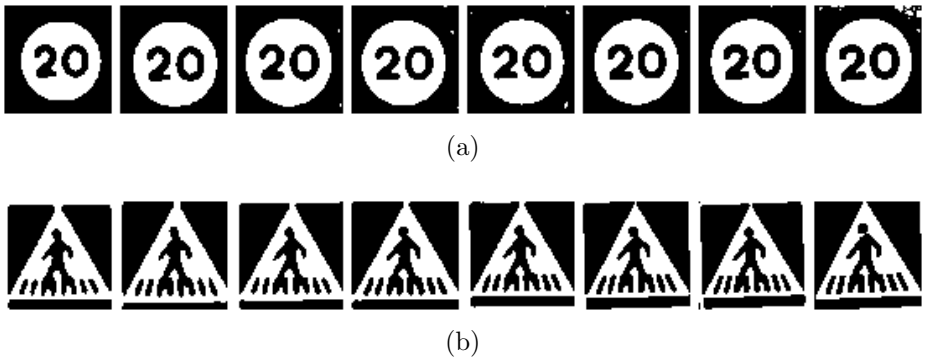


(a)

(b)

Figure 4.7: Example of candidates in Figure 4.2 processed trough P-tile at 15% thresholding for a) Prohibition road sign and b) Crosswalk

G4. The fourth set of candidates depicted in Figure 4.8 is done using an Otsu thresholding, where there is a detailed explanation of this technique in [51].



(a)

(b)

Figure 4.8: Example of candidates in Figure 4.2 processed using Otsu thresholding for a) Prohibition road sign and b) Crossalk

G5. A bimodal binarization is used to build the 5th set of images. In this case the histogram is evaluated in the following way: first of all, a search is performed to find the highest peak this is then followed by a search for the second highest peak which is separated by a certain number of grayscale levels, where this value has been obtained experimentally and is set to be greater than 100 so that the symbol and in-background has sharp contrast level. This value (100) was selected as it was considered to be adequate for all signs under real conditions. The threshold is set to the halfway point between these two peaks. Figure 4.10 shows an example of the conversion following the scheme depicted in Figure 4.9.

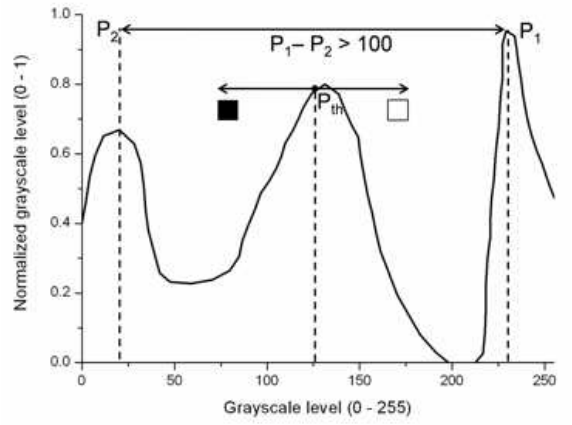


Figure 4.9: Scheme of the bimodal binarization. The highest peak is taken as the reference, then the peak which is separated by more than 100 grayscale levels is taken as the second peak. The mean value of these is the threshold level which separates black and white.

Independently of the method used, template matching or NN, the final processing of candidates has been developed by applying two different masks, M1 and M2.

– M1 is built from the information provided by the detection stage regarding the class of the road sign located. As the candidate is normalized to 64×64 before recognition, only the shape information is required. Applying masks which represent different shapes of road signs is a straightforward process and separates the road sign from the

(a)

(b)

Figure 4.10: Example of candidates in Figure 4.2 processed using Bimodal thresholding for a) Prohibition road sign and b) Crosswalk

background. The masks which were used in this thesis are presented in Figure 4.11a.

– M2 is based on the extraction and binarization of the rim from the enhanced grayscale image obtained from the detection stage (Fig. 4.11b). Since the location and size of the road sign in the image is known, it is possible to extract the rim from the converted image which is applied as a mask over the candidate.

Masks are applied in different ways depending on the method used for recognition, these are:

– Template matching: The M1 mask must be applied to the candidates to set the out-background to a do-not-care value. By doing this the influence of the out-background on the result of the normalized correlation is avoided, this will be described in detail later. An example of this process may be seen in Figure 4.11a.

– Neural networks: all the information that is located outside of the symbol is considered to be noise. To avoid the noise due to the in-background and out-background the two masks M1 and M2 are applied as shown in Figure 4.13. The first, M2 from the cropped grayscale converted image erases the rim of the road sign. This mask is binarized by setting all pixels other than 0 high. The second mask, M1 is similar to those presented in 4.11a which were used for template matching,

(a)

(b)

Figure 4.11: Example of masks and rims used in the preprocessing stage: M1 and M2. a) M1: masks for different shapes, from left to right: Danger, Crosswalk, Yield, Obligation. b) M2: rims extracted from the converted grayscale image, from left to right: Danger, Crosswalk, Yield and both Prohibition and Obligation. Obligation road sign have no rim as they only have in-background

Figure 4.12: Example of applying mask over a Danger road sign. The grayscale candidate may be seen along with the mask in the center and finally the result of the mask over the image in which the out-background has been set to do-not-care values.

in this case the do-not-care values are substituted for high level values, 255, this completely isolates the symbol from the background. Finally, and with knowledge of the shape of the road sign obtained from the detection stage, it is possible to extract the specimens which will be used as the input to the NN. The dimensions of the extracted regions will depend on the shape of the road sign. This is because the location of the symbol is different for each type of sign. The dimensions of the regions are defined to cover every case of symbol, this is due to the large variation of their size.

(a) (b) (c) (d) (e) (f) (g)

Figure 4.13: Example of candidate preprocessing for NN. a) Grayscale image. b) Image of the road sign rim cropped from the picture resulting from the conversion of the given color space (see Chapter 3 on Detection). c) Binarization of image in b). d) Result of the application of c) over a). e) Mask to erase the out-background. f) result of the application of e) over d). g) Final result using Otsu thresholding on image f).

4.2 Recognition using Template Matching

The template matching technique for recognition follows the same scheme as that presented for the detection stage. The only difference is that the models used in detection will be the recognition candidates and the converted image where the matching which occurs in detection will be the templates for the class of the road sign in the recognition stage. 135 road signs may be recognized using this technique, however, new road signs may be added at any time by drawing the picture (i.e template). The template matching method is highly flexible and re-sizable.

Two examples of templates may be seen in Figure 4.15 where it may be observed that there are more differences than those related to the class of the road sign:

– Figure 4.15a and 4.15d represent the templates that are used for the groups **G1** and **G2**. In this case the out-background of the template is set to a low level since in some cases there may be a space between the mask and plate which belongs to the out-background. An example of this is presented in Figure 4.14. As a result the influence of the misalignment between the road sign and the background will have no impact on the correlation results. It may also be observed that the symbol color is slightly different to the rim, this has been done to take into account the different grayscale levels which the rim and symbol have, in general the rim has a higher GSV value than the symbol.

Figure 4.14: Example of a road sign with a space between the out-background and the plate. This space is usually different to the rim of the road sign since it is generally different to the enhanced color.

– In the second case, Figure 4.15b, the template has been binarized to fit the groups **G3**, **G4** and **G5**. The out-background has been set to a high level and the the in-background, rim and symbol have been set to a low value. The objective here being the same, i.e. represent a similar profile of the road road sign candidate and template.

4.2.1 Experimental Outline

The different groups in which the candidates are divided into are the result of the final detection experiment described in Chapter 3. Once the candidates have been cropped, they are converted using the different methods explained previously, 5 sets of candidates are produced **G1** - **G5**.

The basis of the experiments has been to obtain the value of the normalized correlation of the candidates over the templates, the best result is the output of the recognition stage. If the result is below a preset threshold, 20% in the experiment, the candidate will not be considered as a road sign and hence rejected. If the result is over the given threshold, the candidate will be

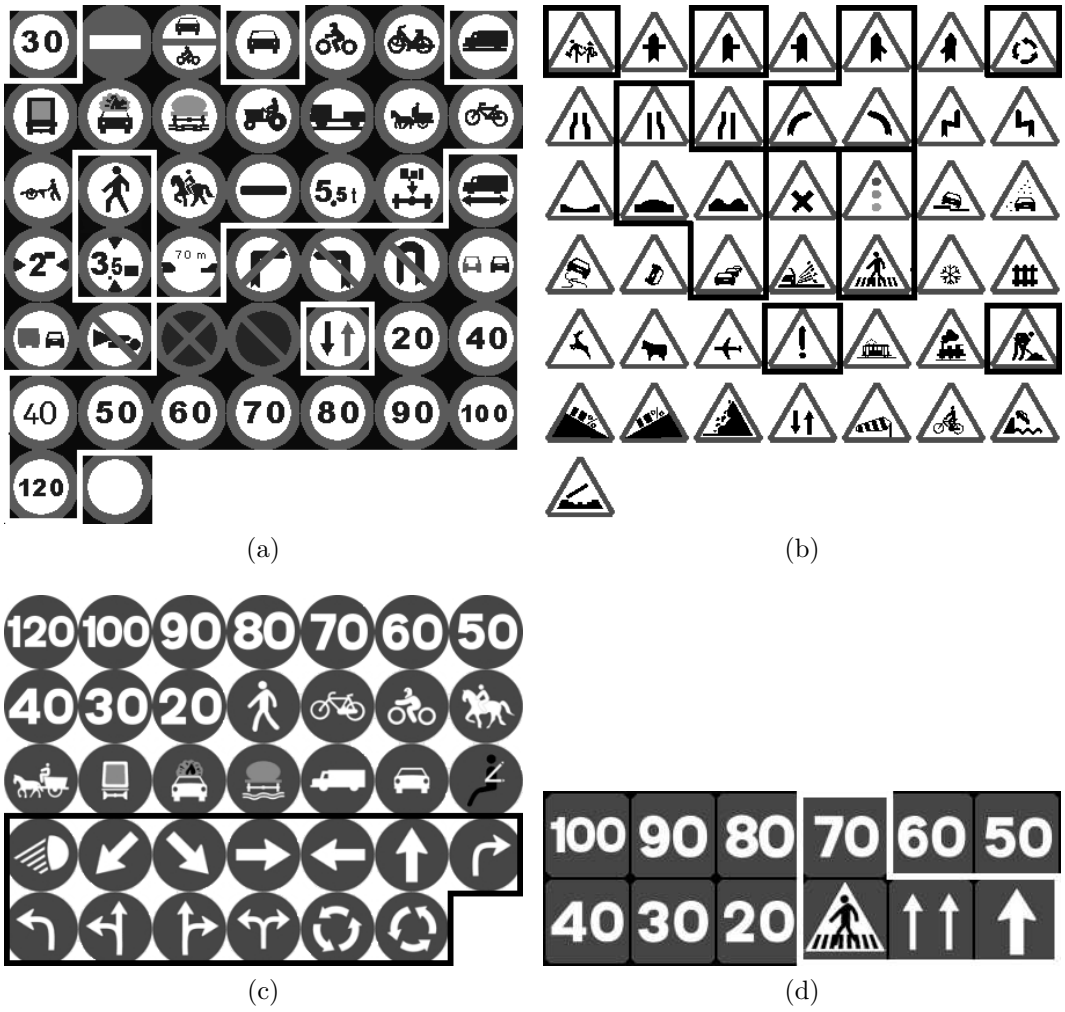


Figure 4.15: [(dibujo y pie de pagina)Example of templates for recognition to correlate with candidates. a) Prohibition road signs. b) Danger road signs. c) Obligation road signs. d) Indication road signs. The road signs which are framed correspond to the sets for NNs.

considered as a road sign and the information will be stored for subsequent comparison with the database of annotated images, this procedure was also carried out in the detection experiments, here the final result on reliability of the recognition is provided.

In recognition, the definition of the FP_{rate}, will change from the definition which was presented in Chapter 3, to the total number of both negatives incorrectly classified and incorrect recognitions over the total amount of detections. In addition, the information of TN_{rate} is also provided and indicates the rejection rate over the total amount of negatives.

4.2.2 Results

Red Road Signs

In Table 4.1 and Figure 4.16 the recognition results using the template matching technique of red candidates are presented. It may be seen that the best performance corresponds to the raw grayscale candidates, this is followed by the histogram stretching and finally the Otsu thresholding which is also observed to obtain satisfactory results, 6% and 9% less than previous case. In general, the option for binarization of candidates is not effective for template matching. The normalized correlation severely punishes the difference of the pixel values and it seems clear that grayscale models are the best choice for template matching.

In general the problems associated with template matching for recognition are related to the FP_{rate} which is very high for grayscale candidates and the other preprocesses. Also, from Table 4.2 it may be observed that it is not possible to reduce the number of false positives and maintain a high level of true recognitions. As a result, no filtering may be done using this technique. These disadvantages give rise to two problems:

- Results on the false and true positives are low, thus the main objective here would be to find a method to increase the results for true positives while at the same time reduce the false positive results.

- The false recognitions are also composed of the incorrect recognitions, one important area which should be considered here is to find a more robust method to convert these incorrect recognitions into true positives.

The ROC curves corroborate the aforementioned results from the experimental trials and also indicate the deficiency of the results obtained from the bimodal algorithm. This is due to the fact that it is not possible to set a constant separation between peaks as the histograms vary dramatically with the illumination conditions of the road sign i.e. exposed to sunlight or shadowing.

Red					
Group	Processing	TP_{rate}	FP_{rate}	TN_{rate}	area
G1	Raw grayscale	0.83	0.25	0.27	0.86
G2	Histogram stretching	0.77	0.22	0.29	0.79
G3	P-tile	0.75	0.34	0.36	0.69
G4	Otsu	0.76	0.34	0.19	0.77
G5	Bimodal	0.27	0.10	0.88	0.66

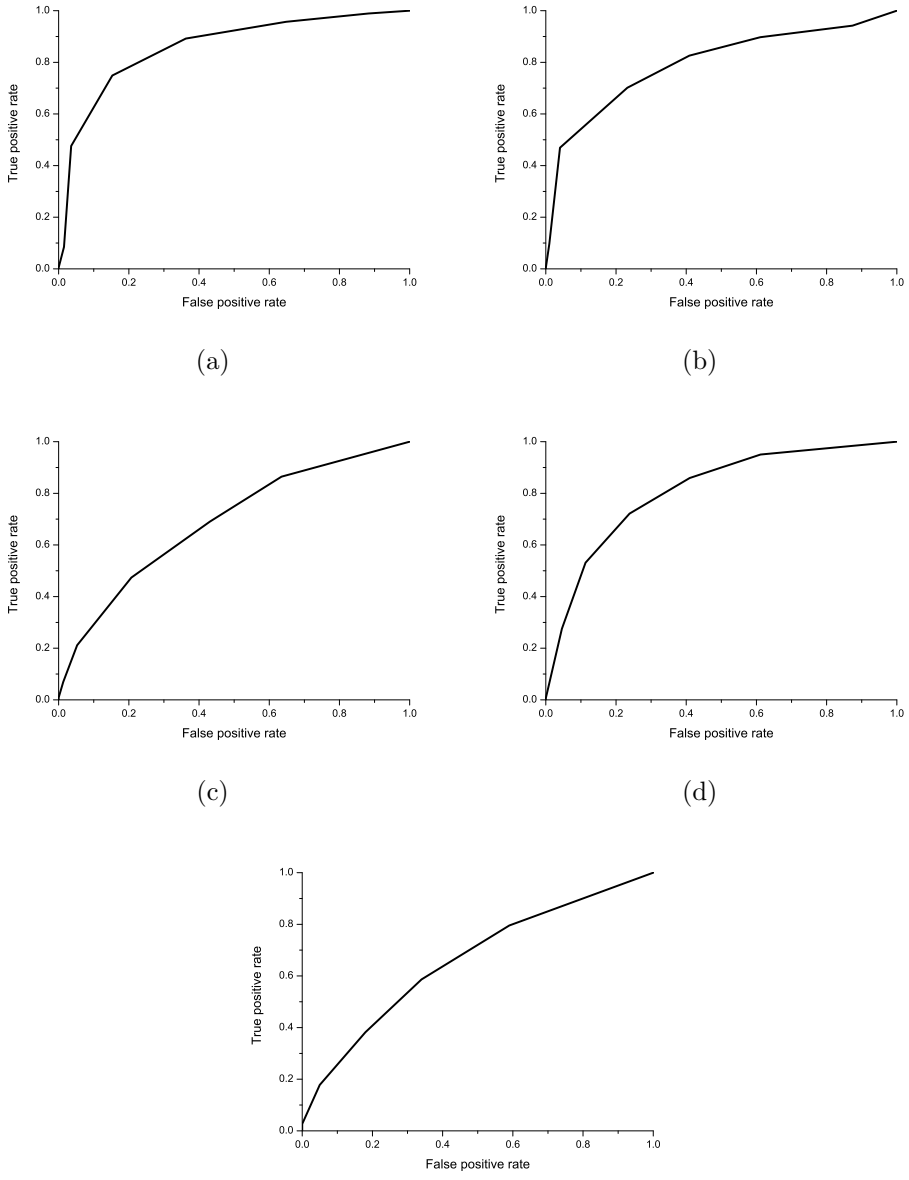
Table 4.1: Results on recognition of red road signs regarding the processing applied to the candidate. The true positive rate (TP_{rate}) takes into account the rate of recognitions over the total amount of true detections. The false positive rate (FP_{rate}) is the rate of false recognitions and incorrect recognitions over the total amount of detections. (TN_{rate}) considers the rejection rate of candidates over the total amount of false detections. Area stands for the area below the ROC curve normalized to 1.

Processing	FP_{rate}	TP_{rate}
Raw grayscale	100	58
Histogram stretching	27	50
P-tile	84	65
Otsu	71	58

Table 4.2: % of true and false positives in the first interval $[20 - 30]\%$ for the best processing of red road signs. It can be seen that it is not possible to reduce the false detections without affecting the true recognitions.

Blue road signs

The results for recognition using template matching of blue road signs are presented in Table 4.3 and in Figure 4.17. In the case of blue road signs the



(a) (b)

(c) (d)

(e)

Figure 4.16: ROC curves which correspond to the recognition results for red road signs using template matching for each of the preprocesses. a) Grayscale. b) Histogram stretching. c) P-tile. d) Otsu. e) Bimodal.

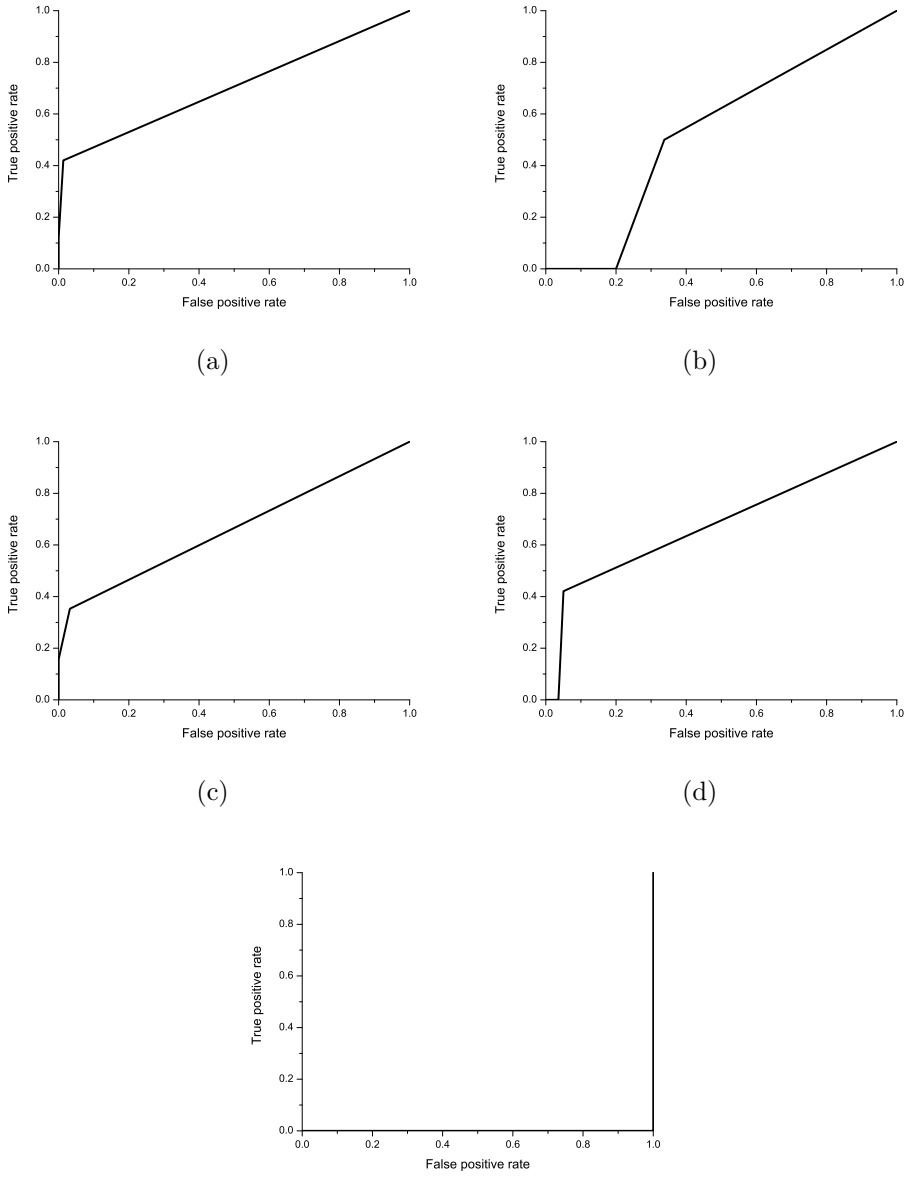
results would not be valid for a recognition system as the true positive rate is very low. The true negative rate is very high, this is the optimum result for filtering of false specimens, however this is due to deficient binarization and the majority of the incoming candidates are rejected as road signs.

Blue					
Group	Processing	TP_{rate}	FP_{rate}	TN_{rate}	area
G1	Raw grayscale	0.45	0.44	0.47	0.70
G2	Histogram stretching	0.29	0.16	0.81	0.53
G3	P-tile	0.42	0.30	0.65	0.66
G4	Otsu	0.31	0.22	0.75	0.68
G5	Bimodal	0.03	0.00	1.00	0.00

Table 4.3: Results on recognition of blue road signs regarding the processing applied to the candidates. The true positive rate (TP_{rate}) takes into account the rate of recognitions over the total amount of true detections. The false positive rate (FP_{rate}) is the rate of false recognitions and incorrect recognitions over the total amount of detections. (TN_{rate}) considers the rate of candidate rejections over the total amount of false detections. Area stands for the area below the ROC curve normalized to 1.

The ROC curves show poor results when considering the most important parameters of the recognition process (TP_{rate},FP_{rate}). In general, the TP_{rate} should be much higher to avoid false negatives; the FP_{rate} should be lower to avoid false warnings being sent to the driver, especially in the case of incorrect recognitions which are more likely to confuse rather than assist the driver. Finally, the TN_{rate} should be high as a result of correct rejections by the system as opposed to system failures.

As has been pointed out in chapter 3, the normalized correlation punishes both the difference of grayscale levels and the inhomogeneity of the road sign. In general the blue road signs are easily missed, this is especially true for several cases such as the obligation road signs where there is a lot more in-background area than symbol area. Also, since the size of the symbol is small, any slight damage can easily lead to an incorrect recognition. Another cause of false negatives, when considering crosswalk road signs which contain white stripes, is due to slight movements or rotations of the road sign. These account for the large amount of false negatives.



(a) (b)

(c) (d)

(e)

Figure 4.17: ROC curves corresponding to the results for recognition of blue road signs using template matching for each of the preprocesses. a) Grayscale. b) Histogram stretching. c) P-tile. d) Otsu. e) Bimodal.

4.3 Recognition using Neural Networks

As has already been presented in the state of the art, NNs have been the most widespread and popular technique for road sign recognition even in the presence of other more straightforward classification techniques. In the work presented in this thesis NNs are considered as an alternative method for recognition using template matching.

However, the use of NNs imply certain disadvantages when used for general recognition and especially road sign recognition applications. These are:

1. Difficulty in generalizing results. After a network has been trained and validated the results cannot be extended to different sets of images, i.e. different sets of images may provide different results. The only way to allow a certain level of generalization of the results is to make the validation set as big as possible and to consider a large variety of cases. It is not possible to prove that the recognition results for a different set of images are always going to be the same. The only thing that can be done here is to enlarge the validation set as much as possible so that the likelihood of obtaining a incorrect classification from subsequent tests is very low.

2. Understanding and knowledge of the recognition process which the NN performs is complex, i.e. the features being considered and the reason for considering them, there is no control of the input parameters although after several trials the operator develops a certain level of intuition.

3. Adding new samples to the NN is not a straightforward process. The addition of only one object for the recognition process requires the network to be retrained and revalidated.

4. Finding a suitable amount of samples to effectively train the network and to perform the validation is difficult, this is especially true when considering that the validation set must be large.

5. The image sets used to train and validate the NN are obtained manually, the operator must obtain a large and varied set of samples, otherwise the NN may become rigid and the final results obtained would not reflect the true flexibility and ability of a NN system.

Points 1 and 2 are easily solved, however points 3, 4 and 5 require special attention to reduce their impact on the final system.

Previously mentioned in section 2.4 in the State of the Art have been several attempts to overcome the obstacles associated with the training process when concerned with the search for samples by using artificial specimens or a mixture of artificial noise added to real images. Also other researchers have used an ART net to solve the previously stated point 4, however the main disadvantage with this net is over-fitting. The initial advantage of the continuous training becomes a disadvantage, as the routes the vehicle takes are generally the same, the road signs during the route are also the same and hence the net will learn some specimens effectively (the most commonly encountered) however the system will not learn new road signs as successfully.

For these reasons, a common Multilayer Perception Neural Network (MLP NN) has been chosen for the recognition process where emphasis has been placed on the processing of the specimen required to train the NN.

The initial research of this thesis on NN's was developed in the Vislab[21] where the NN used was a LWN++ [57] this is an open source implementation which uses a general feed forward NN with backpropagation training. The neurons of each layer compute synchronously the weighted mean of the outputs of the previous layer and then apply the activation function to the result which is given by equation 4.3. Only the logistic part of the activation function has been used providing results between 0 and 1.

$$f(x) = \frac{1}{1 + e^{-x}} \tag{4.3}$$

The input samples are 64×64 pixel specimens which directly enter a specific NN and depend on the shape detected. To breakdown the recognition problem it is possible to have smaller NN instead of using just one for the complete set of road signs. Five different NNs have been developed which correspond to Danger, Yield, Prohibition, Obligation and Indication road sings. A hidden layer has also been added where the number of neurons is modified to obtain the best performance of the net.

The complete set of different road signs is very large and also several road signs are limited to specific areas and located in very few places, for example, the road signs indicated in Figure 4.18 are not very common. For this reason, the problem indicated in point 4 arises, to find real specimens of road signs, and moreover, to find a large enough set of different real road signs to train the system and to validate the results is not a straightforward

task. The total number of road signs which can be recognized using a NN is 54, see figure 4.15.

Figure 4.18: Example uncommon road signs. There is a great difficulty not only in finding them but to acquire enough different samples for training and testing the NN.

This factor not only affects the uncommon road signs but the complete set of them. For this reason, regarding the inputs of the NN, the training set was reduced by providing the net with the most important information, thus reducing the noise and simplifying the training process. This process is similar to that of the application of a principal components analysis which reduces the information to the most important. In the current application the final objective is the same, however the preprocessing stage is assigned the task of preserving the most important information.

The results of the different preprocesses on a set of images may be seen in Figure 4.2, 4.5, 4.7, 4.8 and 4.10. The large amount of images presented throughout this chapter, which correspond to different size road signs, have been resized to 64×64. It may be observed from this resize process that the final results are not affected and that the images are very similar. Thus, if the specimens which introduce noise naturally, i.e. grayscale and histogram stretching are neglected, what remains are the the sets of binarized candidates.

It has been observed from experimental trials that, contrary to the expected behavior, no important information is lost when the noise is removed. The data which remains is more than sufficient for the system and almost the same when referring to the different distances from the camera to the road sign for the range required by the proposed system.

This procedure reduces the effects of the problems which have been discussed in point 4 as the training set is considerably reduced, this reduction also solves the problems associated with point 5 as the election of different and useful samples is much more straightforward and avoids redundancy.

4.3.1 Experimental Outline

Initially a first approach to obtain the best configuration for the final network is performed. Three different preprocessing stages for binarization have been applied (Bimodal, Otsu, P-tile) to each of the networks which correspond to the shape of the road sign (Yield, Prohibition, Danger, Indication, Obligation). Also for each one of these cases, two different errors will be tested which will stop the training. The rate of false positives over the complete set has been calculated to be 2.5% and 5%.

These errors have been set manually after a large number of trials and correspond to a successful method to avoid overfitting. The NN is trained with only one set where the training error has been obtained, while at the same time the NN is being validated each time for different sets where a validation error is obtained. As a result the optimum number of times the algorithm is executed is that which provides the minimum error in the validation set, as this last one is the most effective test of the NN over a set of new samples. It is important to note that the error in training of the backpropagation NN decrease continuously reproducing each time, using an improved method, the patterns used during training, so stopping in a number of epoches or at a given minimum may not lead to an optimal network. An example of this may be seen in figure 4.19.

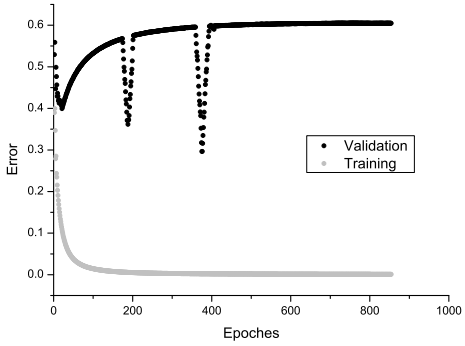


Figure 4.19: Training example. It may be seen that though the training error low, the validation is still high. It is also important to note that since there are local minimums in the validation curve, special attention must be paid to defining the point where the training stops.

Two cut-off threshold values have been used, the first, 2.5% is a satisfac-

tory result to avoid over fitting, and the second 5% is used to verify that a certain error value provides less effective results. For each of these sets the number of neurons in the hidden layer has been varied from 1 to 2000 to find the best performance. The number of samples for training and validation are presented in table 4.4.

Class	Training	Validation
Yield	12	191
Prohibition	208	546
Danger	188	959
Obligation	182	476
Indication	33	93

Table 4.4: Training and validation a number of samples for the NN approach.

This type of training requires NN error certification, this is due to the fact that even though the NN is always trained with the same set, the training set stops when the validation error is below the set threshold. As a result it appears that different validation sets may lead to different results. This certification error will be produced from a certification set which is that of the annotated images used in chapter 3.

4.3.2 Single-Stage Neural Networks

First, a Single-Stage Neural Network (SSNN) has been trained, this implies that for a given input the final recognition is performed from only one network. This approach has been developed for each one of the possible road sign shapes. Use has been made of the information on the shape provided by the output of the detection stage. By using different networks for each shape the problem is reduced from the recognition of one road sign from a large group to sorting one road sign from five different smaller groups, this reduces the size of the networks and the number of outputs required for a net, where this leads to a more straightforward classification.

The outputs depicted in Figure 4.20 correspond to the different classes: Yield, Prohibition and Danger for red road signs, and Obligation and Indication for blue road signs.

The results of the validation are presented in Table 4.5. In this table it may been seen that the best results are obtained from the Otsu and P-tile

Figure 4.20: Examples of the output for the SSNN for road sign recognition. From left to right: Yield, Prohibition, Danger, Obligation and Indication classes.

Preprocessing	Class	2.5%		5%	
		N	TP_{rate}	N	TP_{rate}
Bimodal	Yield	622	0.91	817	0.84
	Prohibition	1937	0.89	1887	0.86
	Danger	1337	0.91	1022	0.87
	Obligation	812	0.98	1337	0.85
	Indication	622	0.65	962	0.48
Otsu	Yield	287	1.00	482	0.98
	Prohibition	1737	0.99	1232	0.97
	Danger	1637	0.99	1207	0.93
	Obligation	657	0.98	1222	0.95
	Indication	312	0.93	642	0.90
P-tile	Yield	242	1.00	412	1.00
	Prohibition	1272	1.00	1182	0.97
	Danger	1232	0.99	1147	0.97
	Obligation	657	0.99	1207	0.95
	Indication	282	1.00	622	0.97

Table 4.5: Initial results of the preprocessing, class of road sign and error allowed. TP_{rate} stands for the positive rate over the complete set of specimens. N is the number of neurons in the hidden layer.

methods. It may also be observed from this table that the results obtained using the bimodal are not as effective as those from the aforementioned methods, the reason for this lack of accuracy has been explained previously in subsection 4.2.2. The initial setting of the distance between peaks which depends on the illumination conditions may not be fulfilled, thus reducing the recognition capabilities of the system. When compared to the template matching technique the results obtained from the neural network are seen to be a lot more effective.

From the experimental trials it has been observed that there is no correlation between the number of neurons in the hidden layer and the percentage of error allowed, this has also been seen for the preprocessing, for example, for yield road signs the number of neurons for bimodal preprocessing is more than twice that of the Otsu or P-tile methods, 622, 287 and 242 respectively, while for Danger road signs, the P-tile method requires 10% less neurons and Otsu 20% more neurons in the hidden layer when compared to the bimodal process.

Results for both P-tile and Otsu methods are seen to be satisfactory, especially when considering the case where the error is 2.5%, here the best results have been obtained. The recognition process is affected as a consequence of using low error, this is because the NN is overtrained, i.e. the samples used for training are well understood by the network, however the system is not efficient for specimens other than those used for this training process. The best performance is achieved when a certain amount of error is allowed, this preserves the flexibility of the net avoiding missclassifications.

Though in the case of P-tile previous assumptions have to be stated, i.e. the % of gray pixels used for each shape to distinguish the symbol from the background, good results have been achieved here. Initially, it was thought that the best outcomes would be those obtained from the Otsu binarization as it is the only algorithm, which has been tested, that does not require any setting, i.e. it works automatically without any need for manual setting of parameters as opposed to the peak distance required for the bimodal method and the % of gray pixels for the P-tile method. This leads to a generalization of the problem which is always expected in research modeling. But contrary to this first assumption, the P-tile method has achieved the best validation results which are similar to those obtained from the Otsu binarization method where for a 2.5% error the P-tile model also achieves satisfactory results for the rate of true recognitions.

After the initial results were obtained from the experimental trials, de-

scribed throughout this chapter, the remaining tests which correspond to the SSNN and cascade NN have been carried out. From here on only the Otsu and P-tile preprocessing for an error of 2.5% will be considered.

The certification set has provided the results presented in table 4.6 and the corresponding ROC curves presented in figure 4.21.

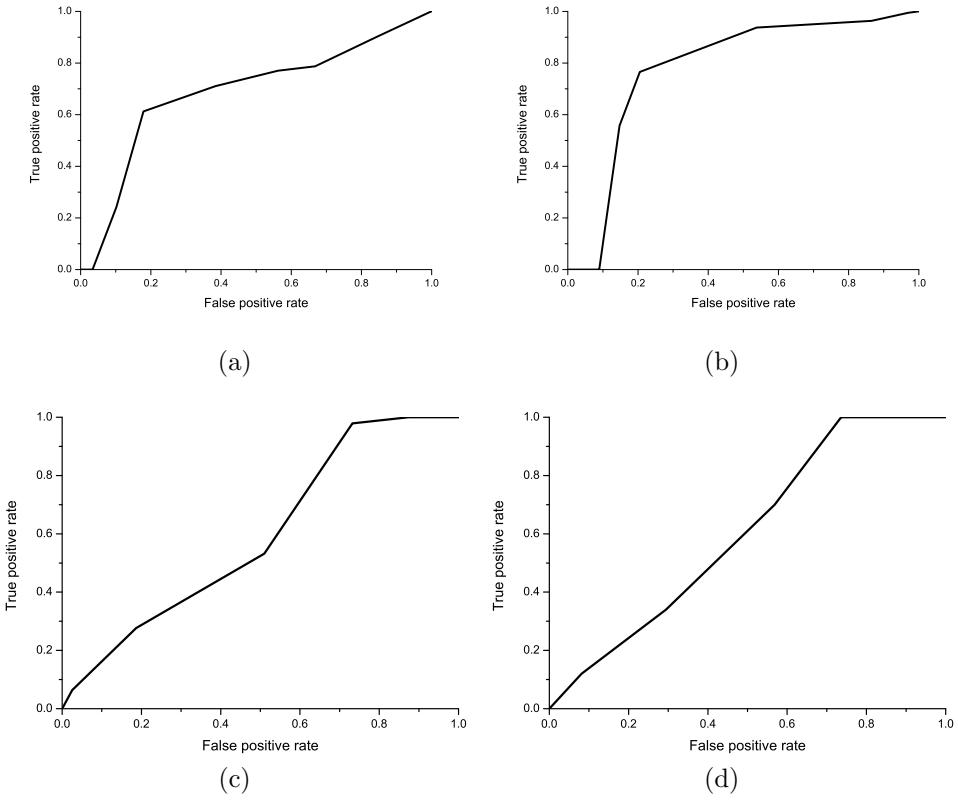


Figure 4.21: ROC curves corresponding to the recognition results for red and blue road signs using the Otsu and P-tile binarization from the single-stage approach. a) Red road signs, P-tile binarization. b) Red road signs, Otsu binarization. c) Blue road signs, P-tile binarization. d) Blue road signs, Otsu binarization.

Preprocessing	Class	TP_{rate}	FP_{rate}	Area
Otsu	Red	0.23	0.81	0.78
	Blue	0.48	0.80	0.60
P-tile	Red	0.28	0.77	0.68
	Blue	0.47	0.81	0.59

Table 4.6: Results of the SSNN approach for recognition. TP_{rate} stands for the positive rate over the total amount of specimens. FP_{rate} stands for the incorrect recognition rate over the total amount of detections.

4.3.3 Cascade Neural Networks

The cascade approach simplifies the problem of recognition by arranging smaller groups of road signs where confusion among the symbols occurred using the single stage approach. This results in more specialized subnetworks and so it is expected that the number of incorrect recognitions will decrease. The cascade approach has been performed using the data obtained from the single-stage network when different road signs were misclassified. Using these groups two networks have been developed for each group where in the first case the candidates are taken as the inputs of the network and in the second case small windows designed for each case are taken as the inputs of the network. The most common confusions observed are described below and are also presented in Figure 4.22.

- **Yield**. **Y1**: in this case, since there is only one output it is not necessary to use the cascade approach, the network is exactly the same as in the single-stage case.
- **Prohibition**. **P1**: Speed limit road signs. **P2**: No trucks. **P3**: Parking, Stop and Driving.
- **Danger**. **D1**: Merging. **D3**: Uneven road, Hump. **D4**: Other dangers, traffic lights.
- **Obligation**. **O1**: Roundabout.
- **Indication**. **I1**: in this case no cascade approach has been developed as the most important indication road signs are those of speed and Crosswalk.

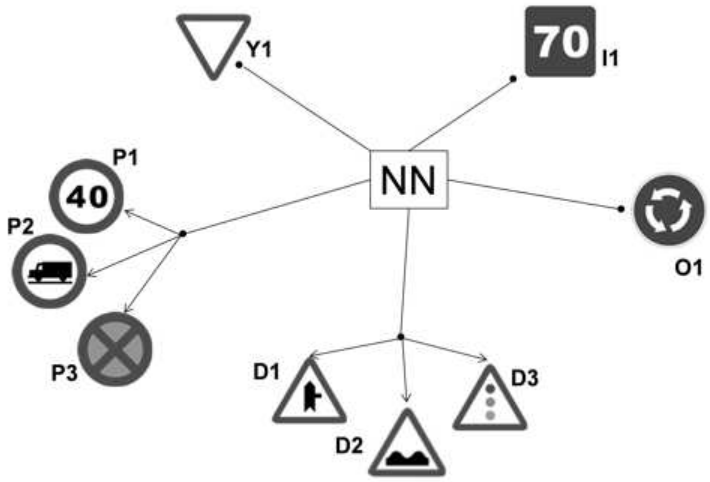


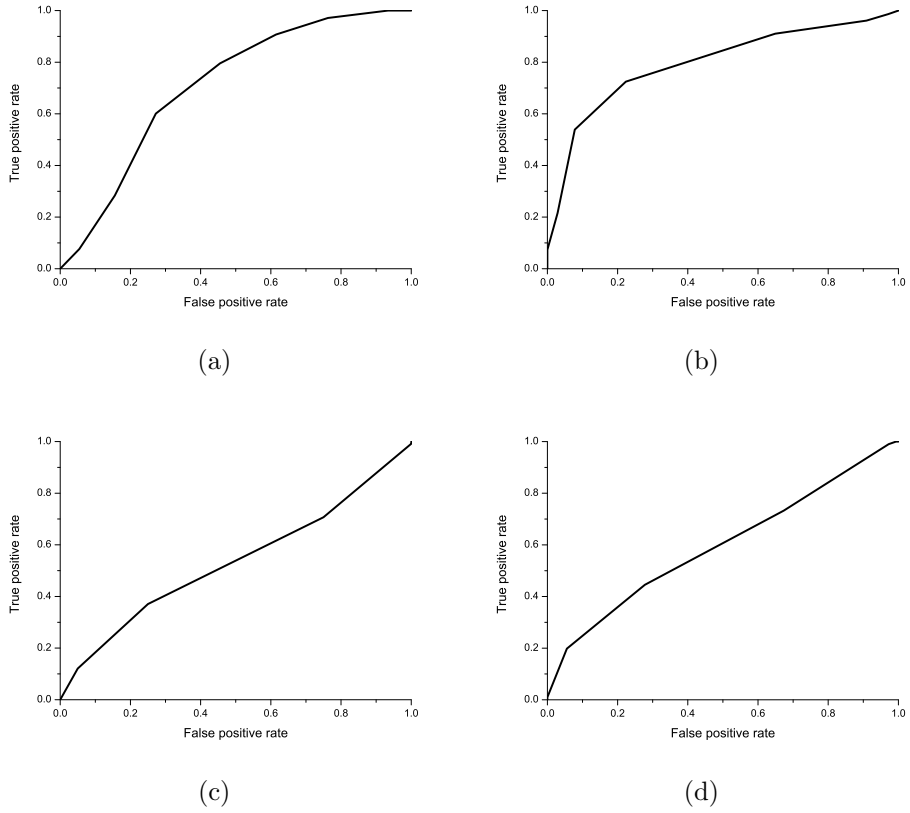
Figure 4.22: Examples of the output of the cascade network for road sign recognition. Each figure shows the group it belongs to.

As previously stated, in this case the candidates are directly sent to the second network for recognition. The results of this experiment may be found in table 4.7 along with the corresponding ROC curves in Figure 4.23. It may be observed that the results for blue road signs are highly satisfactory where the true positive rate is 1 for both cases. Unfortunately, it has not been possible to reduce the amount of false positives, this is mainly due to classification of negative candidates. This may also be seen from the ROC curves which show, especially in the blue road sign case and P-tile method, that the line is close to the diagonal which indicates that the probability of true and false positives cannot be predicted. When compared to the template matching technique the results obtained are seen to be more effective when using NNs.

For red road signs the results are also satisfactory, again more accurate than those obtained from the template matching method, this is especially true for the P-tile binarization case. The Otsu is seen to fail, this is primarily due to the samples which are under shadow conditions,where the algorithm is capable of setting the threshold.

Preprocessing	Class	TP_{rate}	FP_{rate}	Area
Otsu	Red	0.52	0.56	0.79
	Blue	1.00	0.59	0.60
P-tile	Red	0.87	0.26	0.71
	Blue	1.00	0.26	0.53

Table 4.7: Results of the cascade approach for recognition. TP_{rate} stands for the positive rate over the total amount of specimens. FP_{rate} stands for the incorrect recognition rate over the total amount of detections. Area is the value of the **integer** under the curve ROC.



(a) (b)

(c) (d)

Figure 4.23: ROC curves corresponding to the recognition results for red and blue road signs using the Otsu and P-tile binarization in a cascade approach. a) Red road signs, P-tile binarization. b) Red road signs, Otsu binarization. c) Blue road signs, P-tile binarization. d) Blue road signs, Otsu binarization.

4.4 Conclusions on Recognition

In this chapter two approaches to the recognition process have been described, these were template matching and NNs. The principal conclusion observed from experimental trials is that the NN recognition process demonstrates better performance when compared to template matching. Both advantages and disadvantages exist for each technique, this is due to their inherent properties. For example, template matching provides less effective results and less true recognitions; the recognition of blue road signs is very difficult due to the large similarity among the road signs. However the main advantage to this technique is that new road signs may easily be added to the system by simply drawing another template.

Neural networks achieve very good recognition results when the recognition problem is divided into smaller groups. It has been shown that results obtained from a single-stage or cascade networks are different. It has not been possible to find a correlation between the number of neurons in the hidden layer and the improvement of results, it does not seem to depend on the number of samples or the error allowed thus the behavior of a NN in function of the number of neurons in the hidden layer cannot be predicted.

Some disadvantages arise when NN are used, for example, the missclassification problem. Blue road signs belong again to the group which achieve the least accurate results, this is due to the similarities between the different road signs. Another important error is the confusion of the system between indication and obligation road signs. This causes false detections which still have not been adequately avoided. There is also difficulty in defining the correct amount of samples to train the network, validate and certify the results. The addition of just one road sign requires that the system be completely trained and optimized. Finally, the execution time required for the recognition process to recognize road signs is long, where this time is further increased when the recognition is based on two stages.

In conclusion the use of a hybrid system is advised which would be based on both template matching for red road signs and NN for blue road signs, at least until methods are devised to reduce the execution time of the NN recognition process, for example, by reducing the size of the samples. In the case of inventory road sign systems the execution time is not an important issue allowing the addition of more stages to the neural network to achieve better recognition results. It is also possible in this case to increase the size of the specimens providing more accurate information as the time required

to carry out the process is not important as it is an off-line application. Also for both applications, speed supervisor and automatic road sign inventory, it is necessary to find enough samples to train or to reduce the set of possible recognized road signs.

4.5 Summary

- A study of different preprocessing methods has been carried out, this has been based on two approaches, grayscale images and binarized images. The images converted using these approaches are the inputs for the recognition stage for both template matching and neural networks.
- An experiment on recognition using template matching with normalized correlation as the evaluator has been shown, satisfactory results have been obtained in this case for red road signs but unsatisfactory results have been obtained for blue road signs.
- A second experiment, based this time on NNs has been developed. The main feature here has been the use of a small number of training specimens to ease the election of sample sets as well as the use of a cascade network based on the results of the single-stage system. This second net contains only the specimens where confusion arises from the first NN, as a result incorrect recognitions are solved more efficiently. The results show that the cascade network provides the best results when compared to the other approaches. Here only an improved achievement of a false detections filter is required.
- Finally, It has been stated that the different performances of the NN and template matching can be mixed in the case of a speed supervisor to fulfill the time requirements, while in the case of automatic inventory time is not important, a NN is the most effective approach as long as it is possible to store a large amount of different road signs.

Chapter 5

Tracking

In this chapter an approach to the current research trends for road sign recognition will be presented. Tracking is considered as the final stage for road sign recognition systems, this may also be applied to the system presented in this thesis. Tracking is used to guarantee the recognition process and to solve any confusions that may occur in the recognition process, both through repeated positive classifications. In this chapter a model, that has been developed for this purpose, is presented. The aim of this research is to prove that it is possible to properly track road signs therefore no direct application of the method has been performed.

The information from both the inertial sensor, depicted in Appendix D and a camera motion model are used to predict the on-screen location of the road sign from the information provided by the first recognition. The tracking will be carried out separately for each road sign, i.e. tracking for each of the recognitions until the road sign is lost.

The basic movement of the vehicle which has been modeled is the depicted in Figure 5.1. The motion of the car may be described by a curve but for the tracking system it will be considered as a straight line defined by $\mathbf{t}$ the translation vector and θ the direction angle between two different measurements.

5.1 Camera Motion Model

The camera motion model is based on the origin in a Euclidean space (the camera) and a reference point (the road sign). The origin is continuously

changing due to the movement of the vehicle, note that the road surface is approached to flat world. Thus all possible camera movements are limited to the XZ plane. The following equations define the movement of the vehicle.

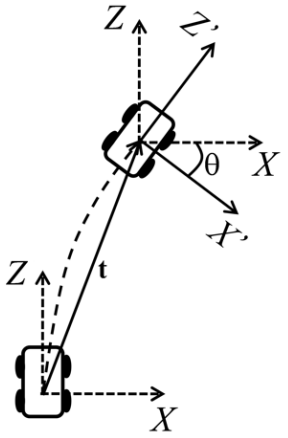


Figure 5.1: Trajectory of the vehicle between two points. The movement is described by a translation vector $\mathbf{t}$ and the angle between two measurements θ.

5.1.1 Notation and Previous Statement

- Homogeneous coordinates have been used to simplify multiplication.
- Matrices: are defined using capital letters and subscripts are used when distinguishing between different transformations. For example, $\mathbf{Rot}_\theta$ is the rotation on the variable θ.
- Vectors: are defined using small bold letters and subscript is used for additional information where required. For example, $\mathbf{t_i}$ stands for a particular *initial* vector.
- Points: are defined by capital letters with respect to the transformations along with a subscript which refers to the type of conversion. For example, X_{rs} refers to the first component of a point which belongs to the road sign. Other points may be defined by small letters, for example, x_0 refers to the first component on the center of the screen.

- Reference System. The euclidean space is defined following the scheme shown in 5.2.

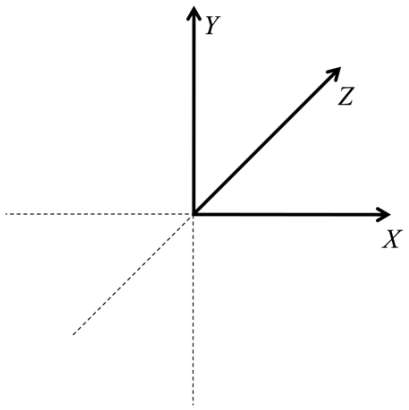


Figure 5.2: World reference system.

- To simplify calculations, the Z_{rs} component to the road sign, as depicted in Figure 5.3, will be approached by the distance from the road sign road sign to the camera D_{rs}. To show that this approach is adequate, the error has been calculated for the worst case scenario, this occurs when the biggest road sign model is located at the side of the screen. Applying equation 5.2 for $s_{rs} = 64$ and using the pinhole model described equation 5.4 and by letting $D_{rs} = Z_{rs}$ both D_{rs} and X_{rs} are obtained. Applying Pythagoras's theorem the real value of Z_{rs} can be calculated. The error committed is obtained from equation 5.1.

$$error = 1 - \frac{Z_{rs}}{D_{rs}} = 1 - \frac{\sqrt{{D_{rs}}^2 - {X_{rs}}^2}}{D_{rs}} = 1 - \sqrt{1 - \left(\frac{X_{rs}}{D_{rs}}\right)^2} \quad (5.1)$$

For the previously described worst case scenario the error committed was $error = 0.003$, the low error value obtained ensures that the approximation that $D_{rs} = Z_{rs}$ is valid and hence the performance of the tracking system will not be affected.

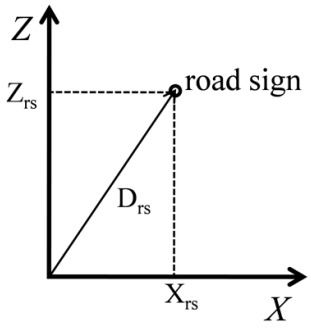


Figure 5.3: Scheme for the approach Z_{rs} to D_{rs}.

5.1.2 World Location

The first step in this process is to allocate the road sign. To do this a series of points as depicted in Figure 5.4 are used, s_{rs} is the size of the road sign in the image, x and y are the coordinates of the center of the road sign and (x_0, y_0) the center of the camera.

Figure 5.4: Image reference system.

s_{rs} and the real road sign size S_{rs} are known, thus the value of the road sign world coordinates X_{rs} and Y_{rs} can be calculated using equation 5.2 and 5.3:

$$X_{rs} = \frac{S_{rs} \cdot (x - x_0)}{s_{rs}} \tag{5.2}$$

$$Y_{rs} = \frac{S_{rs} \cdot (y - y_0)}{s_{rs}} \tag{5.3}$$

The coordinate Z_{rs} may be calculated using the pinhole approach, where this has been explained in Appendix B and using equation 5.4. An image depicting the model is presented in Figure 5.7.

$$Z_{rs} = \frac{f \cdot F \cdot S_{rs}}{s_{rs}} \tag{5.4}$$

5.1.3 Model Description

Once the road sign has been located a camera motion model is used to track this road sign. This model is composed of two parts, the first is based on a translation and rotation which defines the movement of the vehicle on the road and hence the camera placed within the car. The second part uses a description of the world points on the screen. The problem to be solved here is that of the location of a point in the screen given by a world fixed point (the road sign) and a moving reference system (the camera).

- **Translation**

  The translation of the world point from an initial to final point may be described, as in Figure 5.5, using the the vector $\mathbf{t} = \mathbf{t_i} - \mathbf{t_f} = (t_x, t_y, t_z)^T$. The transformation follows equation 5.5:

$$P_t = \mathbf{Tr_t} \cdot P_{rs} \tag{5.5}$$

  this can be extended to:

$$\begin{pmatrix} X_t \\ Y_t \\ Z_t \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{rs} \\ Y_{rs} \\ Z_{rs} \\ 1 \end{pmatrix} \tag{5.6}$$

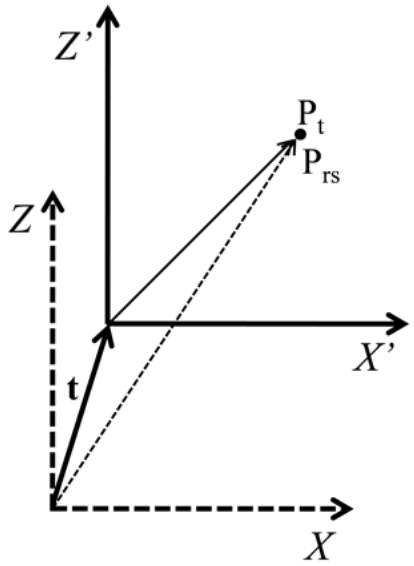


Figure 5.5: Translation of the reference system by the vector **t** in the XZ plane.

- **Rotation**

  The rotation is defined by the yaw angle, it is important to remember that it has been considered the approach of flat world in this part. This way the only angle dependence is due to rotations about Y axis which is represented by the variable θ.

  The movement follows equation 5.7:

$$P_r = \mathbf{Rot}_\theta \cdot P_t \tag{5.7}$$

  which may be extended to equation 5.8:

$$\begin{pmatrix} X_r \\ Y_r \\ Z_r \\ 1 \end{pmatrix} = \begin{pmatrix} cos\theta & 0 & -sin\theta & 0 \\ 0 & 1 & 0 & 0 \\ sin\theta & 0 & cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_t \\ Y_t \\ Z_t \\ 1 \end{pmatrix} \tag{5.8}$$

  The rotation transformation is depicted in Figure 5.6

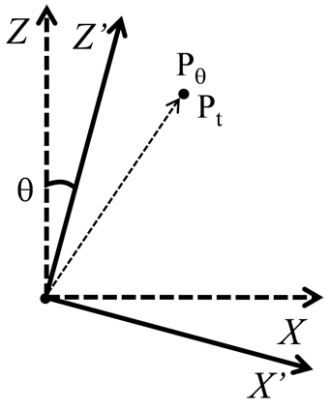


Figure 5.6: Rotation of the reference system about the Y axis and angle θ.

- **Pinhole Approach**

  The resizing over the XY plane is done using the pinhole model approach. This represents the relations between an object on the space and its projection on a plane. An image depicting this model is presented in Figure 5.7.

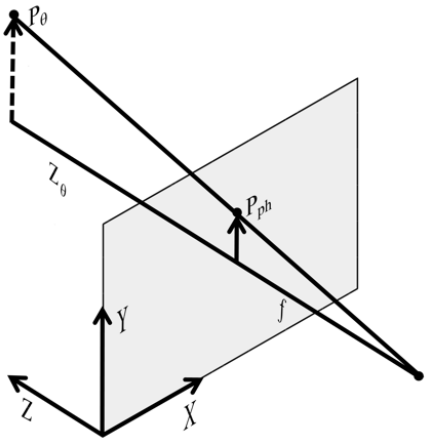


Figure 5.7: Pinhole scheme. The picture represents the variables of the model. It has been used a plane before the focal point to represent the image.

The pinhole model follows equation 5.9 with the matrix form of equation B.6:

$$P_{ph} = \mathbf{Res_t} \cdot P_r \tag{5.9}$$

$$\begin{pmatrix} X_{ph} \\ Y_{ph} \\ Z_{ph} \\ 1 \end{pmatrix} = \begin{pmatrix} \frac{f \cdot F}{Z_\theta} & 0 & 0 & 0 \\ 0 & \frac{f \cdot F}{Z_\theta} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_r \\ Y_r \\ Z_r \\ 1 \end{pmatrix} \tag{5.10}$$

where f is the focal length and F a pixel to mm conversion factor. Z_θ is the euclidean distance from the camera to the road sign. The subscript ph stands for the pin hole.

- **On Screen Location**

  To adequately reallocate the on screen image two further operations must be carried out:

  - Translation of the coordinates, this is done because the usual screen reference system used has its origin located at the top left corner.
  - Reflection of the XZ plane, this operation is done because the Y axis is now taken positive as opposed to the previous negative direction.

  The translation of the screen depicted in Figure 5.8 follows the vector $\mathbf{t_s} = (x_0, y_0)$ where x_0 and y_0 are the center point of the screen.

  The movement is defined by equation 5.11 which, in matrix form, is represented by equation 5.12.

$$P_{st} = \mathbf{Tr_0} \cdot P_{ph} \tag{5.11}$$

$$\begin{pmatrix} X_{st} \\ Y_{st} \\ Z_{st} \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & x_0 \\ 0 & 1 & 0 & -y_0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{ph} \\ Y_{ph} \\ Z_{ph} \\ 1 \end{pmatrix} \tag{5.12}$$

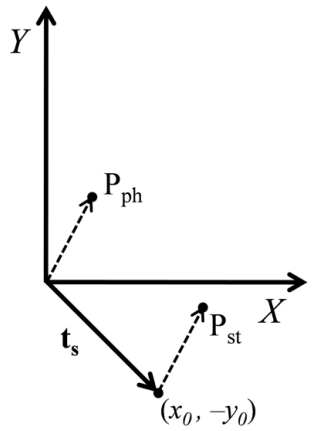


Figure 5.8: Translation of the object to the new origin which is the top left corner of the screen.

The reflection of the XZ plane provides the final on screen points following the scheme presented in Figure 5.9 and using the condensed and extended equations 5.13 and 5.14.

$$P_s = \mathbf{Ref_{xz}} \cdot P_{st} \tag{5.13}$$

$$\begin{pmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{st} \\ Y_{st} \\ Z_{st} \\ 1 \end{pmatrix} \tag{5.14}$$

The composition of all equations gives the condensed form 5.15 and the extended equation 5.16:

$$P_s = \mathbf{Ref_s} \cdot \mathbf{Tr_0} \cdot \mathbf{Res_{ph}} \cdot \mathbf{Rot}_\theta \cdot \mathbf{Tr_t} \cdot P_{rs} \tag{5.15}$$

$$\begin{pmatrix} X_s \\ Y_s \\ Z_s \\ 1 \end{pmatrix} = \begin{pmatrix} \gamma \cdot cos\theta & 0 & \gamma \cdot sin\theta & \gamma \cdot t_x \cdot cos\theta - \gamma \cdot t_x \cdot sin\theta + x_0 \\ 0 & -\gamma & 0 & -\gamma \cdot t_y + y_0 \\ sin\theta & 0 & cos\theta & t_x \cdot sin\theta t_z \cdot cos\theta \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X_{rs} \\ Y_{rs} \\ Z_{rs} \\ 1 \end{pmatrix} \tag{5.16}$$

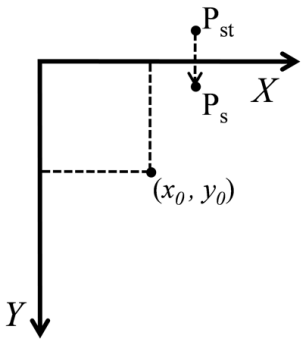


Figure 5.9: Reflection of the translated point used to obtain the coordinates of the screen reference system.

where $\gamma = \frac{f \cdot F}{Z_\theta}$
Substitution using:

- $A = t_x + X_{rs}$
- $B = t_y + Y_{rs}$
- $C = t_z + Z_{rs}$

provides a simplified set of formulas which relate a world point with a point on the screen after the complete movement. These are given by equations 5.17 and 5.18.

$$X_s = x_0 + \gamma \left[A \cdot cos\theta - C \cdot sin\theta \right] \tag{5.17}$$

$$Y_s = y_0 - \gamma(1 + B) \tag{5.18}$$

5.2 Experimental Trials

The algorithm has been tested on different sequences where the following sets of data have been stored:

- Data from the inertial sensor: speed and angle.
- Data from the camera: location of the road sign.

The following two experiments provide a brief summary of the behavior of the algorithm. The images were recorded within a controlled environment, i.e. the grounds of University Carlos III de Madrid, this has been carried out for safety reasons.

The first sequence of data is obtained from a steady curve which follows the trajectory presented in Figure 5.10a, this corresponds to a commonly used driving manoeuvre. The second experiment contains a series of sharp bends as depicted in Figure 5.10b this experimental trial has been designed to test the robustness of the prediction system. The distance between the expected and predicted coordinates have been measured for these two cases and for road sign sizes between 30 an 64 pixels, where these sizes have been chosen to satisfy the requirements of the system with respect to the detection distance.

The results of these experiments are presented in Table 5.1 along with the two tracking examples which are presented in Figure 5.10.

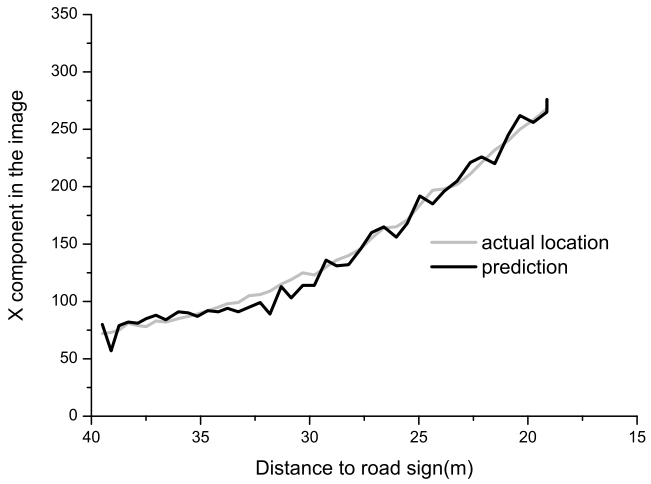
Trajectory	$\overline{s}$	Δs
sharp bends	10	6
steady curve	6	5

Table 5.1: Results from tracking for a series of sharp bends and a steady curve, $\overline{s}$ is the mean value of the distance in pixels between the real and predicted measurements, Δs is the standard deviation of the data.

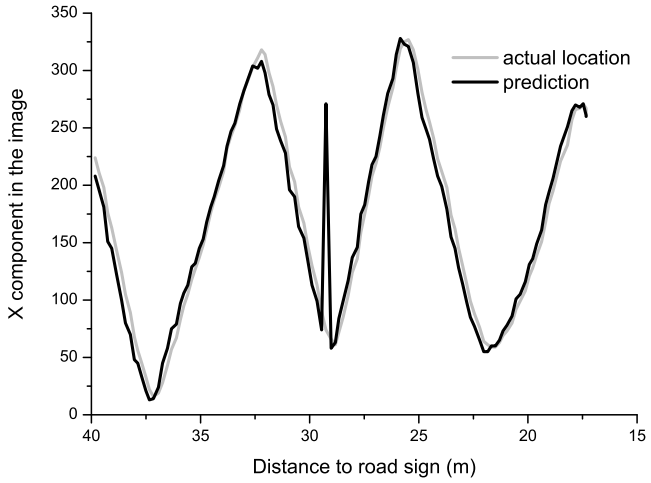
It may be seen that satisfactory results have been obtained, although eventually the situation arises which goes beyond the logical prediction range, this is seen to occur for the series of sharp bends. It must be emphasized that this particular trial has been designed to test the robustness of the system where the sharp bends are not commonly present in real life driving conditions, see Figure 5.10.

Figure 5.11 shows an example of a sequence where tracking has been applied. Three images are seen together which correspond to three different frames which are depicted in graph 5.10a. The first prediction, as may be observed from the graph, is not accurate however the rest of the predictions are observed to indicate the real location of the road sign in the image.

The use of this tracking system for real applications goes beyond the scope of this thesis. However, it must be highlighted that the integration of this technique provides many advantages when considering road sign detection



(a)



(b)

Figure 5.10: Examples of tracking. a) Steady curve. b) Series sharp bends.

systems. The two main areas of investigation using this technique are, first, to take into account subsequent recognitions of the same road sign to avoid false recognitions due to miss classification or missed detections. Secondly, it is possible to establish a reduced search area where the location of a road sign is predicted from the tracking algorithms. This also reduces the execution time as well as the number of possible false detections. When used along with the template matching detection technique, it is possible to restrict the number of models by using the prediction of the size of the new road sign to be found.

Figure 5.11: Three different frames are presented where the behavior of the tracking algorithm is seen. The circles represent the prediction of the model for the size and location of the road signs, the black line is the real trajectory of the road sign taken from subsequent images.

5.3 Summary

- A model which predicts the location of the road sign in the image has been presented. This model is based on the motion of the camera with

respect to the road sign. The data is obtained from both an inertial sensor, providing speed and angle, and from the screen which provides the location of the road sign.

- Results have been presented which indicate the accuracy of the tracking algorithm where the robustness has been tested when the system is exposed to adverse conditions, i.e. a series of sharp bends.
- Two applications of the tracking algorithm have been identified, these are the assurance of the road sign recognition which avoids false detections and reducing the search area of the image which reduces the execution time of the application.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

In the research presented in this thesis the complete implementation of a vehicle security system has been developed. This system is based on the retrieval of information from road signs.

The serious problems associated with traffic accidents have been presented highlighting the large amount of fatal accidents which occur each year and the socioeconomic consequences. The government concerns on this issue have been indicated along with a list of the measures to face this growing problem.

In most cases, incorrect vehicle speed is a major factor associated with traffic accidents, this parameter increases the danger and effects of accidents. For these reasons, special attention is being focused on this area, and different safety precautions have been employed to reduce vehicle speed such as: proposals for awareness campaigns, different engineering techniques and new technologies. The latter of these has motivated the research work performed and presented in this thesis where the present work has been focused on Advanced Driver Assistance Systems. Such systems provide the driver with important information which aid in the correct application of driving maneuvers thus increasing the level of safety for the occupants of the vehicle, other vehicles and pedestrians.

By means of an extensive and complete literary review in the areas of detection, recognition and tracking of road signs, it has been observed that there has been no direct application to security systems. Also mentioned, at the beginning of this thesis, have been the commercial attempts to develop

systems similar to that presented in this thesis, however only one system has been implemented, this may be found on-board an Opel Insignia, the goals of this particular commercial system are not as far reaching as those presented in this thesis as only a limited number of road signs may be detected.

The system proposed in this thesis has been based on the recognition of road signs using a computer vision system. The system has been designed to warn the driver when the speed of the vehicle does not satisfy the information provided on the recognized road sign. The system not only fulfills the requirements of a speed supervisor for both urban and non-urban environments but it also acts as a platform for developing further applications. One particular industrial application, which is achieving much interest, is a system that automatically stores road signs for inventory purposes.

The wide scope of this project, which has tackled the problems associated with detection and recognition, for all types of environments has been a new and difficult challenge. From the literary review it has been observed that the proposals has not included urban environments, this is due to the positioning of road signs and their conservation which is poor when compared to those located on highways. To enhance the detection efficiency, particular attention has been paid to the effect illumination has on the color of the road signs and also the presence of other objects, which are similar to road signs, within the captured images.

Special attention is focused on the theoretical background of this application as well as real life implementations. A study of the commercial viability has also been performed by looking up data from different driver surveys to obtain information on the real interests that a system, like the one proposed here, has on the current market.

In the work presented in this thesis, the three stages which form a road sign detection system have been studied in detail. for each one of these stages (detection, recognition and tracking), it has been observed that there are still many areas of unresolved problems, with no general solutions but *ad hoc* methods focused on specific problems, which require considerable compromises to be made between the different system parameters such that the results obtained are optimized.

6.2 Contributions

6.2.1 Detection

It has been mentioned in the state of the art regarding detection that there are two main approaches: the first uses color information and the second uses grayscale images. These two scopes are completely different as the color provides information on the components of each color space required to initially segment the objects, in the second case, grayscale images, the method relies on the extraction of geometrical characteristics which belong to the objects of interest.

When dealing with color, the color spaces used depend on the illumination conditions, i.e. RGB and a variety of other color spaces that are not as dependent on illumination changes, for example HSL, HCL, HSV or HSI. These color spaces have been used to extract the information from the road sign color under different weather conditions, times of the the day and seasons.

From the literary review it has been observed that the most commonly used methods are based on hue also many cases take advantage of the saturation, however no in depth study or comparative analysis on the influence of these components taken either separately or together has been presented.

In this thesis, a complete study has been performed on the color of the road signs under different illumination conditions where three possible weather conditions have been defined, Sunlight Shadows and Half-light. These three conditions have been sufficient to describe the variety of situations that may be found with respect to the road sign color. When direct sunlight is incident on the road sign the color generally presents a high level of saturation, this level is generally lower for shadow conditions. The half-light scenario has been added to identify the behavior of the system when the road sign is located between the vehicle and the sun, and when it is not possible to distinguish if the sample belongs to the sun or shadows, this is commonly observed due to the presence of objects, such as trees, which create shadows on a sunny day.

An in depth study of the color of the main sources of false positives has been performed, where these have been identified as bricks when searching for red road signs and the sky and asphalt when detecting blue road signs. Due to the absence of‘published results and experimental trials carried out in real conditions, it was considered an interesting task to obtain information

belonging to objects which are generally sources of false detections. From this analysis, it has been possible to obtain information on the limitations of the system and to determine methods to avoid these objects while at the same time loosing as little information as possible on the color provided by the image of the road signs. It has been demonstrated that for the general case it is not possible to avoid enhancing other objects in the image, however, by focusing on specific parameters a conditional search based on the color components provides satisfactory results.

It has been proved that the use of the hue information is generally more than sufficient to adequately enhance the road sign color for different illumination conditions, however in certain cases where the illumination conditions are known or when the aim is to recover only the road signs which are subject to environmental conditions described above (sun, shadow and half-light), it is possible to use the second component of the illumination independent color space to enhance the road sign color i.e. the saturation and chroma components.

Different algorithms have been introduced which are used for color normalization where in the case of a non real time system or when applying different assumptions can provide satisfactory detection results for different illumination conditions. This technique can also be used to detect achromatic road signs, this is a further objective of current research within this area.

It has been shown that the distance from the camera to the road sign and weather conditions affect the edges of the road signs and in many cases produce blurring. The method used in the detection stage has been based on template matching which depends on the image for which the matching is to be performed and the model used to do it. The integration of these two ideas has been based on a study of the borders from different profiles which can be used for detection using template matching. This original work has provided useful information for developing the system presented in this thesis. It has also been shown that when depending on detection distance or when accuracy is not regarded as one of the limiting parameters of the system it is possible to obtain improved results from the different models.

6.2.2 Recognition

In the state of the art, presented previously in this thesis, a list of the most common techniques used to develop the recognition stage have been pre-

sented. It has been shown that the most widespread methods to perform this task are based on template matching and neural networks. Currently, new techniques involving support vector machines have been used to classify road signs. The success of the neural network approach for recognition has one major set back, this is the small variation of road signs that can be recognized. Another drawback to this technique is the amount of samples required to train the network and also the number needed to validate it is usually large, this leads to difficulties in carrying out this task.

The investigation presented here for the recognition stage has been focused on dealing with these unresolved problems. A comparative analysis has been carried out between the template matching and the neural network methods. This study, which has not been found in the literary review, has provided vital indications on which method is most adequate for the road sign recognition application and has considered the conditions imposed on each technique. A review of the advantages and disadvantages for each method has been presented in detail.

Preprocessing of the input candidates for both methods has been developed to optimize the results obtained. For the template matching case, grayscale images and binarized images have been used to obtain the differences between the two preprocessing approaches. It has been observed that the final results are highly dependent on this process.

This is especially true in the case of neural networks where the main goal has been to obtain smaller training sets. The grayscale images have been omitted and more emphasis has been focused on the optimization of results using different thresholding techniques. This original methodology has involved obtaining samples which are free from noise, this leads to images which only contain the most important information of each sample. Using this technique it is possible to manually select the most adequate training samples, this removes the necessity of using methods of principal component analysis or other similar algorithms. It has been demonstrated that neural networks are a powerful tool for classification, however two problems remain unsolved: the execution time of the stage and incorrect classifications of false candidates.

Finally a cascade based classification has been developed using the samples which lead to confusion, for example, speed limits. By doing this a coarse to fine classification can be developed using a small number of networks, this is because the number of confusion groups is also small.

6.2.3 Tracking

Tracking is the final stage that is currently under investigation for integration along with road sign recognition systems. This stage is primarily used for the following two tasks: assure detection and recognition and to reduce the road sign search region.

The first of these tasks involves the probability of predicting a region in a subsequent frame where a road sign is most likely to be located, using this method it is less probable that road signs are missed after being detected in previous frames. This technique may also be used as a filter to remove incorrect recognitions when a series of positive classifications have been carried out. The second task involves using information from the tracking technique to predict the location of the road sign on the image, this allows the search area to be reduced, resulting in reduced execution time as the road sign search algorithms are omitted from regions on the image, also a reduced number of models are required since it is possible to predict the size of the road sign that must be found. More models may also be used in a shorter range instead or less models in the same range, this improves the accuracy which is an important parameter of tracking systems.

To develop tracking systems, temporal integration is usually carried out on the data, more recently a Kalman filter has been applied, this indicates the likelihood that a recognized road sign is found in a specific area of subsequent frames. This has been shown to be an effective method to track road signs though the amount of operations involved in the predictions is very large.

In the work carried out on tracking systems, presented in this thesis, a tracking model has been developed using data obtained from the inertial sensor and a camera motion model. The road sign is taken as a reference system around which the camera is moved. Equations which take into account this movement, assuming a flat world, have provided predictions on the location of the road sign in subsequent images. The integration of this sensor has not been carried out before and it has been demonstrated that a straightforward model which is combined along with the data from the inertial sensor is enough to adequately track the road sign, the main goal here has been to demonstrate the accuracy of the method, thus indicating the viability of the tracking system for further integration in recognition platforms.

6.3 Future Work

The scope of this thesis covers a wide range of research areas, as a result this leads to a great number of fields for ongoing research. Following is a brief list of possible areas of future work which may be carried out in each stage of the recognition system.

6.3.1 Detection

Understanding color and its associated detection techniques is an important area of investigation not only for this application but also for a wide range of other detection schemes. Further research in this area would involve defining a color space which provides color components under constant illumination changes. Further areas of investigation involving grayscale images would incorporate algorithms based on geometrical features such as edges or shapes. Using other parts of the electromagnetic spectrum for detection, i.e. near infra red, is another area of future investigation.

A theory based on the achromatic colors would also aid in the detection of end of limits road signs which are gray (gray, black and white). Until now this type of road sign has been detected by searching for geometrical features instead of enhancing the color.

Little research work has been performed on obtaining information related to the maintenance of road signs. Information provided by the changes in color due to illumination, or spaces where the illumination has less influence, can be used to identify road signs which have degraded color properties and which need to be repaired or replaced.

The use of different cameras would also provide a great amount of real parameter information on which detection relies, it is possible that data may be filtered using the information obtained from different sensors, or at least provide the possibility of generalizing results obtained from a variety of different sensors. Another area of investigation would employ cameras based on CMOS technology, the main advantage associated with these types of camera when compared to standard devices is that the information from one pixel does not affect the information contained in neighboring pixels.

The addition of different photodetectors to evaluate illumination may solve the aforementioned problem of grouping the different illumination conditions. From this additional information specific algorithms may be used to achieve correct detections.

6.3.2 Recognition

The recognition process could be improved in many ways. In general, many advances may be achieved from further investigation on the preprocessing of candidates used by the neural networks and template matching. This preprocessing would help in the presence of adverse illumination conditions, i.e. shadows and back light, where candidates have been observed to lose information thus reducing the possibility of classification. By focusing on the cases where the thresholding is severely affected by the illumination it may be possible to find a pattern and reduce the cases where no recognition is possible due to the quality of the images acquired.

Template matching may be further developed using more realistic model templates for the template matching instead of perfect models which have been used up to now.

An inverse template matching technique using a cascade approach may also be performed using particular model features instead of using the complete model which has been presented in this thesis. The parts selected from the model would include the most characteristic features which define the road sign where a search would be performed for these features on the candidates. The cascade approach may be carried out by grouping different characteristics. These models would have smaller sizes to the samples which are currently being used thus reduce the execution time.

In the case of neural networks there is a large variety of improvements that may be carried out. This includes a study of the optimal size of the samples that would lead to a reduction in the number of matrix inputs and a reduced execution time.

To improve results it may also be possible to use part of the image instead of the complete image, i.e. by reducing the image to fit the numbers located on the speed limit road signs or other symbols which define the road signs. This could be achieved by designing smaller windows which vary in size depending on the group of road signs to be recognized. Exploiting these characteristics in a cascade approach among the confusion objects would improve both the results and execution time. This process would also lead to a more simplified selection of objects required for training and hence improved control and results.

6.3.3 Tracking

Though the model for tracking has been shown to be sufficiently accurate to follow the changes in direction of the vehicle, it is possible to improve the results by defining a more complex but realistic model which includes curve translations. This methodology would be very useful and its implementation would rely on the inertial sensor which provides quick and accurate data on the movement of the vehicle. Another area of interest here would be to include changes in the Y direction, this can be achieved by adding information from the pitch angle and would lead to both improved results and more accurate tracking.

A comparative analysis of the results between the Kalman filter and temporal integration would provide vital information on both the advantages and disadvantages of each method with respect to the complexity of the algorithm, time of execution and results accuracy.

Another area of investigation in this field would involve an analysis of the camera vibrations for different speeds and situations. By extracting this information the noise associated with the vibrations can be reduced or removed using specifically designed filters or removed mechanically by improved on board design and optimized sensor and camera locations.

Different approaches to the tracking technique may be carried out and compared to the methods presented in this document. Other approaches may involve on screen evaluation of the speed and direction instead of prediction using world variables.

Appendix A

Annotation of Images

One of the problems encountered when developing this system has been the laborious and tedious task of testing the reliability of the system i.e. the retrieval of data to generate the ROC curves and the analysis of the statistics. Straightforward methods to deal with this problem exist: annotation of images. In general, this is based on storage of image data which may be retrieved using a semantic search. For the research work carried out in this thesis the image annotation was used to build a database containing the main road sign features obtained from different images. This permits a more controlled management of the lists of data which may be required for further applications, i.e. to compare the results of the image annotation to those which are obtained from other applications. Using this methodology, it is possible to automatically test the functioning of the system as the annotation provides a basis for the ground truth.

The annotation of images has been used in previous research work where good results have been obtained, for example, in [130] the authors present a database which contains more than 20000 images. Each road sign image has been manually recorded and located within small protocol files. The road signs have been labeled using an identification number as well as a set of additional attributes such as position, size and visibility.

Another example has been presented in [113] where a semiautomatic framework has been developed for annotation which interpolates the information using cubic splines between every fifth frame. This assists the task of manual annotation.

The annotation process of this thesis follows the methods used by the aforementioned work from Professor Priese in [130] and was carried out using

a Linux application developed by the Arbeitsgruppe Aktives Sehen (AGAS) from Koblenz Universität.

The user interface developed for this task has been designed to be user friendly, see (Fig. A.1). Within this interface a list of different general attributes may also be found, there are: camera, type of illumination, weather conditions, project name and object. Each different class allows the addition of subcategories to optimize the organization of the data, this is done by simply inserting the names which are then automatically assigned an identification number. This application has been designed to perform a straightforward conditional search for image retrieval using the features contained in the lists.

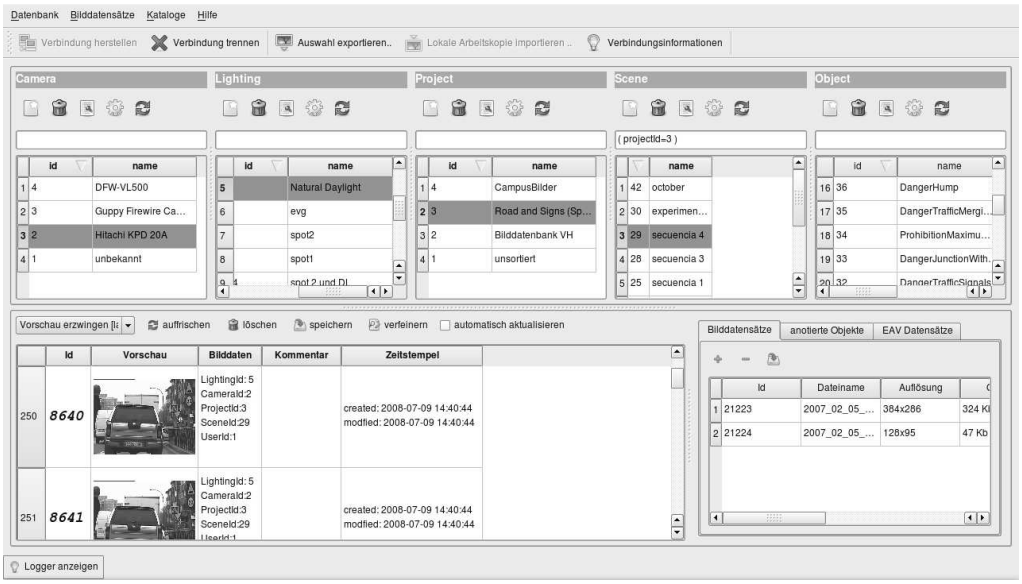


Figure A.1: General view of the annotation interface.

The most important part of this application is the annotation of objects of interest. This is done by choosing an object from the class list "Object" (where the list is opened to store a new object when necessary). Once the objects in the image have been selected, another window appears with the image and the list of objects of interest (Figure A.2). Clicking on one of the objects form the list allows it to be edited and after the user only has to define a frame, using the mouse, which contains the object of interest. This frame, in real time, provides the size of the object which is required for

selection, thus it is possible to directly reject candidates which do not fulfill the size requirements of the application. This last point has been essential to carry out the work presented in this thesis. Once the frame has been defined, the location of its corners are stored automatically.

The retrieval is carried out using SQL commands which permit a text file to be saved along with the data. Mathematical operations may also be carried out here, for example, extract the location of the center of the road sign and apply conditional statements to the search so that every possible data organization can be carried out.

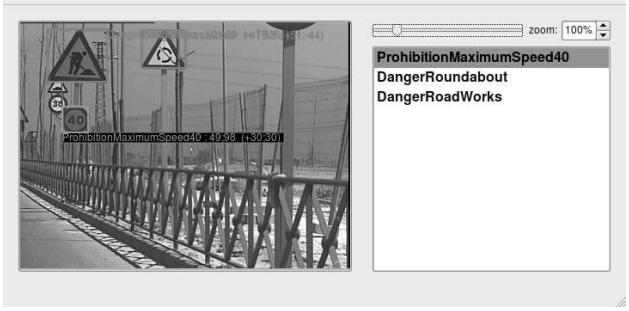


Figure A.2: Annotation Bounding Box Editor. On the left, the objects that are already annotated may be seen. The image shows the framed road signs with the size of the bounding box. In this way it is possible to reject candidates due to size considerations.

The different types of data which can be retrieved using this tool for the current application are: road sign location, road sign shape (danger, prohibition, yield, indication and obligation) and type of road sign. Using different sets of images which contain the complete range of real situations can be used to test the algorithms by simply storing the data of the application and comparing it to the annotation database.

Appendix B

System Settings

One of the main objectives of the system designed and presented in this thesis has been to perform speed supervision. Hence, the system must supply the information well in advance so that the driver can react adequately to the road sign which is detected. It is thus necessary to determine the minimum distance required to provide the information, here the speed of the vehicle for the worst case scenario must be considered.

The distance may be calculated in two different stages:

1 Distance due to the reaction time of the driver t_r which follows B.1:

$$s_r = \alpha \cdot v_i \cdot t_r \tag{B.1}$$

where v_i is the speed of the vehicle in m/s with $\alpha = 0.28$ the conversion factor from km/h to m/s and s_r the distance in meters traveled in t_r seconds.

2 Distance due to the braking time until the legal speed v_e is achieved, this is denoted by s_b.

The complete distance must take into account the contribution of both of these values, so:

$$\Delta s = s_r + s_b \tag{B.2}$$

As regards the driver reaction time, there is extensive literature in this area which state that this time is affected by different stimulus and by a

variety of circumstances. Two papers in this area have been used as reference sources as they confirm data which has been found from other non official sources.

In [81] the goal has been to measure the response of drivers to different situations. Among the different experiments, the most interesting one is that of the dual task condition. First, the person under study must memorize a telephone number. While driving in a circuit the driver has been requested to enter the first digit of that number in a cell phone, meanwhile an indication light will appear on the road which indicates that the driver must brake. This hypothetic case indicates the focusing ability of the driver in the presence of a distraction.

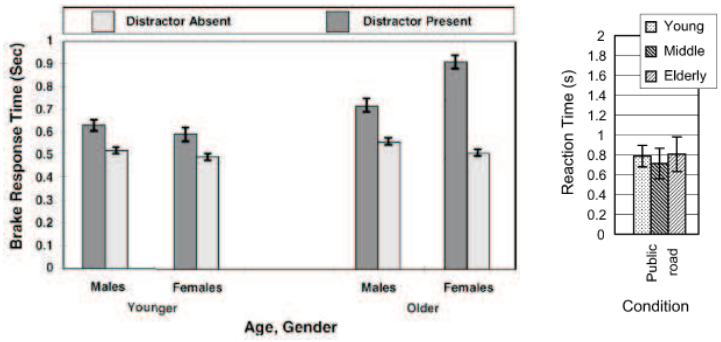
Among the many parameters which have been studied, a specific case has been highlighted which is presented in Figure B.1a. This is the brake response time (t_r) which indicates the time from when the brake signal appears to when the driver reacts by pressing on the brake. Among all possible groups of drivers, the slowest times corresponded to the group of older females, achieving a $t_r = 0.91s$.

In [107] the goal has also been to determine the driver reaction time (t_r), in this case the signal to stop has been provided acoustically. This experiment is more suitable to the application presented in this thesis as the driver is warned via recorded voice messages. The experimental procedure has taken place for different scenarios and for different aged and gendered drivers. Included within these experimental procedures are trials on public roads. From Figure B.1b the slowest times are demonstrated by older drivers where the average response time was $t_r = 0.8s$.

The reaction time from the experimental procedure involving the optical signaling is longer than that of the second scenario where no element of distraction is used. It has been decided to assign the reaction time in the work of this thesis as the mean value of both of these results. Thus the reaction time is calculated to be $t_r = 0.85s$.

To calculate the second distance s_b, the commission directive 2006/27/EC [2] is used which is an amendment to other directives regarding different laws which include braking devices. Here the equations used are briefly described and used to measure the stopping distances from the quality tests B.3. For further information see [1].

$$a_m = \frac{{v_i}^2 - {v_e}^2}{\beta \cdot s_b} \tag{B.3}$$



(a) (b)

Figure B.1: a) Brake response time vs. gender and age. b) Reaction time vs. age.

Which is easily derived from the well known:

$$v^2 = 2as \tag{B.4}$$

The parameters in B.3 are: a_m: fully developed deceleration in m/s^2. v_i: initial vehicle speed in km/h. v_e: final vehicle speed in km/h. $s_b = s_e - s_i$: braking distance which is the distance traveled between v_i and v_e in meters. The constant β is equal to 25.92 and stands for the conversion from km/h to m/s.

The parameters in B.4 are: v: speed. a: acceleration. s: distance.

Rewriting equation B.2 with the information from B.1 and B.3 defines B.5:

$$\Delta s = s_r + s_b = \frac{{v_i}^2 - {v_e}^2}{\beta \cdot s_b} + \alpha \cdot v_i \cdot t_r \tag{B.5}$$

Also, the distance from the vehicle to the road sign may be calculated using the well known pinhole model approach [51] which is depicted in Figure B.2 and follows equation B.6:

$$Z = \frac{F \cdot f \cdot D_{rs}}{d_{rs}} \tag{B.6}$$

where Z: distance from the camera to the road sign is in meters. F: resize non-dimensional factor, 0.5099. f: focal length in pixels, 3880.79. D_{rs}: Road sign real size in meters. d_{rs}: road sign size on the screen in pixels.

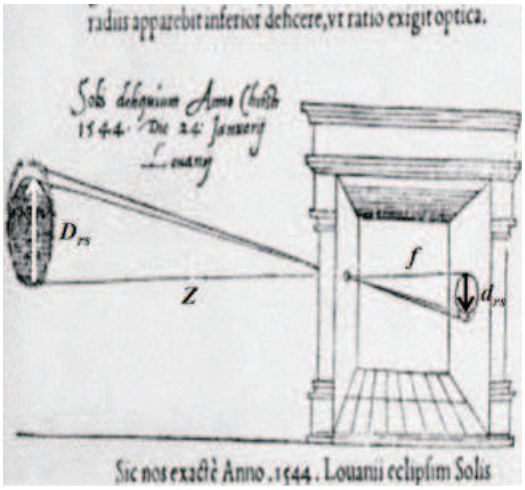


Figure B.2: Description of the magnitudes that define the pin hole model superimposed on a picture extracted from an obscure camera article by Dutch physician Gemma Frisius in the year 1544. In this case D_{rs} represents the real size of the road sign. Z is the road sign to lens distance. f the focal length and d_{rs} the size of the road sign on the screen.

Inserting eq. B.5 into eq B.6 and letting $Z = \Delta s$ and calculating the value of d_{rs} in terms of the speed gives equation B.7:

$$d_{rs} = \frac{\beta \cdot a_m \cdot f \cdot F \cdot D_{rs}}{{v_i}^2 - {v_e}^2 + \beta \cdot a_m \cdot \alpha \cdot v_i \cdot t_r} \tag{B.7}$$

Using this equation it is possible to calculate the size, in pixels, of a road sign using information from the speed difference between two points.

The only parameter left to define is a_m, this may be found in [136], where a graph which from B.3 describes the vehicle deceleration taking into account whether there is a brake assistant or not, and also the experience of the driver. It may be seen that the deceleration of a standard driver is on average equal to $7m/s$, this value will be used for further calculations.

The experimental calculation of the minimum d_{rs} in the worst case scenario is for a speed of 120km/h being reduced to 80km/h where the vehicle is a car with no braking assistant and is driven by an older female. The values of the speed limits have been selected according to the Spanish standard *Norma 8.11C de señalización vertical* [110] here it is stated that there cannot be two subsequent speed limit road signs with a speed difference of

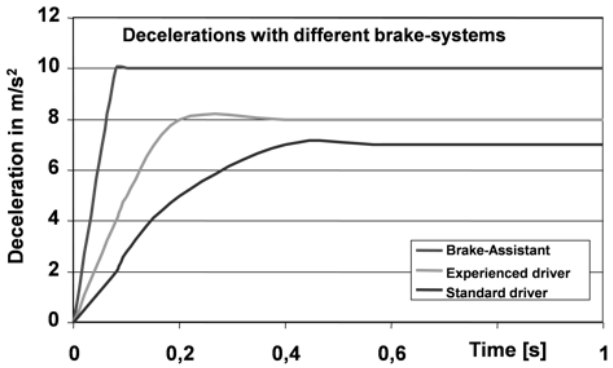


Figure B.3: Deceleration in terms of the experience of the driver and the use of a braking assistance system.

more than 40km/h.

The diameter of the smallest road sign on a highway is the prohibition road sign which is 1200 mm. Thus the result obtained is $d_{rs} = 33$ value required is satisfied by the camera since the minimum model is 30 pixels.

Appendix C

Geometrical Limitations and Lens Distortion

It is important to remember that following standard rules, road signs should be placed so that they are in the direction facing the driver which should be almost perpendicular to the direction of the moving vehicle. Unfortunately, as has been mentioned previously in this thesis, the aforementioned rules are not required to be fulfilled in urban environments due to the limitations of road sign placement. This causes slight skewing and rotations for both in and out planes. Road signs may also appear distorted because of the optical aberrations of the camera lens or simply by the movement of the vehicle. All these problems affect detection but fortunately in the majority of cases the location and recognition of road signs is still possible.

Several tests to verify the detection robustness have been carried out, three of these regarding different rotations and another related to lens aberrations.

C.0.4 Rotations

The tests have been designed and developed within the grounds of University Carlos III de Madrid. To simplify the experimental process real road signs have always been used. The three different shapes (circular, triangular, rectangular) have been moved by different operators, as shown in Figure C.2, for the three possible rotations roll, pitch and yaw. An image depicting the directions of these angles is presented in Figure C.1.

The experiment has been repeated several times for different distances

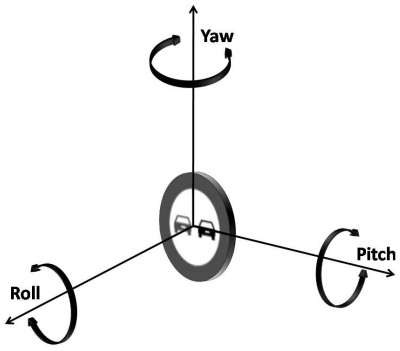


Figure C.1: The three angles under study. Roll causes in-plane rotations while pitch and yaw cause out-plane rotations.

of the camera to the road sign, this was done to determine the limits of the system where the candidates corresponding to the last detection before the road sign is not recognizable are stored. These candidates have been manually measured to calculate the maximum values for the angles. To carry out this experiment the cut-off value of the normalized correlation set in chapter 3 has been respected, where this is 40% for red road signs and 50% for blue road signs.

Some images from the result of this experiment may be found in Figure C.2. The quantitative results may be looked up in table C.1. From the values obtained it is seen that the algorithms discussed in the state of the art which deal with perspective changes are not required, this is because the system is capable of recognizing road signs in unfriendly environments which have a very low likelihood of appearing in real situations.

The detection conditions are dependent on the recognition of the road signs where the results may change from one road sign to another. However, the signs chosen for the experimental trials may be taken as a representative sample as the set of images annotated and used for detection and recognition also include road signs which suffer from in and out of plane rotations .

C.0.5 Lens distortion

To prove that there is no need to apply an undistortion algorithm to correct some lens aberrations the following test has been developed. First the camera has been calibrated to extract the main parameters. The method used was

Shape	Roll	Pitch	Yaw
Circular	16	42	47
Triangular	13	41	51
Squared	20	38	47

Table C.1: Results on angle changes in degrees in the two possible directions of each angle.

(a) (b) (c)

Figure C.2: a) Examples of detections under angle changes. a) Roll. b) Pitch. c) Yaw.

the common calibration based on the location of crossing points on a chess board. The calibration pattern was placed on a road sign to perform this experiment, see Figure C.3. To obtain the intrinsic parameters (focal length, center of the image and distortions) the Matlab Camera Calibration Toolbox developed by CalTech [13] has been used. Among other methods this toolbox implements the technique developed by Zhang, see [163] and[164]. The algorithm requires a flat calibration pattern and at least two images which are taken from different positions. First, the focal length and the center of the image are calculated analytically, this is followed by a non linear optimization method which is used to obtain the distortion parameters.

For the distortion calculation the tangential and radial components are taken into account by considering equation C.1 and C.2:

$$x_r = x_i + \overbrace{x_i\left(k_1 r^2 + k_2 r^4\right)}^{radial} + \overbrace{2p_1 x_i y_i + p_2\left(r^2 + 2{x_i}^2\right)}^{tangential} \tag{C.1}$$

$$y_r = y_i + y_i\left(k_1 r^2 + k_2 r^4\right) + p_1\left(r^2 + 2{x_i}^2\right) + 2p_2 x_i y_i \tag{C.2}$$

where (x_r, y_r) are the real coordinates of the pixels and (x_i, y_i) the ideal ones. r is the distance to the center of the image, k_1 and k_2 are the radial distortion coefficients, p_1 and p_2 are the tangential distortion coefficients.

To undistort the image a simple bilinear interpolation is carried out using the cvRemap function in the OpenCV libraries [15] which is documented in [14].

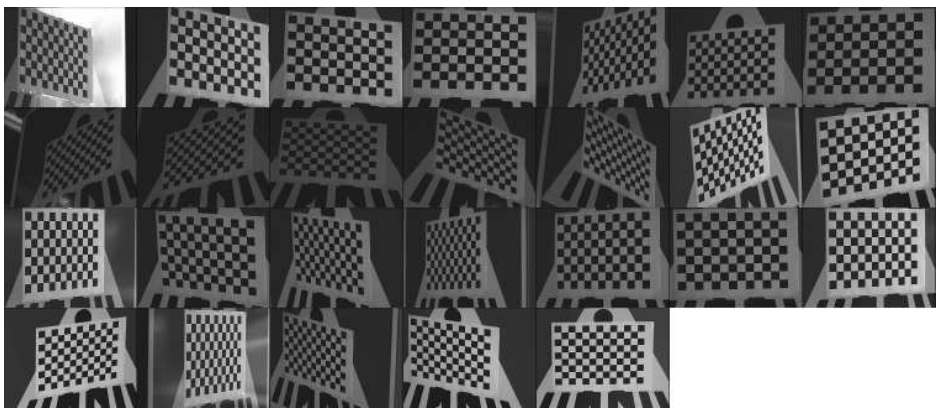


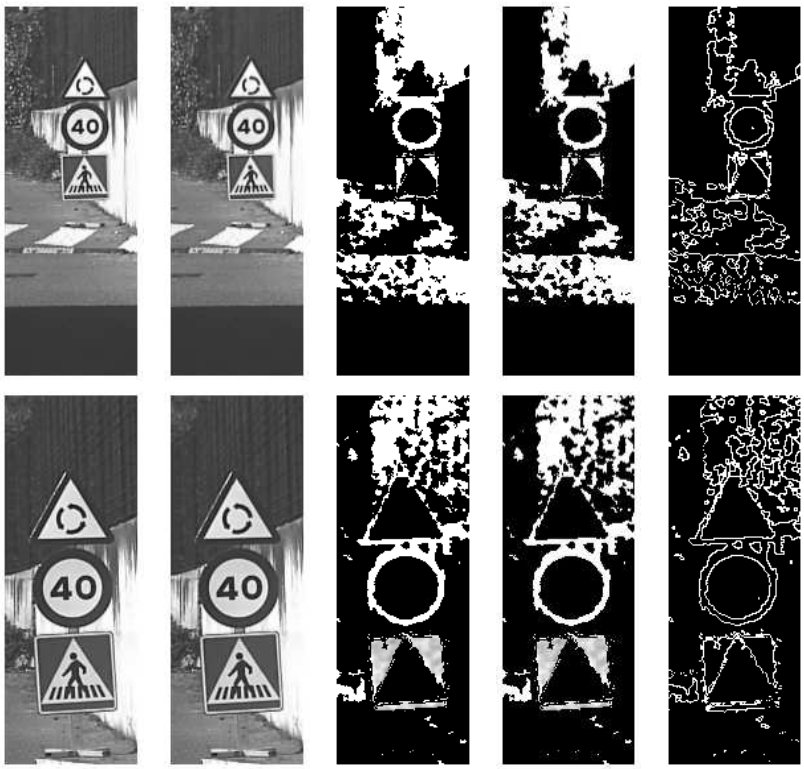
Figure C.3: Changes in the angle of view of the chessboard pattern to calibrate the camera.

The experimental set up has been carried out in the grounds of the University Carlos III de Madrid, here it is possible to perform experiments without endangering other vehicles.

The road signs used in this experiment belong to the ISL and have been selected to perform trials using the three main road sign shapes, this is done to ensure that all shapes are located within the same image and under the same conditions. The tests consisted in rectifying two images where the road signs appeared at the edge of the image. This was done because the sharpest effects of tangential and radial optical aberrations occur here.

The two images, Figure C.4a, correspond to the smallest model used in detection (30px) and the largest model (64px), these sizes correspond to the detection limits. Only the largest and smallest have been investigated as the intermediate sizes are expected to provide similar behavior.

The images are rectified to correct the distortion believed to be produced by the lens, see Figure C.4b. It has been observed that there is no apparent difference between both cases. After rectification, the color characteristics of both the converted and original images are enhanced with the color space chosen and discussed in the Detection chapter (HCL(H) for red road signs and HSL(H) for blue road signs). The effect of this enhancement may be seen in Figure C.4c and C.4d. Finally, subtracting the distorted and undistorted images is done to provide an improved visual understanding of the results. This may be seen in Figure C.4e where the differences observed are insignificant, only one pixel in every case, this test has verified that undistortion algorithms are not required for this application.



(a) (b) (c) (d) (e)

Figure C.4: Process of sample undistortion corresponding to the smallest and biggest models (30 and 64 px.) used in detection and for the three shapes used. a) Original images. b) Rectified images. c) Enhancement of a). d) Enhancement of b). e) Subtraction and binarization of c) and d)

Appendix D

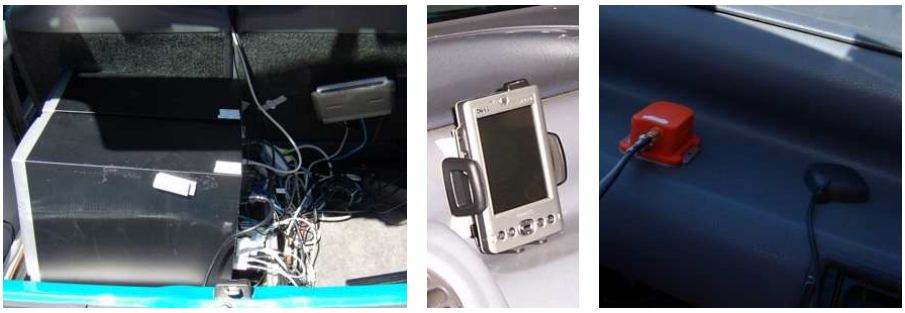
IVVI

IVVI is the acronym for Intelligent Vehicle based on Visual Information. To perform all on-board vehicle tests a Renault Twingo has been used as the experimental platform to develop and test security systems based on computer vision. The vehicle used may be seen in Figure D.1.

Figure D.1: IVVI. Intelligent Vehicle based on Visual Information. Experimental platform for research on ADAS based on computer vision.

The vehicle is equipped with two computers which manage and control the information provided by the different systems, and a WiFi router for

wireless communications (see figure D.2a). Recently a PDA has been added to the road sign recognition system along with an inertial sensor, the latter has substituted the previous GPS device. This new equipment provides more accurate results on the location of the vehicle and at a faster rate. The two devices are shown in figure D.2b and figure D.2c.

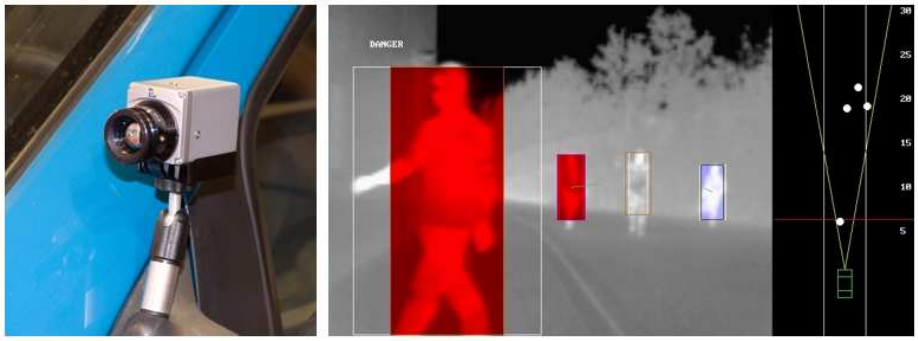


(a) (b) (c)

Figure D.2: System equipment. a) Computer and WiFi wireless. b) PDA. c) Inertial sensor with GPS.

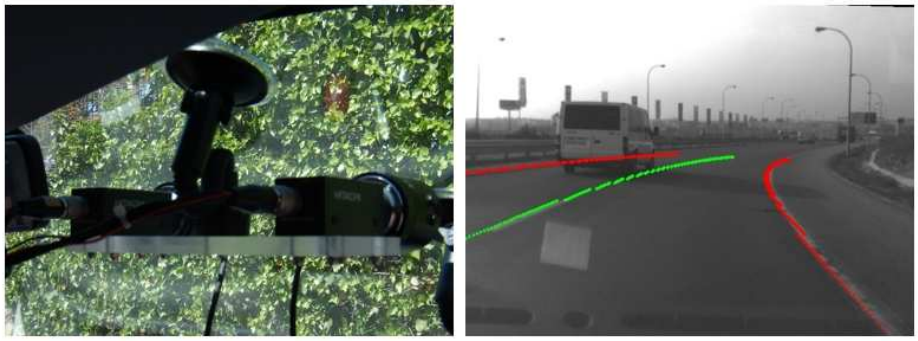
As part of the on-board IVVI system there are different cameras, each on performs separate tasks (see Figure D.3). A stereo system composed of two B&W cameras is used for the lane departure warning system and for detection of pedestrians. An infrared and visible camera are located in front of the driver to detect drowsiness and distractions via an ocular analysis. There is also a camera located on the wing mirror which is used for detection of pedestrians in nocturnal conditions. The interface for each of these systems is also presented in Figure D.3.

Finally there are two monitors Figure D.4, A large one (14 inches) located in the backseat of the car which is used for demonstrations and developing applications, and a smaller (7 inches) portable screen used for tests.



(a)

(b)



(c)

Figure D.3: Examples of the security systems on-board the IVVI. a) Lane departure System. b) Driver monitoring . c) Pedestrian detection for nocturnal conditions.

(a) (b)

Figure D.4: Screens used for the development and test of the system. a) Backseat screen. b)Co-driver screen.

Appendix E

System and Interface

E.0.6 System

There are several different components used to form the complete IVVI system. A Hitachi KPD-20A color camera used for acquisition of images. The information of the camera is sent directly to a AMD 64, 2 GHz, 1 GB computer. The inertial sensor, Xsens MTI-G is continuously broadcasting data on angles, times and speed. The two sets of data obtained from the camera and inertial sensor are controlled by the computer. When a message is to be sent it is done in two different ways which depend on the status of the equipment, i.e. testing or driving mode. The first uses the large screen located in the backseat of the car, while the second uses the PDA. All of these devices may be seen in Figure D.2.

E.0.7 Interface

There are two possible interfaces which allow interaction with the system:

- Developer interface and test interface.
- Driver interface.

Developer ant test interface. This interface, either on the large or small screen, shows the current image of the road along with information on the detections and recognitions (see Figure E.1). The objects which have been detected are surrounded by a green frame while the recognition is presented in the form of an image of the recognized road sign. A reliability color

scale is also included here which is positioned at the right hand side of the image, the color of the scale beside each road sign indicates the reliability in the form of the tone of the frame which goes from poor recognition reliability (red) to definite recognition reliability (blue).

If required, the system sends a warning message via the vehicles speaker system along with an on-screen message which provides information of the warning.

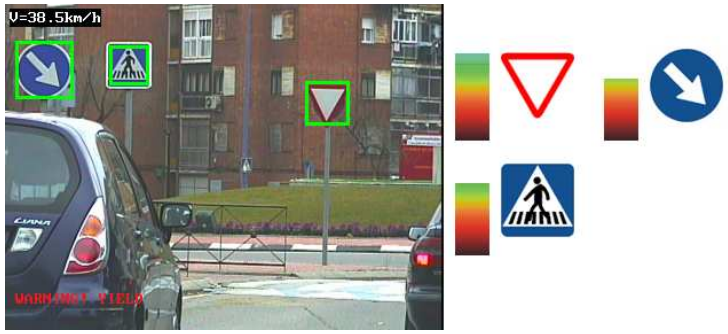


Figure E.1: Developer and tester interface. The framed detections may be seen in green and the recognitions on the right side with the reliability level. A warning message appears on the bottom left side of the screen indicating that a yield has been identified

Driver interface. In this mode of operation the computer screen is replaced by the PDA , see Figure E.2, which is placed in the front part of the vehicle. When the system has recognized a road sign and the speed of the vehicle is not appropriate, the road sign appears on the screen along with a voice message that warns and advises the driver on the appropriate manoeuvres to be made. There are a variety of warning messages, for example, The speed messages advises the driver to “reduce your speed” or in the case of a road works road sign “Attention, road works ahead”.

Figure E.2: Driver interface. The current detection may be seen on the screen. A voice message is also sent from the PDA.

Bibliography

[1] Official journal of the european union. commission directive 1998/12/ec, 1998. http://eur-lex.europa.eu/JOIndex.do?year=1998&serie=L&textfield2=81&Submit=Search&_submit=Search&ihmlang=en.

[2] Official journal of the european union. commission directive 2006/27/ec, 2006. http://eur-lex.europa.eu/JOIndex.do?year=2006&serie=L&textfield2=66&Submit=Search&_submit=Search&ihmlang=en.

[3] Commission internationale de l´eclairage, Last access: April 2009. http://www.cie.co.at/index_ie.html.

[4] Easyrgb, Last access: April 2009. http://www.easyrgb.com/.

[5] Ivvi, Last Access: April 2009. http://turan.uc3m.es/uc3m/dpto/IN/dpin04/ISL/SIT.html.

[6] Mobileye, Last access: April 2009. http://www.mobileye.com.

[7] Blaupunkt navigator device, Last access: February 2009. http://www.blaupunkt.com/de/7612201639_main.asp.

[8] Distronic plus, Last access: February 2009. http://www4.mercedes-benz.com/e/cars/s-class/betriebsanleitung/erleben/distronic_plus.html.

[9] Driver monitoring, Last access: February 2009. http://www.nissan-global.com/EN/TECHNOLOGY/INTRODUCTION/DETAILS/DPCC/.

[10] Mercedes blind spot warning system, Last access: February 2009. http://www4.mercedes-benz.com/e/cars/s-class/betriebsanleitung/erleben/totwinkel.html.

[11] Parking assitant device, Last access: February 2009. http://rb-k.bosch.de/de/sicherheitkomfort/komfort_beim_fahren/fahrerassistenzsysteme/einparkhilfe/index.html.

[12] Road signs database, Last access: February 2009. http://bancoimagenes.isftic.mepsyd.es.

[13] Camera calibration toolbox for matlab, Last access: March 2009. http://www.vision.caltech.edu/bouguetj/calib_doc/.

[14] cvremap documentation, Last access: March 2009. http://proquest.safaribooksonline.com/9780596516130/remap.

[15] Opencv, Last access: March 2009. http://sourceforge.net/projects/opencvlibrary/.

[16] Continental, Last access: May 2009. http://www.vdo.de/press/releases/chassis_safety/2008/pm-2008-08-09-safety-days-modul-3-de.htm.

[17] Daimler news, Last access: May 2009. http://www.gavrila.net/Media_Coverage/NUZ_2008y07m05d_Daimler.pdf.

[18] Opel eye news 1, Last access: May 2009. http://www.elmundo.es/elmundomotor/2008/06/20/seguridad/1213963150.html.

[19] Opel insignia web-site, Last access: May 2009. http://www2.opel.de/insignia/.

[20] Lane departure system, Last acess: February 2009. http://www.iteris.com/ldws.aspx?q=10137&c=10136.

[21] Vislab, università degli studi di parma, Last acess: March 2009. http://vislab.it/.

[22] R. Ach, N. Luth, and A. Techmer. Real-time detection of traffic signs on a multi-core processor. *Intelligent Vehicles Symposium, 2008 IEEE*, pages 307–312, June 2008.

[23] B. Alefs, G. Eschemann, H. Ramoser, and C. Beleznai. Road sign detection from edge orientation histograms. *Intelligent Vehicles Symposium, 2007 IEEE*, pages 993–998, June 2007.

[24] J. An and T. Y. Choe. Dominant color transform and circular pattern vector for traffic sign detection and recognition(special section of papers selected from itc-cscc'97). *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 81(6):1128–1135, 1998.

[25] Y. Aoyagi and T. Asakura. A study on traffic sign recognition in scene image using genetic algorithms and neural networks. *Industrial Electronics, Control, and Instrumentation, 1996., Proceedings of the 1996 IEEE IECON 22nd International Conference on*, 3:1838–1843 vol.3, Aug 1996.

[26] P. Arnoul, M. Viala, J. Guerin, and M. Mergy. Traffic signs localisation for highways inventory from a video camera on board a moving collection van. *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE*, pages 141–146, Sep 1996.

[27] T. Asakura, Y. Aoyagi, and O. Hirose. Real-time recognition of road traffic sign in moving scene image using new image filter. *SICE 2000. Proceedings of the 39th SICE Annual Conference. International Session Papers*, pages 13–18, 2000.

[28] H. Austermeier, U. Büker, B. Mertsching, and S. Zimmermann. Analysis of traffic scenes by using the hierarchical structure code. pages 561–570, 1992.

[29] V. authors. Prevention of road traffic accidents. In *Twenty-seventh World Health Assembly*, Geneva, May 1974. World Health Organization. accesed 1 ocotober 2008.

[30] V. authors. United nations general assembly resolution a/res/58/9 on global road safety crisis. New York, November 2003.

[31] V. Authors. Use of event data recorder (edr) technology for highway crash data analysis. Technical report, National Cooperative Highway Research Program, 2004.

[32] V. Authors. Training system on new safety technologies for road transport adreessed to proffesional bodies of the automotive sector. Technical report, RASCC Automobile Club, 2007.

[33] V. authors. *Speed Management*. World Health Organization, Geneva, 2008.

[34] S. Azami, S. Katahara, and M. Aoki. Route guidance sign identification using 2-d structural description. *Intelligent Vehicles Symposium, 1996., Proceedings of the 1996 IEEE*, pages 153–158, Sep 1996.

[35] S. Backaitis. Economic consequences of traffic accidents in the baltic countries. *Lithuanian Quarterly Journal of Arts and Sciences*, 46, 2000.

[36] C. Bahlmann, Y. Zhu, V. Ramesh, M. Pellkofer, and T. Koehler. A system for traffic sign detection, tracking, and recognition using color, shape, and motion information. *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, pages 255–260, June 2005.

[37] N. Barnes and A. Zelinsky. Real-time radial symmetry for speed sign detection. *Intelligent Vehicles Symposium, 2004 IEEE*, pages 566–571, June 2004.

[38] F. Beainy, R. Akkari, and J. Sluss. Vehicle road sign controller system: A man-machine interface to reduce collisions. *Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE*, pages 1016–1021, 30 2007-Oct. 3 2007.

[39] M. Benallal and J. Meunier. Real-time color segmentation of road signs. *Electrical and Computer Engineering, 2003. IEEE CCECE 2003. Canadian Conference on*, 3:1823–1826 vol.3, May 2003.

[40] B. Besserer, S. Estable, and B. Ulmer. Multiple knowledge sources and evidential reasoning for shape recognition. *Computer Vision, 1993. Proceedings., Fourth International Conference on*, pages 624–631, May 1993.

[41] B. Besserer, S. Estable, B. Ulmer, and D. Reichqardt. Shape classification for traffic sign recognition. *Proc. 1st Int. Workshop on Intelligent Autonomous Vehicles (IFAC)*, pages 487–492., 1993.

[42] M. Betke and N. C. Makris. Recognition, resolution, and complexity of objects subject to affine transformations. *Int. J. Comput. Vision*, 44(1):5–40, 2001.

[43] A. Broggi, P. Cerri, P. Medici, P.and Porta, and G. Ghisio. Real time road signs recognition. pages 981–986, June 2007.

[44] C. Caraffi, E. Cardarelli, P. Medici, P. Porta, G. Ghisio, and G. Monchiero. An algorithm for italian de-restriction signs detection. pages 834–840, June 2008.

[45] J. P. Carrasco, A. de la Escalera, and J. M. Armingol. Driving supervision through traffic sign analysis. *Vehicular Electronics and Safety, 2008. ICVES 2008. IEEE International Conference on*, pages 242–247, Sept. 2008.

[46] D. Casasent and D. Psaltis. Position, rotation and scale invariant optical correlation. *Applied Optics*, 15:1795–1799., 1976.

[47] L. Csink, D. Paulus, U. Ahlrichs, and B. Heigl. Color normalization and object localization. pages 49–55, 1998.

[48] B. Cyganek. Recognition of road signs with mixture of neural networks and arbitration modules. pages 52–57, 2006.

[49] B. Cyganek. A real-time vision system for traffic signs recognition invariant to translation, rotation and scale. pages 278–289, 2008.

[50] Y. Damavandi and K. Mohammadi. Speed limit traffic sign detection and recognition. *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, 2:797–802, 2004.

[51] A. de la Escalera. *Visión por computador. Fundamentos y métodos.* 2001.

[52] A. de la Escalera, J. Armingol, J. Pastor, and F. Rodriguez. Visual sign information extraction and identification by deformable models for intelligent vehicles. *Intelligent Transportation Systems, IEEE Transactions on*, 5(2):57–68, June 2004.

[53] A. de la Escalera, J. Armingol, and M. Salichs. Traffic sign detection for driver support systems. *International Conference on Field and Service Robotics*, 2001.

[54] A. de la Escalera, L. Moreno, M. Salichs, and J. Armingol. Road traffic sign detection and classification. *Industrial Electronics, IEEE Transactions on*, 44(6):848–859, Dec 1997.

[55] de la Escalera A., A. J.M., and M. M. Traffic sign recognition and analysis for intelligent vehicles. *Image and Vision Computing*, 21:247–258, 2003.

[56] M. de Saint Blancard. Road sign recognition: a study of vision-based decision making for road environment recognition. pages 162–172, 1992.

[57] U. di Firenze. Lightweight neural network ++, Last access: March 2009. http://lwneuralnetplus.sourceforge.net/.

[58] P. Douville. Real-time classification of traffic signs. *Real-Time Imaging*, 6(3):185–193, 2000.

[59] V. Editors, editor. *World report on road traffic injury prevention*. World Health Organization, Geneva, 2004.

[60] L. W. Estevez and N. Kehtarnavaz. Real-time object specific recognition via raster scan video processing. *Real-Time Imaging*, 5(1):49–62, 1999.

[61] A. Etemadi. Robust segmentation of edge data. *Image Processing and its Applications, 1992., International Conference on*, pages 311–314, Apr 1992.

[62] D. E. Falcoff and L. R. Canali. Method for traffic-sign detection within a picture by color identification and external shape recognition. 3837:210–217, Aug. 1999.

[63] C.-Y. Fang, S.-W. Chen, and C.-S. Fuh. Road-sign detection and tracking. *Vehicular Technology, IEEE Transactions on*, 52(5):1329–1341, Sept. 2003.

[64] O. Faugeras. *Three-Dimensional Computer Vision*. MIT press., 1993.

[65] T. Fawcett. An introduction to roc analysis. *Pattern Recogn. Lett.*, 27(8):861–874, 2006.

[66] G. Finlayson, B. Schiele, and J. Crowley. Comprehensive colour image normalization. *Computer Vision – ECCV 98*, pages 475–490, 1998.

[67] L. Fletcher, L. Petersson, N. Barnes, D. Austin, and A. Zelinsky. A sign reading driver assistance system using eye gaze. *Robotics and Automation, 2005. ICRA 2005. Proceedings of the 2005 IEEE International Conference on*, pages 4655–4660, April 2005.

[68] H. Fleyeh. Color detection and segmentation for road and traffic signs. *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, 2:809–814, 2004.

[69] U. Franke, D. Gavrila, S. Gorzig, F. Lindner, F. Puetzold, and C. Wohler. Autonomous driving goes downtown. *Intelligent Systems and their Applications, IEEE*, 13(6):40–48, Nov/Dec 1998.

[70] K. Fukunaga. *Introduction to statistical pattern recognition (2nd ed.)*. Academic Press Professional, Inc., San Diego, CA, USA, 1990.

[71] X. Gao, K. Hong, P. Passmore, L. Podladchikova, and D. Shaposhnikov. Colour vision model-based approach for segmentation of traffic signs. *J. Image Video Process.*, 8(2):1–7, 2008.

[72] X. W. Gao, L. Podladchikova, D. Shaposhnikov, K. Hong, and N. Shevtsova. Recognition of traffic signs based on their colour and shape features extracted using human vision models. *J. Visual Communication and Image Representation*, 17(4):675–685, 2006.

[73] M. Garcia-Garrido, M. Sotelo, and E. Martin-Gorostiza. Fast traffic sign detection and recognition under changing lighting conditions. *Intelligent Transportation Systems Conference, 2006. ITSC '06. IEEE*, pages 811–816, 2006.

[74] D. Gavrila. Multi-feature hierarchical template matching using distance transforms. *Pattern Recognition, 1998. Proceedings. Fourteenth International Conference on*, 1:439–444 vol.1, Aug 1998.

[75] D. Gavrila and V. Philomin. Real-time object detection for 'smart' vehicles. *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1:87–93 vol.1, 1999.

[76] D. M. Gavrila. Real-time object detection using distance transforms. page 998, 1998.

[77] D. Ghica, S. W. Lu, and X. Yuan. Recognition of traffic signs by artificial neural network. *Neural Networks, 1995. Proceedings., IEEE International Conference on*, 3:1444–1449 vol.3, Nov/Dec 1995.

[78] P. Gil-Jimenez, H. Gomez-Moreno, P. Siegmann, S. Lafuente-Arroyo, and S. Maldonado-Bascon. Traffic sign shape classification based on support vector machines and the fft of the signature of blobs. *Intelligent Vehicles Symposium, 2007 IEEE*, pages 375–380, 2007.

[79] P. Gil-Jimenez, S. Lafuente-Arroyo, H. Gomez-Moreno, F. Lopez-Ferreras, and S. Maldonado-Bascon. Traffic sign shape classification evaluation. part ii. fft applied to the signature of blobs. *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, pages 607–612, June 2005.

[80] A. Hanbury and J. Serra. A 3d-polar coordinate colour representation suitable for image analysis. *Computer Vision and Image Understanding*, 2781:124–131, 2003.

[81] P. A. Hancock, M. Lesch, and L. Simmons. The distraction effects of phone use during a crucial driving maneuver. *Accident Analysis & Prevention*, 35(4):501 – 514, 2003.

[82] G. Hartmann. Hierarchical contour coding and generalization of shape. *International Conference on Robot Vision Sensory Controls*, pages 108–115, 1983.

[83] G. Hartmann. Recognition of hierarchically encoded images by technical and biological systems. *Biological Cybernetics*, 57:73– 84, 1987.

[84] T. Hibi. Vision based extraction and recognition of road sign region from natural colour image, by using hsl and co-ordinates transformation. *29th International Symposium on Automotive Technology and*

Automation, Robotics, Motion and Machine Vision in the Automotive Industries, ISATA, pages 201–206, 1996.

[85] W.-C. Huang, C.-H. Wu, and J. Irwan. Recognition of colorful objects in variant backgrounds and illumination conditions. *Industrial Technology, 1996. (ICIT '96), Proceedings of The IEEE International Conference on*, pages 849–853, Dec 1996.

[86] H. Ishida, T. Takahashi, I. Ide, Y. Mekada, and H. Murase. Identification of degraded traffic sign symbols by a generative learning method. *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, 1:531–534, 2006.

[87] G. Jacobs, A. Aeron-Thomas, and A. Astrop. *Estimating global road fatalities*. Crowthorne, 2000.

[88] J. Janet, M. White, T. Chase, R. Luo, and J. Sutton. Pattern analysis for autonomous vehicles with the region- and feature-based neural network: global self-localization and traffic sign recognition. *Robotics and Automation, 1996. Proceedings., 1996 IEEE International Conference on*, 4:3598–3604 vol.4, Apr 1996.

[89] F. Jeun-Haii and L. Gang. A vision-aided vehicle driving system: establishment of a sign finder system. *Vehicle Navigation and Information Systems Conference, 1994. Proceedings., 1994*, pages 33–38, Aug-2 Sep 1994.

[90] G.-Y. Jiang and T. Y. Choi. Robust detection of landmarks in color image based on fuzzy set theory. *Signal Processing Proceedings, 1998. ICSP '98. 1998 Fourth International Conference on*, 2:968–971 vol.2, 1998.

[91] D. Kang, N. Griswold, and N. Kehtarnavaz. An invariant traffic sign recognition system based on sequential color processing and geometrical transformation. *Image Analysis and Interpretation, 1994., Proceedings of the IEEE Southwest Symposium on*, pages 88–93, Apr 1994.

[92] N. Kehtarnavaz and A. Ahmad. Traffic sign recognition in noisy outdoor scenes. *Intelligent Vehicles '95 Symposium., Proceedings of the*, pages 460–465, Sep 1995.

[93] N. Kehtarnavaz, N. Griswold, and D. Kang. Stop-sign recognition based on color-shape processing. *Machine Vision and Applications*, 6:206–208, 1993.

[94] D. Kellmeyer and H. Zwahlen. Detection of highway warning signs in natural video images using color image processing and neural networks. *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference on*, 7:4226–4231 vol.7, Jun-2 Jul 1994.

[95] S. Kim. An efficient roadsign detection and recognition algorithm. *SPIE Conference on Enhanced and Synthetic Vision*, 3364:336–347, 1998.

[96] E. Kopits and M. Cropper. Traffic fatalities and economic growth. Technical report, The World Bank, Washington, 2003.

[97] A. Kouzani. Road-sign identification using ensemble learning. *Intelligent Vehicles Symposium, 2007 IEEE*, pages 438–443, June 2007.

[98] U. Kressel, F. Lindner, C. Wohler, and A. Linz. Hypothesis verification based on classification at unequal error rates. *Artificial Neural Networks, 1999. ICANN 99. Ninth International Conference on (Conf. Publ. No. 470)*, 2:874–879 vol.2, 1999.

[99] W.-J. Kuo and C.-C. Lin. Two-stage road sign detection and recognition. *Multimedia and Expo, 2007 IEEE International Conference on*, pages 1427–1430, July 2007.

[100] S. Lafuente-Arroyo, P. Gil-Jimenez, R. Maldonado-Bascon, F. Lopez-Ferreras, and S. Maldonado-Bascon. Traffic sign shape classification evaluation i: Svm using distance to borders. *Intelligent Vehicles Symposium, 2005. Proceedings. IEEE*, pages 557–562, June 2005.

[101] S. Lafuente-Arroyo, S. Maldonado-Bascon, P. Gil-Jimenez, J. Acevedo-Rodriguez, and R. Lopez-Sastre. A tracking system for automated inventory of road signs. *Intelligent Vehicles Symposium, 2007 IEEE*, pages 166–171, June 2007.

[102] S. Lafuente-Arroyo, S. Maldonado-Bascon, P. Gil-Jimenez, H. Gomez-Moreno, and F. Lopez-Ferreras. Road sign tracking with a predictive

filter solution. *IEEE Industrial Electronics, IECON 2006 - 32nd Annual Conference on*, pages 3314–3319, Nov. 2006.

[103] Y. Lauziere. Autonomous physics-based color learning under daylight. *SPIE Proceedings of 9th World Congress on Intelligent Transport Systems*, 3826:86–100, 1999.

[104] Y. Lauziere, D. Gingras, and F. Ferrie. A model-based road sign identification system. *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*, 1:1163–1170, 2001.

[105] G. Loy and N. Barnes. Fast shape-based road sign detection for a driver assistance system. *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, 1:70–75 vol.1, Sept.-2 Oct. 2004.

[106] R. C. Luo, H. Potlapalli, and D. Hislop. Translation and scale invariant landmark recognition using receptive field neural networks. *Intelligent Robots and Systems, 1992., Proceedings of the 1992 lEEE/RSJ International Conference on*, 1:527–533, Jul 1992.

[107] H. Makishita and K. Matsunaga. Differences of drivers' reaction times according to age and mental workload. *Accident Analysis & Prevention*, 40(2):567 – 575, 2008.

[108] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Siegmann, H. Gomez-Moreno, and F. Acevedo-Rodriguez. Traffic sign recognition system for inventory purposes. *Intelligent Vehicles Symposium, 2008 IEEE*, pages 590–595, June 2008.

[109] R. Malik, J. Khurshid, and S. Ahmad. Road sign detection and recognition using colour segmentation, shape analysis and template matching. *Machine Learning and Cybernetics, 2007 International Conference on*, 6:3556–3560, Aug. 2007.

[110] Ministerio de Fomento, Madrid, Spain. *Señalización vertical. Instrucción 8.1-IC*, 2000.

[111] J. Miura, T. Kanda, and Y. Shirai. An active vision system for real-time traffic sign recognition. *Intelligent Transportation Systems, 2000. Proceedings. 2000 IEEE*, pages 52–57, 2000.

[112] G. Mo and Y. Aoki. Recognition of traffic signs in color images. *TENCON 2004. 2004 IEEE Region 10 Conference*, B:100–103, Nov. 2004.

[113] S. Muller-Schneiders, C. Nunn, and M. Meuter. Performance evaluation of a real time traffic sign recognition system. *Intelligent Vehicles Symposium, 2008 IEEE*, pages 79–84, June 2008.

[114] C. Murray and A. Lopez, editors. *The Global Burden of Disease*. Harvard University Press, Boston, 1996.

[115] G. Nilsson. *Traffic safety dimensions and the power model to describe the effect of speed on safety*. Sweden, Lund Institute of Technology, lund university edition, 2004. Bulletin 221.

[116] L. Norman. Road traffic accidents: epidemiology, control and prevention. Technical report, World Health Organization, 1962.

[117] J.-C. Noyer, P. Lanvin, M. Yeary, and Y. Zhai. Sequential monte-carlo techniques and vision-based methods for road signs detection. *Instrumentation and Measurement Technology Conference Proceedings, 2007. IMTC 2007. IEEE*, pages 1–6, May 2007.

[118] H. L. Oei and P. H. Polak. Intelligent speed adaptation (isa) and road safety. *IATSS Res (Int Assoc Traffic Saf Sci)*, 26(2):45–51, 2002.

[119] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics*, 9(1):62–66, January 1979.

[120] P. Paclík and J. Novovičová. Road sign classification without color information. 2000.

[121] P. Paclík, J. Novovičová, P. Pudil, and P. Somol. Road sign classification using laplace kernel classifier. *Pattern Recogn. Lett.*, 21(13-14):1165–1173, 2000.

[122] D. Paulus and L. Csink. On color normalization. pages 222–229, 1997.

[123] D. Paulus, L. Csink, and H. Niemann. Color cluster rotation. *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on*, 1:161–165 vol.1, Oct 1998.

[124] D. Paulus and J. Hornegger. Pattern recognition of images and speeech in c++. 1997.

[125] E. Pérez and B. Javidi. Nonlinear distortion-tolerant filters for detection of road signs in background noise. *Vehicular Technology, IEEE Transactions on*, 51(3):567–576, May 2002.

[126] L. Petersson, L. Fletcher, N. Barnes, and A. Zelinsky. An interactive driver assistance system monitoring the scene in and out of the vehicle. *Robotics and Automation, 2004. Proceedings. ICRA '04. 2004 IEEE International Conference on*, 4:3475–3481 Vol.4, 26-May 1, 2004.

[127] G. Piccioli, E. D. Micheli, and M. Campani. A robust method for road sign detection and recognition. pages 495–500, 1994.

[128] G. Piccioli, E. D. Micheli, P. Parodi, and M. Campani. Robust method for road sign detection and recognition. *Image and Vision Computing*, 14(3):209 – 223, 1996.

[129] T. Pomierski and H. Gross. Verfahren zur empfindungsgemassen farbumstimmung (cambiar la dieresis). pages 473–480, 1995.

[130] L. Priese, J. Klieber, R. Lakmann, V. Rehrmann, and R. Schian. New results on traffic sign recognition. *Intelligent Vehicles '94 Symposium, Proceedings of the*, pages 249–254, Oct. 1994.

[131] L. Priese, R. Lakmann, and V. Rehrmann. Ideogram identification in a realtime traffic sign recognition system. *Intelligent Vehicles '95 Symposium., Proceedings of the*, pages 310–314, Sep 1995.

[132] L. Priese and V. Rehrmann. A fast hybrid color segmentation method. pages 297–304, 1993.

[133] S. Prince and R. Bergevin. Road sign detection and recognition using perceptual grouping. *International Symposium on Automotive Technology & Automation. Robotics, motion and machine vision in the automotive industries*, pages 163–170, 1997.

[134] B. Ran, H. Liu, and W. Martono. A vision-based object detection and recognition system for intelligent vehicles. *SPIE Conference on Sensing and Controls with Intelligent Transportation Systems*, 3525:326–336, 1998.

[135] V. Rehrmann and L. Priese. Fast and robust segmentation of natural color scenes. (1351):598–606, 1998.

[136] K. Rentschler. Extract from the gidas study for pedestrian safety, 2003. `http://www.google.com/search?hl=es&rls=com.microsoft%3Aes%3AIE-SearchBox&rlz=1I7GGLD_es&q=Karl+Rentschler+gidas&lr=`.

[137] W. Ritter. Traffic sign recognition in color image sequences. *Intelligent Vehicles '92 Symposium., Proceedings of the*, pages 12–17, Jun-1 Jul 1992.

[138] C. Schiekel. A fast traffic sign recognition algorithm for gray value images. pages 588–595, 1999.

[139] C. Seifert, L. Paletta, A. Jeitler, E. Hödl, J.-P. Andreu, P. M. Luley, and A. Almer. Visual object detection for mobile road sign inventory. pages 491–495, 2004.

[140] M. Shirvaikar. Automatic detection and interpretation of road signs. *System Theory, 2004. Proceedings of the Thirty-Sixth Southeastern Symposium on*, pages 413–416, 2004.

[141] Z. Shuang-dong, Z. Yi, and L. Xiao-feng. Detection for triangle traffic sign based on neural network. *Vehicular Electronics and Safety, 2005. IEEE International Conference on*, pages 25–28, Oct. 2005.

[142] Z. Shuangdong and J. Tian-tian. Intelligence approach of traffic sign recognition based on color standardization. *Vehicular Electronics and Safety, 2005. IEEE International Conference on*, pages 296–300, Oct. 2005.

[143] P. Siegmann, R. Lopez-Sastre, P. Gil-Jimenez, S. Lafuente-Arroyo, and S. Maldonado-Bascon. Fundaments in luminance and retroreflectivity measurements of vertical traffic signs using a color digital camera. *Instrumentation and Measurement, IEEE Transactions on*, 57(3):607–615, March 2008.

[144] G. Siogkas and E. Dermatas. Detection, tracking and classification of road signs in adverse conditions. *Electrotechnical Conference, 2006. MELECON 2006. IEEE Mediterranean*, pages 537–540, 0-0 2006.

[145] A. Soetedjo and K. Yamada. Fast and robust traffic sign detection. *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, 2:1341–1346 Vol. 2, Oct. 2005.

[146] A. Soetedjo and K. Yamada. Traffic sign classification using ring partitioned method. *IEICE Transactions*, 88-A(9):2419–2426, 2005.

[147] J. Torresen, J. Bakke, and L. Sekanina. Efficient recognition of speed limit signs. *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*, pages 652–656, Oct. 2004.

[148] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. 2001.

[149] P. Viola and M. Jones. Robust real-time object detection. *International Journal of Computer Vision*, 2002.

[150] S. Vitabile, V. Conti, F. Gennaro, and F. Sorbello. Efficient mlp digital implementation on fpga. *Digital System Design, 2005. Proceedings. 8th Euromicro Conference on*, pages 218–222, Aug.-3 Sept. 2005.

[151] S. Vitabile, A. Gentile, S. Siniscalchi, and F. Sorbello. Efficient rapid prototyping of image and video processing algorithms. *Digital System Design, 2004. DSD 2004. Euromicro Symposium on*, pages 452–458, Aug.-3 Sept. 2004.

[152] S. Vitabile, A. Gentile, and F. Sorbello. A neural network based automatic road signs recognizer. *Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, 3:2315–2320, 2002.

[153] S. Vitabile, G. Pollaccia, G. Pilato, and E. Sorbello. Road signs recognition using a dynamic pixel aggregation technique in the hsv color space. *Image Analysis and Processing, 2001. Proceedings. 11th International Conference on*, pages 572–577, Sep 2001.

[154] S. Vitabile and F. Sorbello. Pictogram road signs detection and understanding in outdoor scenes. *Proceedings of the SPIE Conference on Enhanced and Synthetic Vision 1998*, 3364:359–369, 1998.

[155] N. Yabuki, Y. Matsuda, Y. Fukui, and S. Miki. Region detection using color similarity. *Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium on*, 4:98–101 vol.4, Jul 1999.

[156] N. Yabuki, Y. Matsuda, H. Kimura, Y. Fukui, and S. Miki. Region extraction using color feature and active net model in color image (special section on selected papers from the 11th workshop on circuits and systems in karuizawa). *IEICE transactions on fundamentals of electronics, communications and computer sciences*, 82(3):466–472, 19990325.

[157] H. Yamauchi, A. Kojima, T. Miyamoto, H. Takahashi, and K. Fukunaga. A far road sign recognition by applying super-resolution to extracted regions from successive frames. *Mechatronics, ICM2007 4th IEEE International Conference on*, pages 1–5, May 2007.

[158] H.-M. Yang, C.-L. Liu, K.-H. Liu, and S.-M. Huang. Traffic sign recognition in disturbing environments. pages 252–261, 2003.

[159] S. Yang and M. Wang. Identification of road signs using a new ridgelet network. *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*, pages 3619–3622 Vol. 4, May 2005.

[160] L. Yi-Sheng, D. Der-Jyh, C. Shu-Yuan, L. Ru-Sheng, and H. Jun-Wei. Scale and skew-invariant road sign recognition. *International Journal of Imaging Systems and Technology*, 17:28–39, 2007.

[161] A. Yuille, D. Snow, and M. Nitzberg. Signfinder: using color to detect, localize and identify informational signs. *Computer Vision, 1998. Sixth International Conference on*, pages 628–633, Jan 1998.

[162] M. M. Zadeh, T. Kasv, and C. Y. Suen. localization and recognition of traffic signs for automated vehicle control systems.

[163] Z. Zhang. Flexible camera calibration by viewing a plane from unknown orientations. *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on*, 1:666 – 673, 1999.

[164] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, 2000.

[165] Y.-J. Zheng, W. Ritter, and R. Janssen. An adaptive system for traffic sign recognition. *Intelligent Vehicles '94 Symposium, Proceedings of the*, pages 165–170, Oct. 1994.

[166] S. Zhu, L. Liu, and X. Lu. Color-geometric model for traffic sign recognition. *Computational Engineering in Systems Applications, IMACS Multiconference on*, pages 2028–2032, Oct. 2006.

[167] T. Zin and H. Hama. Robust road sign recognition using standard deviation. *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on*, pages 429–434, Oct. 2004.