



Article Traffic Sign Detection System for Locating Road Intersections and Roundabouts: The Chilean Case

Gabriel Villalón-Sepúlveda¹, Miguel Torres-Torriti¹ and Marco Flores-Calero^{2,3,*}

- ¹ Departamento de Ingeniería Eléctrica, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Casilla 306-22, Santiago, Chile; procesosestocasticosespe@gmail.com (G.V.-S.); mtorrest@ing.puc.cl (M.T.-T.)
- ² Departamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas-ESPE,
 Av. Gral. Rumiñahui s/n, PBX 171-5-231B Sangolquí, Pichincha, Ecuador
- ³ Departamento de Sistemas Inteligentes, Tecnologías *I&H*, CP 050102, Latacunga, Cotopaxi, Ecuador
- * Correspondence: mjflores@espe.edu.ec; Tel.: +593-23-989400 (ext. 1850)

Academic Editor: Simon X. Yang Received: 2 January 2017; Accepted: 22 May 2017; Published: 25 May 2017

Abstract: This paper presents a traffic sign detection method for signs close to road intersections and roundabouts, such as stop and yield (give way) signs. The proposed method relies on statistical templates built using color information for both segmentation and classification. The segmentation method uses the RGB-normalized (ErEgEb) color space for ROIs (Regions of Interest) generation based on a chromaticity filter, where templates at 10 scales are applied to the entire image. Templates consider the mean and standard deviation of normalized color of the traffic signs to build thresholding intervals where the expected color should lie for a given sign. The classification stage employs the information of the statistical templates over YCbCr and ErEgEb color spaces, for which the background has been previously removed by using a probability function that models the probability that the pixel corresponds to a sign given its chromaticity values. This work includes an analysis of the detection rate as a function of the distance between the vehicle and the sign. Such information is useful to validate the robustness of the approach and is often not included in the existing literature. The detection rates, as a function of distance, are compared to those of the well-known Viola-Jones method. The results show that for distances less than 48 m, the proposed method achieves a detection rate of 87.5% and 95.4% for yield and stop signs, respectively. For distances less than 30 m, the detection rate is 100% for both signs. The Viola–Jones approach has detection rates below 20% for distances between 30 and 48 m, and barely improves in the 20-30 m range with detection rates of up to 60%. Thus, the proposed method provides a robust alternative for intersection detection that relies on statistical color-based templates instead of shape information. The experiments employed videos of traffic signs taken in several streets of Santiago, Chile, using a research platform implemented at the Robotics and Automation Laboratory of PUC to develop driver assistance systems.

Keywords: statistical template; traffic signs; color; road intersection; roundabouts; accidents

1. Introduction

Traffic accidents are the primary cause of death for young people between 15 and 29 years old. Between 20 to 50 million people are injured each year, while 1.3 million died due to traffic accidents, of which 91% take place in low and medium income countries [1–3]. Latin America is a region with high rates of road traffic accidents [2,4], due to diverse reasons which include driver education and behavior, law enforcement, and lack of adequate road infrastructure. However, technology can also play an important role in driver assistance systems that contribute to the alertness of the driver and better driving behaviors [5,6].

Most traffic accidents occur in urban areas, especially at road intersections and roundabouts. Country statistics for traffic accidents show that a significant number happened at road intersections, for instance, 22% in the USA [7,8], 58.7% in Japan during 1995 [9], 13.75% in Ecuador during 2015 [10], and 9.22% in Chile during 2014 [11]. Thus, the importance of developing systems for road intersections detection [12], which unlike other aspects such as pedestrian detection, lane-tracking and driver drowsiness or distraction [5,6] has not received enough attention. Prior work has focused on pavement segmentation to detect intersections [6] by analyzing the continuity and curvature of the road boundaries. However, occlusions in urban environments make the analysis of edges difficult or impossible. Therefore, the implementation of an advanced driving assistance system (ADAS) [13] requires a module capable of detecting road signs in general, and specifically those found at intersections.

We propose a traffic sign detection approach based on statistical templates built using normalized color information. The novelty of the approach lies in the probabilistic model of the sign (or object) conditioned over the intensity of the normalized color channels instead of using traditional shape descriptors. The results show that this approach is robust to variations in distance between the car and the traffic sign, as well as variation in illumination. Unlike deep learning techniques, the results show that it is possible to implement the proposed traffic sign detection approach with small datasets of a few hundred images.

This paper is organized as follows. First, the state-of-the-art of traffic sign detection algorithms is discussed in Section 2. Section 3 describes a new system for traffic sign detection and its modules based on color information. The experimental results, where an analysis between the detection rate and the distance is performed to verify the quality of this system, are presented in Section 4. Finally, Section 5 is devoted to the conclusions and discussion of future work.

2. State-of-the-Art

Traffic sign detection using visible-spectrum cameras may take different approaches. Some works implement feature classifiers. This means that a sliding-window method is used to compute features on different overlapping regions, which are then fed to the previously trained classifier [14,15]. The drawback of this strategy is that many positions and scales have to be tested using classifiers that may need computationally demanding training phases. More recent methods formulate a two-stage strategy, in which candidate or proposal regions are computed first by some "class-agnostic" segmentation process, i.e., extracting groups of pixels that share some characteristic without necessarily identifying whether they truly belong to the same class of object. In a second stage, some classification or decision process is used to complete the detection deciding whether some classes of objects sought are present or not. The methods proposed in [16–24] can be found among recent approaches for traffic sign detection employing regional proposal strategies together with classifiers. The most recent approaches to segmentation and classification employed in traffic sign detection are discussed next.

2.1. Segmentation for ROI Generation

In the context of traffic sign detection, blob generation and color analysis are the main techniques employed to segment regions of interest. Special efforts have been placed on making the color-based segmentation robust to large variations in illumination and weather conditions. Greenhalgh et al. [16] transform RGB into grayscale images using the red and the blue components and experimentally obtained thresholds to generate ROIs. Salti et al. [17] employ three color spaces derived from the RGB, the first to highlight road signs with a predominance of blue and red colors, the second one is for signs with intense red and the third one for the bright blue. Li et al. [18] have built the Gaussian space ($EE_{\lambda}E_{\lambda\lambda}$), where objects dominated by the green-red and blue-yellow colors are highlighted. The preselected regions are in turn transformed to normalized values $C_{\lambda} = E_{\lambda}/E$ and $C_{\lambda\lambda} = E_{\lambda\lambda}/E$, which are fed to a *k*-means clustering [25] to generate the ROIs. Zaklouta et al. [19] implement two RGB-based chromatic filters for ROIS generation, one for signs that have a red color prevalence, and another filter for red-yellow predominance; in both cases, thresholds are defined in terms of mean and variance. Lillo et al. [26] have used the L*a*b* space to detect signs in which the blue, green, yellow and red colors dominate. Based on the *k*-means clustering algorithm, the authors build a classifier that employs the a^* and b^* components. Fleyeh et al. [23] use the *H* and *S* components of the HSV space to train a classifier and implement the color segmentation that yields ROIs. More recently, Han et al. [24] have used the *H* component of the *HSI* space, in which the traffic signs are highlighted in order to build a grayscale image where a set of ROIs is generated. Keser et al. [27] have used the HSV filter intervals to generate a set of ROIs. Finally, Zhu et al. [28] employ three different object proposal strategies (Selective Search, Edge Boxes and BING) and convolutional neural networks for classification, achieving an accuracy of 88% on average.

2.2. Recognition

The recognition stage typically employs feature classifiers, and therefore requires a feature descriptor and a classification algorithm. One of the most popular feature descriptors is the histogram of oriented gradients (HOG) [15], which provides information about objects' shape. Recent works in traffic sign detection that employ the HOG descriptor include [16,17,19,29]. Li et al. [18] employ the PHOG descriptor, a variant of the HOG descriptor. Other descriptors are based upon the discrete Fourier transform [26], the Hough transform [23], the SURF method [30], the values of the neighboring pixels in a ROI [31], or predefined contour descriptors for basic shapes (circular, triangular, or rectangular) [27].

Concerning classifiers, most of the recent work in traffic sign detection employs SVM (support vector machine) classifiers [25]; see for example [16–19,26]. Another popular classification approach relies on artificial neural networks (NN). For example, recent work by Huang et al. [29] combines an NN-classifier with ELM (Extreme Learning Machine), and Pérez et al. [32] relies on MLPs (MultiLayer Perceptrons). The simpler *k*-NN (*k*-nearest neighbors) algorithm [25] is employed in the traffic sign detection method proposed in [24]. Recently, Deep Learning techniques are being used for simultaneous detection and recognition of traffic signs. CNN (Convolutional Neural Networks) is also employed in many of the most recent papers—Lau et al. [31], Jung et al. [33], Zeng et al. [34], Zhu et al. [28]—which propose new architectures for automatic sign detection. Other strategies, such as the one employed by Li and Yang [35] rely on a combination of DBM (Deep Boltzmann Machine) and CCA (Canonical Correlation Analysis) for feature extraction and classification. Lau et al. [31] have also experimented with *RBNN* (Radial Basis Neural Networks) for classification.

2.3. Databases

The main traffic sign databases correspond to the following countries: Germany [17,19,20,29,32], United Kingdom [16], Spain [22,26], Japan [36], China [28] or Malaysia [31]. Each country has its own regulations and standards concerning traffic signs, divided in regulatory, prevention and information categories [17,23,26,28,36]. Generally, they do not follow the Vienna Convention-Complaint for traffic signs [37].

Thus country-specific databases are required for the development of traffic sign detection systems. However, there is a lack of databases with traffic signs in Latin America. Therefore, another goal of this work is to contribute to the development of traffic sign detection systems that can be validated also on traffic signs of the Latin American region.

3. Proposed Approach for Segmentation and Recognition of Traffic Signs at Road Intersections and Roundabouts

The proposed computational strategy for detecting signs found at road intersections and roundabouts is composed of two parts. The first part generates ROIs in which traffic signs could be found calculating and analyzing color statistics in the normalized RGB space. The second part solves the recognition of signs in ROIs using a statistical template matching strategy using templates also in the normalized RGB space.

The detection must be done at the furthest possible distance, so that the driver has enough time to react and to stop in time. Examples of typical testing scenarios for the proposed approach are presented in Figure 1, which shows a distant stop sign and a yield (give way) sign.



Figure 1. Examples of testing scenarios showing stop (Left) and yield (Right) signs near a road intersection and a roundabout.

3.1. Chromaticity Filter for the Selection of ROIs

Under adequate illumination conditions, such as daylight or artificial lighting, the color of traffic signs is a feature that can be used to generate ROIs. Comparing the histograms of traffic signs in four color spaces, RGB, HSV, YCbCr and ErEgEb (the normalized RGB space) [38,39], it possible to observe in Figure 2 that some of the color spaces provide better discrimination capability between traffic signs and the image backgrounds. In Figure 2, the histograms labeled RPOS correspond to histograms of the stop sign, while the curves labeled RNEG correspond to the histograms of backgrounds or scenes that do not contain traffic signs.



(a) Histogram of red channel.



(d) Histogram of Y channel.



(b) Histogram of green channel.

80

70

20

10

× 60



(c) Histogram of blue channel.



(f) Histogram of Cb channel.

Figure 2. Cont.

(e) Histogram of Cr channel.

100



Figure 2. Histograms of the average Stop sign images (*c*POS) and their background (*c*NEG) for each channel *c* in the RGB, YCrCb, HSV and ErEgEb color spaces.

The color space that shows the smaller overlap between the histograms of positives (signs) and negatives (non-sign) is the ErEgEb space, where in particular the Er channel shows a clear separation between the two classes. The Eg channel histograms have a small intersection, but likewise it serves to discard a significant portion of negatives. Finally, the Eb does not provide considerable information, but will be considered a part of the classification strategy, see Figure 2j–l. A similar analysis was conducted for the yield sign and the results allow to conclude that the ErEgEb space, and in particular Er and Eg channels provide a better discrimination capability.

Therefore, the candidate regions of interest can be detected using a chromaticity filter, i.e., a filter that works on the variables that define color hue (dominant wavelength) and color purity or saturation (difference between the intensity of the dominant wavelength with respect to white, grey or black) regardless of luminance (magnitude of the color components vector) or psychological perception of illumination brightness or intensity (as an average of the components of the color vector). In other words, a chromaticity filter only requires two variables that describe dominant wavelengths regardless of the total energy by mapping the components of the thrichromacy color model into a subspace of two normalized values. Assuming a normal distribution of the chromaticity channels *Er* and *Eg*, the selection thresholds for extracting ROIs can be defined as intervals [$\mu_c - \alpha \sigma_c$, $\mu_c + \alpha \sigma_c$], c = Er, *Eg*, where μ_c and σ_c are the mean and standard deviation of the channel *c* computed over a set of positives (images with traffic signs) according to:

$$\mu_{i,c} = \frac{\sum_{p} I_{i,c}[p]}{N_{P}},\tag{1a}$$

$$\mu_c = \frac{\sum_i \mu_{i,c}}{N_I},\tag{1b}$$

$$\sigma_c^2 = \frac{\sum_i (\mu_{i,c} - \mu_c)^2}{N_I}.$$
 (1c)

where $I_{i,c}[p]$ is the value of channel *c* at the pixel location *p* within the traffic sign of the *i*-th reference image (positive), $i = 1, 2, ..., N_I$, N_I is the number of positive images, N_P is the number of pixels within the reference traffic sign area, $\mu_{i,c}$ is the mean value of channel *c* for the *i*-th image, μ_c is mean value of channel *c* across the ensemble of N_I images, and μ_c^2 is the variance of the mean values $\mu_{i,c}$, $i = 1, 2, ..., N_I$. The value α is set to minimize false positives and false negatives that will be passed to the recognition stage. A practical value that ensures the lowest amount of false positives while preventing misdetections was found to be $\alpha = 2$. It is to be noted that using the so-called summed integral area tables or integral images [14,40] is possible to reduce the computation time of the average values on $N \times N$ sliding blocks. Two examples showing the generation of preliminary candidates for ROIs using windows sizes of 50×50 and 10×10 pixels are shown in Figure 3 for values of $\alpha = 2, 3$.



(d) Original image.

(e) $10 \times 10, \pm 2\sigma$ (768 pre-candidates).

(f) 10 × 10, ±3 σ (33,589 pre-candidates).

Figure 3. Examples of ROI (regions of interest) candidates generated for the stop sign using the chromaticity filter before region merging.

The last step for the final proposal of ROIs is to eliminate the overpopulation of candidates. To this end, all the candidates that are contained within or are a sub-window of another candidate are discarded, so that only the largest block remains. Next, windows whose mean value is closest to μ_c in each neighborhood are selected so that there is only one candidate per neighborhood. Figure 4 shows the final ROI proposal obtained using 10 window sizes ranging from 10×10 to 50×50 in geometric progression, with a fixed scaling factor between each size. The number of preliminary candidates satisfying the chromaticity filter threshold was 32,849. This number is reduced to only 9 after the pre-candidate selection and merging step.



Figure 4. Example of ROIs obtained after removal/merging of overlapping pre-candidates selected with the chromaticity filter.

3.2. Recognition of Traffic Signs Based on Statistical Templates

The second stage of the proposed traffic sign detection approach is responsible for solving the identification of ROIs as traffic signs of a given type. To this end, a set of images is employed to create two statistical models, one with the mean intensity and the other with its standard deviation for each pixel belonging to the traffic sign. Testing on a sliding block for the percentage of pixels that fall within the expected intensity range for a given location provides a discriminator to detect traffic signs from background and non-traffic sign objects. A flow diagram of the proposed method is shown in Figure 5. The corresponding pseudocode of the algorithm is presented in Algorithm 1. The most effective recognition of traffic signs is achieved applying the algorithm to the Er and Eb channels of the ErEgEb color space. Only two channels conveying the chromaticity information are sufficient because the magnitude normalization satisfies Er + Eg + Eb = 1, making the third channel dependent on the other two (Eg = 1 - Er - Eb). The probabilistic model that defines the discrimination thresholds is discussed next.

In order to develop the probabilistic template matching model to recognize traffic signs, it is first convenient to introduce the following notation:

- $I \triangleq \{I[k], k = 1, .., n\}$ is a block or subwindow composed of pixel values I[k],
- I[k] is a vector with the pixel chromaticity components Er and Eb, k = 1, ..., n,
- $O \triangleq \{O[k], k = 1, ..., n\}$ is the object (sign) class label of subwindow *I*,
- O[k]: the object (sign) class label at every pixel k, k = 1, .., n in the subwindow I.

The probability that a window corresponds to a particular object (sign) is:

$$P(O \mid I) = 1 - P(\overline{O} \mid I[1], I[2], ..., I[n]),$$

= $1 - \prod_{k=1}^{n} P(\overline{O}[k] \mid I[k]),$
= $1 - \prod_{k=1}^{n} [1 - P(O[k] \mid I[k])],$ (2)

where \overline{O} is the complement of O and $P(\overline{O}[k] | I[1], ..., I[n]) = P(\overline{O}[k] | I[k])$ is obtained by assuming independence between $\overline{O}[k]$ and I[j], for all $j \neq k$. This is possible in view of the fact that O and I are random samples [41,42]. Also, this means that the probability of a pixel not belonging to an object only depends on the pixel value and not its neighborhood. In other words, the background (non-sign) pixels are assumed to be conditionally independent with respect to their neighborhood. This assumption is not entirely true in every area of the background, but simplifies the probability computation.



Figure 5. Flow chart of the recognition stage based on statistical templates.

Algorithm 1: Traffic sign recognition algorithm based on statistical templates

Input: *I*_c: candidate image block,

 α : pixel acceptance amplitude parameter, λ_{σ} : background pixels discard threshold, λ_{det} : minimal amount of pixels threshold for detection. **Output**: *Z*_{*det*}: binary detection output. // Loading pre-trained masks $A_M = LoadAverageMask();$ $\sigma_M = \text{LoadStandardDeviationMask()};$ // Pixel mask discarding corresponding to the background $B_M = \sigma_M Y < \lambda_\sigma;$ // Minimum and maximum accepted masks $MAX = \overline{A}_M + \alpha \times \sigma_M;$ $MIN = \overline{A}_M - \alpha \times \sigma_M;$ // Pixel mask accepted $P_M = (MIN \le I_c \le MAX) \times B_M;$ // Final decision if SumPixels(P_M)/SumPixels(B_M) $\geq \lambda_{det}$ then $Z_{det} = true;$ else $Z_{det} = false;$ end

In order to compute the posterior probability P(O[k] | I[k]) that a pixel *k* has label O[k] given its chromaticity value I[k], Bayes' theorem is used

$$P(O[k] | I[k]) = P(I[k] | O[k]) \frac{P(O[k])}{P(I[k])}$$
(3)

to express the posterior probability in terms of the measurement model P(I[k] | O[k]).

The measurement model P(I[k]|O[k]) can be obtained assuming that the pixel values of the object (sign) of interest follow a normal distribution $N(\mu_{Ex}, \sigma_{Ex})$, x = r, b, with mean chromaticity μ_{Ex} and standard deviation σ_{Ex} obtained from a set of reference images, see the last row of Figure 2.

The likelihood or conditional probability that the measured chromaticity values (Er[k], Eg[k]) at pixel k take some value in the interval [$Ex[k] - \beta\sigma_{Ex}[k]$, $Ex[k] + \beta\sigma_{Ex}[k]$], x = r, g, given object class, is then given by:

$$P(I[k] = (Er[k], Eb[k]) | O[k]) = \prod_{x=r,b} \frac{2}{\sqrt{\pi}} \int_{0}^{\frac{E_x[k] + \beta \sigma_{E_x}[k] - \mu_{E_x}[k]}{\sigma_{E_x}[k] \sqrt{2}}} e^{-t^2} dt$$
$$= \prod_{x=r,b} \operatorname{erf}\left(\frac{E_x[k] + \beta \sigma_{E_x}[k] - \mu_{E_x}[k]}{\sigma_{E_x}[k] \sqrt{2}}\right)$$
(4)

The prior probability P(O[k]) of finding the object (sign) in an image can be obtained experimentally from the set of reference images as:

$$P(O[k]) = \frac{TP}{TP + FP'},\tag{5}$$

where *TP* is the true positive rate and *FP* is the false positive rate for the object (sign) of class (type) *O*. This ratio is known as a positive predictive value and it describes the probability of traffic signs being correctly detected [43].

Finally, the probability of the occurrence of chromaticity values P(I[k] = (Er, Eb)) can be obtained empirically or analytically. The empirical approach would require constructing the histograms for the Er and Eb channels using a representative set of training data and normalizing the histograms to obtain the ratio between the number of pixels with the given chromaticity levels with respect to the the total amount of pixels. Analytically, a cumulative density function for the values of each channel can be deduced under the assumption that RGB values distribute according to a uniform distribution within a window block that contains both object and background pixels (see appendix for calculations details). The cumulative function for the chromaticity values under the uniform distribution assumption is given by:

$$F(y) = \begin{cases} 0 & y \le 0\\ \frac{y}{1-y} & 0 < y \le \frac{1}{3}\\ \frac{-21y^3 + 27y^2 - 9y + 1}{6y^2(1-y)} & \frac{1}{3} < y \le \frac{1}{2} \\ \frac{5y^2 + 2y - 1}{6y^2} & \frac{1}{2} < y < 1\\ 1 & y \ge 1 \end{cases}$$
(6)

The probability density function f(y) associated to the cumulative distribution of the chromaticity values I[k] = (Er, Eb) is easily obtained from (6) by deriving F(y) with respect to y. Figure 6 depicts the cumulative and density functions. The density function f(y) reaches a maximum at $y \sim 0.36$, which is the most likely background value without prior knowledge of the object (sign) class. Thus, it is desirable that the objects of interest have their chromaticity levels far away from this value.

The analysis thus far provides enough tools to compute the probability that a window corresponds to the object of interest using (2) and the statistical templates for the mean and standard deviation $\mu_{Ex}[k], \sigma_{Ex}[k]$ over the window block k = 1, 2, ..., n. However, these templates consider both the pixels of the object of interest and the background; therefore, to improve discrimination, it is convenient to discard background pixels.



Figure 6. Background probability model, cumulative density function F(y) (**a**), and probability density function f(y) (**b**), for *ErEgEb* space, considering that each channel of RGB space follows a uniform distribution.



Figure 7. Representative points of a stop sign template obtained by averaging reference samples in the *ErEgEb* space.

To this end, Figure 7 presents the representative points to the comparison of histograms of the pixels corresponding to the background and to the object of interest presented in Figure 8 reveals that the luminance channel Y spreads over the entire range of possible values for background objects. Hence, the standard deviation of the luminance channel is a good indicator to determine whether the pixel is part of the object of interest or part of the background. Figure 9 shows the mean and standard deviation for the luminance channel Y. It is clear that the template built using the variance provides higher contrast between background and foreground than using the mean value of the luminance to create a mask for discarding background regions.



(d)RGB-bottom background at (180,180). (e)YCrCb-bottom background at (180,180).(f)ErEgEb-bottom background at (180,180).

Figure 8. Cont.



Figure 8. Histograms in three color spaces for pixels in the reference areas of Figure 7.



Figure 9. Scaled images of the mean μ_Y (**a**) and standard deviation σ_Y (**b**) of the Y channel.

To create a mask for discarding background pixels, an adequate thresholding value for the standard deviation template is $\sigma_Y = 60$ as may be observed in Figure 10 since it allows to retain most of the pixels of the object of interest and discard all of the background. Assuming the luminance Y would distribute Gaussianly, $\sigma_Y > 60$ would imply that 95% of the samples would fall within ±120 intensity levels, thus it would cover a range of 240 levels, which would be almost the full range for an 8-bit image with 255 levels. Thus, higher values for the threshold on σ_Y are not convenient, while lower values cause part of the object to be labeled as background, as shown in Figure 10a with $\sigma_Y = 55$.



Figure 10. Discarding the background in the Y channel with different thresholds: (**a**) $\sigma_Y = 55$; (**b**) $\sigma_Y = 60$; and (**c**) $\sigma_Y = 65$.

By the foregoing, the channels to be used in the recognition stage have the luminance to remove the background and the chromaticity values Er and Eb to confirm the ROIs are of a given traffic sign, ensuring robustness to illumination variations.

4. Testing Methodology and Experimental Results

4.1. Perception and Processing Systems

The vehicle shown in Figure 11a was employed as a testing platform for the experiments. The system comprises several cameras (visible spectrum, IR, catadioptric), an IMU and a RTK GPS; see [6] for further details. For the experiments presented here, only one camera from Imaging Source, model DFK31BF03 model (see Figure 11b) was used together with the RTK GPS from Navcom Technology, model SF-2050, with decimeter positioning accuracy (see Figure 11c). The camera has a resolution of 1024×768 pixels and delivers images at 30 fps, while the GPS sampling rate is 10 Hz. Thus, the GPS data was interpolated to match the sampling instants of the camera. Registering the position is important in order to compute detection rates as a distance function. The processing of the images was carried out on a PC with an Intel Core 2 Duo processor with a 2.0 GHz frequency, and 3.5 GB of RAM. All the algorithms were implemented in C++ using the OpenCV version 2.2 library.



Figure 11. Experimental platform, vehicle (a); camera (b); and GPS (c).

4.2. Training and Validation Dataset

The training dataset contains 2567 negative examples, 122 images containing stops signs and 80 images containing yield signs. The positive samples were randomly rotated, scaled and translated in order to produce 7000 positive examples; see Figure 12. The validation dataset contains 273 stop signs and 447 yield signs captured in six different driving runs.

The datasets consider a sequence of signs as the car approaches different intersections in the city of Santiago, Chile. Both datasets contain traffic signs in real driving conditions, under varying illumination and partial occlusion.

The database has been made available at RAL [44] to contribute to the making of new studies and the development of ADAS.



Figure 12. Typical traffic signs of road intersections and roundabouts in Chile: stop and yield signs; under different lighting conditions (sunny (**a**,**b**), normal (**c**,**d**) and dark (**e**,**f**)) and observer positions.

4.3. Experiments Employing the Viola–Jones Method and the Proposed Statistical Template Approach

In this section, the proposed traffic sign detection approach based on statistical templates is tested and compared to the well-know Viola–Jones method [45].

4.3.1. Viola–Jones Method:

The detection rates of the Viola–Jones approach for the stop and yield signs are summarized in Table 1 as a function of distance. The detection rate for stop signs is 100% when distances are 20 m or less. However, the detection rate rapidly falls to 0% for distances above 35 m. On the other hand, less that 3.1% of the yield signs was detected for distances below 20 m. For distances above 20 m, it was not possible to detect any yield sign. This is attributed, in part, to the fact that some samples in the dataset contained signs with marks and graffiti on it. These results show that the Viola–Jones approach is highly sensitive to possible modifications in the signs sought. The false alarm rates of the Viola–Jones approach for the stop and yield signs were practically 0%, as shown in Table 2.

Distance to the Intersection [mt]	Yield %	Stop %
>62	0.0%	0.0%
62–55	0.0%	0.0%
55-48	0.0%	0.0%
48–41	0.0%	0.0%
41–34	0.0%	6.5%
34–27	0.0%	21.0%
27–20	0.0%	57.6%
<20	3.1%	100.0%

Table 1. Detection rate based on the Viola–Jones method.

Table 2. False alarm rate per frame.

Method	Yield	Stop
Viola–Jones	0.006	0.0
Statistical template	0.036	0.069

4.3.2. Statistical Template Method:

The proposed algorithm was executed employing eight different templates sized from 14×14 to 50×50 pixels in geometric progression. The detection rates presented in Table 3 show that a high rate of success was achieved for distances up to 37 m in the case of the stop sign and 35 in the case of the yield sign. Detection rates fall to 0% for distances above 52 m in the case of the stop sign and 48 m in the case of the yield sign. Unlike the Viola–Jones approach, false alarm rates were between 3.6% and 6.9% as shown in Table 2. A comparison of the detection rate performance of the proposed method with the Viola–Jones approach is presented in Figure 13. This figure shows the effectiveness of the proposed approach at detecting traffic signs earlier than the Viola–Jones method does.

Tables 1 and 3 were constructed with approximately 430 images for each distance class.

In terms of the computational effort, the Viola–Jones method required 450 ms per frame while the proposed approach required 950 ms per frame. This amount could be decreased for real-time operation using dedicated graphic processing units (GPUs).



Figure 13. Comparison of detection rates versus distance between the Viola–Jones method and the statistical templates method for the stop and yield signs.

Distance to the Intersection [mt]	Yield %	Stop %
>62	0.0%	0.0%
62–55	0.0%	5.2%
55–48	8.5%	28.1%
48-41	50.0%	82.2%
41–34	87.3%	94.7%
34–27	100.0%	100.0%
27–20	100.0%	100.0%
<20	100.0%	100.0%

Table 3. Detection rate based on the statistical template method.

Finally, Figure 14 presents an extended example of the proposed system, in real driving conditions during the day, for both stop and yield signs.



Figure 14. Example of the proposed system in different instants of time, in daytime conditions. Stop sign (**Top**) and Yield sign (**Bottom**), where blue indicates the ROIs and red shows the true sign.

5. Conclusions

This paper presented an approach based on statistical templates computed from chromaticity and luminance values for traffic sign detection near road intersections and roundabouts. The approach is evaluated using a dataset of stop and yield (give way) signs from Chile and compared to the well-known Viola–Jones classification method.

The proposed approach is divided into two stages. The first stage is a segmentation stage based on the chromaticity filter applied to the Er and Eg channels of the ErEgEb color space (the normalized RGB color space). In the second stage, the chromaticity filter returns candidate regions that are responsible for the recognition of traffic signs and which must be tested. The selection of the Er and Eg channels is based on the analysis of the histograms of the color components in four color spaces (RGB, YCrCB, HSV and ErEgEb), which shows that the ErEgEb provides greater capacity to discriminate candidate traffic signs. The segmentation stage employs a selection threshold that can be computed automatically from a reference dataset. The recognition stage is based on a statistical template built with information of the *Er* and *Eb* chromaticity channels and the *Y* luminance channel. The luminance channel is employed to create a traffic sign mask from the variance of pixels that allows background regions within a ROI to be discarded. The Er and Eb channels are used to compute a statistical template that provides the sign selection thresholds. A probabilistic model and probability distribution function were derived to construct a recognition process based on Bayesian inference.

The results obtained show that the proposed approach has higher detection rates than the Viola–Jones method. The experiments considered the evaluation of the detection rate at different distances as the vehicle approaches a traffic sign. On average, the proposed approach exhibits a detection rate of 87.5% for yield signs and 95.4% for stop signs at distances below 48 m.

The two main advantages of the proposed approach are summarized in that it does not require a computationally expensive training or calibration stage, and it is not sensitive to changes in illumination, partial occlusion or marks drawn on the traffic sign. Future work will consider extending the proposed approach to the detection of traffic lights at road junctions and crosswalks, as well as to the detection of other signs not necessarily found at road intersections. Ongoing research is considering joint analysis of lane geometry analysis and edge continuity together with traffic sign detection to improve the detection of road intersections. The main limitation of this model is the computing time, which is about 950 ms per frame. As a future work, we will work to reduce this processing time. **Acknowledgments:** This project has been supported by Comision Nacional de Ciencia y Tecnología de Chile (Conicyt) Grant 11060251, Basal Project FB0008, and by Universidad de las Fuerzas Armadas ESPE through both Plan de Movilidad con Fines de Investigción 2015 Grant and 2014 - PIT - 007 Research Project, and by Tecnologías I&H.

Author Contributions: The framework was proposed by Miguel Torres, and further development and implementation were realized by Gabriel Villalón. Marco Flores and Miguel Torres studied some of the ideas, analyzed the experiment results and prepared the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Deduction of the Background Probability Distribution

To derive the cumulative density function of chromaticity values for any pixel F(y) in Equation (6), it is assumed that each channel of the RGB color spaces follows a uniform distribution. Thus, three random variables X_1 , X_2 and X_3 representing each channel are defined according to the uniform distribution:

$$X_1, X_2, X_3 \sim U(0, S) \Rightarrow f(x) = \begin{cases} 0 & x < 0 \lor x \ge S \\ \frac{1}{S} & 0 \le x < S \end{cases},$$
 (A1)

The cumulative distribution function of the chromaticity channel $E_{X_1} = X_1/(X_1 + X_2 + X_3)$, given by:

$$F(y) = P(Y \le y) = P\left(\frac{X_1}{X_1 + X_2 + X_3} \le y\right),$$
(A2)

is calculated as follows. The procedure for E_{X_2} or E_{X_3} is exactly the same.

In finding a closed-form expression for (A2), the following cases are considered:

• Case y = 0:

$$F(0) = P\left(\frac{X_1}{X_1 + X_2 + X_3} \le 0\right)$$

= $P(X_1 \le 0) = 0.$ (A3)

• Case *y* = 1:

$$F(1) = P\left(\frac{X_1}{X_1 + X_2 + X_3} \le 1\right)$$

= $P(X_2 + X_3 \ge 0) = 1.$ (A4)

• Case 0 < *y* < 1:

$$F(y) = P\left(\frac{X_1}{X_1 + X_2 + X_3} \le y\right) \\ = P\left(X_1 \le \frac{y}{1 - y}(X_2 + X_3)\right).$$
(A5)

The last case amounts to solving the following integral:

$$F(y) = \int_{0}^{S} f(x_3) \int_{0}^{S} f(x_2) \int_{0}^{\frac{y}{1-y}(x_2+x_3)} f(x_1) dx_1 dx_2 dx_3.$$
(A6)

The inner integral of $f(x_1)$ must consider that if $\frac{y}{1-y}(x_2+x_3) \ge S$, then density function $f(x_1) = \frac{1}{5}$ is integrated up to *S*, while if $\frac{y}{1-y}(x_2+x_3) < S$, then the integral is computed for $x_1 \in [0, \frac{y}{1-y}(x_2+x_3)]$.

For the latter to be fulfilled throughout the non-zero domain of $f(x_2)$ and $f(x_3)$, the following must be satisfied:

$$\frac{y}{1-y}(x_2+x_3) \le S \ \forall x_2, x_3 \in [0,S] \Rightarrow y \le \frac{1}{3}$$
(A7)

Similarly, when analyzing the limits of integration for the integrals over x_2 and X_3 , so that the non-null parts are integrated and the result is continuous over the range of integration, the following intervals are obtained for y: $(0, \frac{1}{3}], (\frac{1}{3}, \frac{1}{2}]$ and $(\frac{1}{2}, 1)$.

For $0 < y \le \frac{1}{3}$:

$$F(y) = \int_{0}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{\frac{y}{1-y}(x_{2}+x_{3})} \int_{0}^{\frac{y}{1-y}(x_{2}+x_{3})} \frac{1}{S} dx_{1} dx_{2} dx_{3} = \frac{y}{1-y'},$$
(A8)

For $\frac{1}{3} < y \leq \frac{1}{2}$, the integral is calculated as follows:

$$F(y) = \int_{0}^{\frac{1-y}{y}S-S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{\frac{y}{1-y}(x_{2}+x_{3})} \frac{1}{S} dx_{1} dx_{2} dx_{3}$$

$$+ \int_{\frac{1-y}{y}S-S}^{S} \frac{1}{S} \int_{0}^{\frac{1-y}{y}S-x_{3}} \frac{1}{S} \int_{0}^{\frac{y}{1-y}(x_{2}+x_{3})} \frac{1}{S} dx_{1} dx_{2} dx_{3}$$

$$+ \int_{\frac{1-y}{y}S-S}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} dx_{1} dx_{2} dx_{3}$$

$$= \frac{-21y^{3} + 27y^{2} - 9y + 1}{6y^{2}(1-y)}, \qquad (A9)$$

Finally, for $\frac{1}{2} < y < 1$, the integral is computed as follows:

$$F(y) = \int_{0}^{\frac{1-y}{y}S} \frac{1}{S} \int_{0}^{\frac{1-y}{y}S-x_{3}} \frac{1}{S} \int_{0}^{\frac{y}{1-y}(x_{2}+x_{3})} \frac{1}{S} dx_{1} dx_{2} dx_{3}$$

$$+ \int_{0}^{\frac{1-y}{y}S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} dx_{1} dx_{2} dx_{3}$$

$$+ \int_{\frac{1-y}{y}S}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} \int_{0}^{S} \frac{1}{S} dx_{1} dx_{2} dx_{3}$$

$$= \frac{5y^{2} + 2y - 1}{6y^{2}}.$$
(A10)

The above results in the cumulative distribution presented in Equation (6).

Appendix B. Traffic Sign Detection by Using the Viola-Jones Method

The Viola–Jones object recognition approach [14,40] has been used in multiple computer vision applications [45–47]. In this work, it has been used to build a traffic signs detector, trained for stop and yield signs. Figure A1 shows the first and second convolution masks based on Haar-like features for each of the traffic signs considered in this work.



Figure A1. First and second Haar-like convolution masks for stop and yield signs employed in this work.

References

- 1. World Health Organization (WHO). Road Traffic Injuries. Available online: http://www.who.int/violence_injury_prevention/road_traffic/en/ (accessed on 24 May 2015).
- World Health Organization (WHO). La OMS y la FIA aúnan Esfuerzos Para Mejorar La Seguridad Vial. Available online: http://www.who.int/mediacentre/news/releases/2003/pr11/es/ (accessed on 24 May 2015).
- 3. World Health Organization (WHO). Lesiones Causadas Por el Tránsito. Available online: http://www.who. int/mediacentre/factsheets/fs358/es/ (accessed on 24 May 2015).
- 4. Fraser, B. Traffic accidents scar Latin America's roads. *Lancet* 2005, 366, 703–704.
- 5. Jiménez-Pinto, J.; Torres-Torriti, M. Optical Flow and Driver's Kinematics Analysis for State of Alert Sensing. *Sensors* **2013**, *13*, 4225–4257.
- 6. Tapia-Espinoza, R.; Torres-Torriti, M. Robust Lane Sensing and Departure Warning under Shadows and Occlusions. *Sensors* **2013**, *13*, 3270–3298.
- 7. National Highway Traffic Safety Administration (NHTSA). Traffic Fatalities Up Sharply in 2015. Available online: http://www.nhtsa.gov/ (accessed on 24 May 2015).
- Mesriani Law Group. Accidents Caused by Dangerous Intersections. Available online: http://www.hg.org/ article.asp?id=7652 (accessed on 24 May 2015).
- 9. Wang, Y.; Nihan, N.L. *Quantitative Analysis on Angle-Accident Risk at Signalized Intersections*; Research Associate Department of Civil Engineering University of Washington: Seattle, WA, USA, 2015.
- 10. Agencia Nacional de Tránsito del Ecuador. Estadísticas de Transporte Terrestre Y Seguridad Vial. Available online: http://www.ant.gob.ec/ (accessed on 24 May 2015).
- 11. CONASET. Observatorio de Datos De Accidentes. Available online: https://estadconaset.mtt.gob.cl/ (accessed on 24 May 2015).
- Nie, Y.; Chen, Q.; Chen, T.; Sun, Z.; Dai, B. Camera and lidar fusion for road intersection detection. In Proceedings of the IEEE Symposium on Electrical and Electronics Engineering, Kuala Lumpur, Malaysia, 24–27 June 2012; pp. 273–276.
- Horgan, J.; Hughes, C.; McDonald, J.; Yogamani, S. Vision-Based Driver Assistance Systems: Survey, Taxonomy and Advances. In Proceedings of the IEEE 18th International Conference on Intelligent Transportation Systems (ITSC 2015), Las Palmas de Gran Canaria, Spain, 15–18 September 2015; pp. 2032–2039.
- Viola, P.; Jones, M. Rapid object detection using a boosted cascade of simple features. In Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2001), Kauai, HI, USA, 8–14 December 2001; Volume 1, pp. I-511–I-518.
- Dalal, N.; Triggs, B. Histograms of oriented gradients for human detection. In Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, USA, 20–25 June 2005; Volume 1, pp. 886–893.
- 16. Greenhalgh, J.; Mirmehdi, M. Real-Time Detection and Recognition of Road Traffic Signs. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 1498–1506.

- 17. Salti, S.; Petrelli, A.; Tombari, F.; Fioraio, N.; DiStefano, L. Traffic sign detection via interest region extraction. *Pattern Recogn.* **2015**, *48*, 1039–1049.
- 18. Li, H.; Sun, F.; Liu, L.; Wang, L. A novel traffic sign detection method via color segmentation and robust shape matching. *Neurocomputing* **2015**, *169*, 77–88.
- 19. Zaklouta, F.; Stanciulescu, B. Real-Time Traffic-Sign Recognition Using Tree Classifiers. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 1507–1514.
- 20. Zaklouta, F.; Stanciulescu, B. Real-time traffic sign recognition in three stages. *Robot. Auton. Syst.* **2014**, 62, 16–24.
- 21. Mogelmose, A.; Trivedi, M.M.; Moeslund, T.B. Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey. *IEEE Trans. Intell. Transp. Syst.* **2012**, *13*, 1484–1497.
- 22. Carrasco, J. Advanced Driver Assistance System Based on Computer Vision Using Detection, Recognition and Tracking of Road Signs. Ph.D. Thesis, Laboratorio de Sistemas Inteligentes, Universidad Carlos III de Madrid, Madrid, Spain, 2009.
- 23. Fleyeh, H.; Biswas, R.; Davami, E. Traffic sign detection based on AdaBoost color segmentation and SVM classification. In Proceedings of the 2013 IEEE EUROCON, Zagreb, Croatia, 1–4 July 2013; pp. 2005–2010.
- Han, Y.; Virupakshappa, K.; Oruklu, E. Robust traffic sign recognition with feature extraction and k-NN classification methods. In Proceedings of the 2015 IEEE International Conference on Electro/Information Technology (EIT), Dekalb, IL, USA, 21–23 May 2015; pp. 484–488.
- 25. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 2nd ed.; Springer: New York, NY, USA, 2009.
- 26. Lillo, J.; Mora, I.; Figuera, C.; Rojo, J.L. Traffic sign segmentation and classification using statistical learning methods. *Neurocomputing* **2015**, *1*, 286–299.
- 27. Keser, T.; Kramar, G.; Nozica, D. Traffic Signs Shape Recognition Based on Contour Descriptor Analysis. In Proceedings of the IEEE International Conference on Smart Systems and Technologies (SST), Osijek, Croatia, 12–14 October 2016.
- Zhu, Z.; Liang, D.; Zhang, S.; Huang, X.; Li, B.; Hu, S. Traffic-Sign Detection and Classification in the Wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 27–30 June 2016.
- 29. Huang, Z.; Yu, Y.; Gu, J. A Novel Method for Traffic Sign Recognition based on Extreme Learning Machine. In Proceedings of the IEEE 11th World Congress on Intelligent Control and Automation (WCICA), Shenyang, China, 29 June–4 July 2014; pp. 1451–1456.
- Bay, H.; Ess, A.; Tuytelaars, T.; Gool, L.V. Speeded-Up Robust Features (SURF). *Comput. Vis. Image Underst.* 2008, 110, 346–359.
- Lau, M.M.; Lim, K.H.; Gopalai, A.A. Malaysia Traffic Sign Recognitio on with Convolutional Neural Network. In Proceedings of the IEEE International Conference on Digital Signal Processing (DSP), Singapore, 21–24 July 2015; pp. 1006–1010.
- Perez-Perez, S.E.; Gonzalez-Reyna, S.E.; Ledesma-Orozco, S.E.; Avina-Cervantes, J.G. Principal component analysis for speed limit Traffic Sign Recognition. In Proceedings of the 2013 IEEE International Autumn Meeting on Power Electronics and Computing (ROPEC), Morelia, Mexico, 13–15 November 2013; pp. 1–5.
- Jung, S.; Lee, U.; Jung, J.; Shim, D.H. Real-time Traffic Sign Recognition system with deep convolutional neural network. In Proceedings of the Ubiquitous Robots and Ambient Intelligence (URAI), Xian, China, 19–22 August 2016.
- 34. Zeng, Y.; Xu, X.; Shen, D.; Fang, Y.; Xiao, Z. Traffic Sign Recognition Using Kernel Extreme Learning Machines With Deep Perceptual Features. *IEEE Trans. Intell. Transp. Syst.* **2016**, *PP*, 1–7.
- Li, C.; Yang, C. The research on traffic sign recognition based on deep learning. In Proceedings of the 2016 16th IEEE International Symposium on Communications and Information Technologies (ISCIT), Qingdao, China, 26–28 September 2016; pp. 156–161.
- Nguyen, B.T.; Shim, J.; Kim, J.K. Fast Traffic Sign Detection under Challenging Conditions. In Proceedings of the 2014 International Conference on Audio, Language and Image Processing (ICALIP), Shanghai, China, 7–9 July 2014; pp. 749–752.
- Convention on Road Signs and Signals. Available online: http://www.unece.org/fileadmin/DAM/trans/ conventn/signalse.pdf (accessed on 24 May 2015)

- Flores, M.; Armingol, M.; Escalera de la, A. New probability models for face detection and tracking in color images. In Proceedings of the 2007 IEEE International Symposium on Intelligent Signal Processing (WISP 2007), Alcala de Henares, Spain, 3–5 October 2007; pp. 1–6.
- 39. Jain, A.K.; Li, S.Z. Handbook of Face Recognition; Springer-Verlag Inc.: Secaucus, NJ, USA, 2005; p. 117.
- 40. Viola, P.; Jones, M. Robust real-time face detection. In Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV 2001), Vancouver, BC, Canada, 7–14 July 2001.
- 41. Duda, R.; Hart, P.; Stork, D. Pattern Classification; Wiley: New York, NY, USA, 2001; p. 42.
- 42. Casella, G.; Berger, R. Statistical Inference; Duxbury: Pacific Grove, CA, USA, 2002; p. 207.
- 43. Heston, T.F. Standardizing predictive values in diagnostic imaging research. *J. Magn. Reson. Imaging* **2011**, 2, 506–507.
- 44. Robotics and Automation Laboratory, School of Engineering, Pontificia Universidad Católica de Chile. Available online: http://ral.ing.puc.cl/datasets/intersection/index.htm (accessed on 24 May 2015).
- 45. Viola, P.A.; Jones, M.J.; Snow, D. Detecting pedestrians using patterns of motion and appearance. *Int. J. Comput. Vis.* **2005**, *63*, 153–161.
- 46. Flores, M. Sistema Avanzado de Asistencia a la Conducción Mediante Visión por Computador para la Detección de la Somnolencia. Ph.D. Thesis, Laboratorio de Sistemas Inteligentes, Universidad Carlos III de Madrid, Madrid, Spain, 2009.
- Houben, S.; Stallkamp, J.; Salmen, J.; Schlipsing, M.; Igel, C. Detection of Traffic Signs in Real-World Images: The German Traffic Sign Detection Benchmark. In Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA, 4–9 August 2013; pp. 205–216.



 \odot 2017 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).