

Sistema de detección de señales de tráfico para la localización de intersecciones viales y frenado anticipado

Gabriel Villalón-Sepúlveda, Miguel Torres-Torriti^a, Marco Flores-Calero^{b,*}

^aDepartamento de Ingeniería Eléctrica, Pontificia Universidad Católica de Chile, Vicuña Mackenna 4860, Casilla 306, Correo 22, Santiago, Chile 782 – 0436.

^bDepartamento de Eléctrica y Electrónica, Universidad de las Fuerzas Armadas-ESPE, Av. Gral. Rumiñahui s/n, PBX 171-5-231B, Sangolquí, Pichincha, Ecuador.

Resumen

En este trabajo se presenta un sistema de detección de señales de tráfico aledañas a intersecciones viales y rotundas, y un análisis para conocer su capacidad de detección en función de la distancia. El método propuesto está basado en la segmentación por color sobre el espacio RGB-normalizado (*ErEgEb*) para la generación de regiones de interés (ROIs), la clasificación del tipo de señal usando los espacios *YCbCr* y *ErEgEb* para construir una plantilla estadística, donde para eliminar el fondo se ha propuesto una función de distribución de probabilidad que modeliza el color de los objetos de interés. El sistema está especializado en dos tipos de señales: disco Pare y triángulo invertido de Ceda el Paso. Luego se analiza su desempeño, al compararlo con el método de Viola & Jones, donde se verifica la capacidad de detección a diferentes distancias. Para distancias menores a 48 metros, el método propuesto ha logrado una tasa de detección del 87.5% en el caso de la señal de Ceda el Paso, y del 95.4% en el caso del disco Pare, mientras que para distancias inferiores a 30 metros la detección lograda es del 100%. Estos resultados son superiores a los reportados en la mayoría de la literatura existente, la cual a diferencia del presente trabajo no aborda la sensibilidad de los métodos propuestos ante la variación de la distancia a las señales de tráfico aún cuando esto es uno de los principales desafíos en la implementación de sistemas de asistencia al conductor aplicables al mundo real. Los experimentos han sido desarrollados sobre una base de datos de señales de tráfico, generadas a partir de imágenes tomadas en varias calles de la ciudad de Santiago, Región Metropolitana, Chile, usando un vehículo experimental diseñado para desarrollar sistemas inteligentes.

Palabras Clave: Intersección vial, accidentes, señales de tráfico, plantillas estadísticas, distancia, color, Chile.

1. Introducción

Los accidentes de tráfico causan al año 1.3 millones de muertes y entre 20 y 50 millones de heridos en el mundo, siendo la primera causa de muerte de jóvenes entre 15 y 29 años de edad. El 91% de ellos ocurren en países de bajo y mediano ingresos (World Health Organization WHO, 2015c,a,b).

La mayoría de los accidentes de tráfico se producen en zonas urbanas, de ellas, las *intersecciones* son especialmente peligrosas, es así que, en USA el 22% (National Highway Traffic Safety Administration, 2015; Mesriani Law Group, 2015), en Ecuador, en el año 2015, el 13,75% (Agencia Nacional de Tránsito del Ecuador, 2015) de todos los accidentes sucedieron en estos sitios. En Chile, en el año 2014, el 9,22% fueron por desobediencia a la señalización (CONASET, 2014).

Esto ocurre debido a la presencia de puntos ciegos, a una inadecuada señalización, a carriles mal implementados y a la

falta de concentración del conductor. Consecuentemente, la detección de intersecciones viales es importante para planificar una ruta de navegación segura (Nie et al., 2012) y poder conducir con la suficiente confianza.

Por lo tanto, el principal objetivo de este trabajo es desarrollar un sistema avanzado de asistencia a la conducción (SAAC) (Horgan et al., 2015) para detectar automáticamente las señales de tráfico próximas a las intersecciones viales, en las áreas urbanas, durante el día, utilizando las tecnologías de visión por computador e inteligencia artificial. Es por esto que se propone un método capaz de detectar y alertar sobre la inminencia de un cruce de calles o una rotonda para que el conductor disminuya su velocidad o frene, según corresponda.

Este documento está organizado de la siguiente manera. La primera sección corresponde a la introducción y a la motivación que han generado esta investigación. La siguiente presenta el estado del arte en la detección de señales de tráfico. A continuación, el apartado tres describe un nuevo sistema para la detección de señales de tráfico junto con sus distintos módulos. La siguiente sección exhibe los resultados experimentales, donde se incluye un análisis para verificar la calidad de la propuesta

* Autor en correspondencia

Correos electrónicos: mtorrest@ing.puc.cl (Gabriel Villalón-Sepúlveda, Miguel Torres-Torriti), mjflores@espe.edu.ec (Marco Flores-Calero)

en función de la distancia. Finalmente, la última parte esta dedicada a las conclusiones y los trabajos futuros.

2. Estado del arte

Los vehículos modernos incorporan cada vez más SAACs, es así que los sistemas destinados a la detección de señales de tráfico toman importancia (Nguyen et al., 2014) para mejorar la seguridad y para ganar mercado frente a sus competidores.

Para la detección de señales de tráfico se suele seguir una estrategia compuesta de dos partes, la segmentación (también conocida como detección) y el reconocimiento (Greenhalgh and Mirmehdi, 2012; Salti et al., 2015; Li et al., 2015; Zaklouta and Stanciulescu, 2012, 2014; Mogelmose et al., 2012; Carrasco, 2009; Fleyeh et al., 2013; Han et al., 2015).

En el primer caso, la característica predominante es el color, donde se han utilizado varios espacios de color y distintas técnicas de visión por computador para generar regiones con alta probabilidad de contener una señal de tráfico. En el segundo escenario, se han utilizado algunos métodos de aprendizaje-máquina para clasificar las señales por su forma o para conocer su contenido.

2.1. Segmentación

En este caso, todas las técnicas basadas en el color, buscan robustez frente a las variaciones de iluminación. Así Greenhalgh et al. (Greenhalgh and Mirmehdi, 2012), a partir de la imagen en *RGB*, han construido una nueva imagen en escala de grises, utilizando las componentes roja y azul, luego bajo ciertos umbrales, generados experimentalmente, obtienen las ROIs.

Salti et al. (Salti et al., 2015) han utilizado tres espacios de color derivados de *RGB*, el primero para resaltar las señales de tráfico con predominancia de los colores azules y rojos, el segundo es para las señales con rojo intenso y el tercero para los azules vivos. Li et al. (Li et al., 2015) han construido el espacio Gaussiano $EE_{\lambda}E_{\lambda\lambda}$, donde resaltan los objetos dominados por los colores azul-amarillo y verde-rojo. Posteriormente, vuelven a generar el espacio $C_{\lambda} = E_{\lambda}/E$ y $C_{\lambda\lambda} = E_{\lambda\lambda}/E$, sobre el cual, utilizando el algoritmo *k – means* (Hastie et al., 2009) construyen un método de clasificación por color para la generación de ROIs. Zaklouta et al. (Zaklouta and Stanciulescu, 2012) han construido un filtro cromático a partir de *RGB* para las señales con predominancia roja, y otro para las señales con supremacía de rojo-amarillo, en ambos casos usan umbrales estimados con la media y la varianza de las ROIs. Lillo et al. (Lillo et al., 2015) han utilizado los espacios $L^*a^*b^*$ y *HSI* para detectar señales donde predominan los colores rojo, blanco y amarillo. Con el algoritmo *k – means* (Hastie et al., 2009) han construido un clasificador usando las componentes a^* y b^* . Fleyeh et al. (Fleyeh et al., 2013) han usado las componentes H y S del espacio *HSV* para entrenar el clasificador AdaBoost y construir un método de segmentación por color para la generación de ROIs. Finalmente, Han et al. (Han et al., 2015) han usado la componente H del espacio *HSI*, donde han generado un intervalo, donde resaltan las señales de tráfico, para construir una imagen en grises donde se localizan las ROIs.

2.2. Reconocimiento

Esta etapa se divide en dos partes: *i*) elección del algoritmo de clasificación y, *ii*) método de extracción de características.

En el primer caso, entre los algoritmos preferidos están: SVM (Hastie et al., 2009), usado en los trabajos de Greenhalgh et al. (Greenhalgh and Mirmehdi, 2012), Salti et al. (Salti et al., 2015), Zaklouta et al. (Zaklouta and Stanciulescu, 2012), Li et al. (Li et al., 2015), (Lillo et al., 2015), Lillo et al. (Fleyeh et al., 2013), *k-NN* (Hastie et al., 2009) en las investigaciones de Han et al. (Han et al., 2015), redes neuronales artificiales, empleadas por Huang et al. (Huang et al., 2014) con el caso *ELM*, Perez et al. (Perez-Perez et al., 2013) con la implementación *MLP* y Lau et al. (Lau et al., 2015) experimentando con *CNN* y *RBNN*.

En la segunda cuestión se tiene una amplia variedad de propuestas. Así Greenhalgh et al. (Greenhalgh and Mirmehdi, 2012), Salti et al. (Salti et al., 2015) y Zaklouta et al. (Zaklouta and Stanciulescu, 2012), Huang et al. (Huang et al., 2014) han utilizado el descriptor *HOG* (Dalal, 2006). Li et al. (Li et al., 2015) han usado el descriptor *PHOG*, que es una variación de *HOG*. Lillo et al. (Lillo et al., 2015) han implementado la extracción de características usando la transformada discreta de Fourier. Fleyeh et al. (Fleyeh et al., 2013) han implementado su método con la transformada de Hough. Han et al. (Han et al., 2015) utilizaron el método *SURF* (Baya et al., 2008). Perez et al. (Perez-Perez et al., 2013) usaron *PCA* para la reducción de la dimensión y la elección de características. Finalmente, Lau et al. (Lau et al., 2015) usaron una ponderación de los píxeles vecinos de la ROI.

Las tasas de detección de los distintos métodos se presentan en la tabla 1. En ninguno de estos trabajos se menciona a que distancia se generaron los resultados.

2.3. Bases de datos

Las principales bases de datos corresponden a países como Alemania (Salti et al., 2015; Zaklouta and Stanciulescu, 2012, 2014; Huang et al., 2014; Perez-Perez et al., 2013), Reino Unido (Greenhalgh and Mirmehdi, 2012), España (Lillo et al., 2015; Carrasco, 2009), Japón (Nguyen et al., 2014) o Malaysian (Lau et al., 2015). En las señales de tráfico predominan los colores rojo, amarillo, verde y azul; y las formas circular, triangular, cuadrada, rectangular, flechas, diamante y semi-circular. Asimismo cada país tiene sus propias normativas en cuanto a la señalética de tráfico, divida en las categorías de reglamentación, prevención e información (Lillo et al., 2015; Fleyeh et al., 2013; Nguyen et al., 2014; Salti et al., 2015).

Esto demuestra que existe la necesidad de contribuir con información de señales de tráfico de los países de América Latina, para desarrollar este tipo de investigación y porque esta zona mantiene altas cifras de accidentes de tráfico (Fraser, 2005; World Health Organization WHO, 2015c).

3. Detección y reconocimiento de señales de tráfico en intersecciones viales

3.1. Descripción del sistema

El sistema se compone de dos partes, una dedicada a la detección de ROIs, y otra destinada al reconocimiento de señales

Tabla 1: Tasas de reconocimiento en la detección de señales de tráfico.

Autor	TR ^a	Color ^b	Forma señal ^c
(Greenhalgh and Mirmehdi, 2012)	92.1	r, a	T, C, Ti, R
(Zaklouta and Stan-ciulescu, 2012, 2014)	97.0	r	T, C
(Perez-Perez et al., 2013)	98.9	r	C
(Fleyeh et al., 2013)	98.4	r	C, T
	97.0 ^d		Ti
(Nguyen et al., 2014)	92.5	r, a	Ti, C
(Salti et al., 2015)	78.21	r, a	C, T,
	82.13		Ti, R
	72.78		
(Li et al., 2015)	94.5 ^e	r, a, am	T, C, Ti, D
(Lillo et al., 2015)	97.0	r, b, am	Todas
(Han et al., 2015)	97.54	r, am	C, T, Ti, y D
(Lau et al., 2015)	99.0	r, a, am	C, T, D, Ti, R

^aTR: Tasa de reconocimiento en %.

^bPrincipal color de las señales de tráfico, rojo (r), azul (a), amarillo (am) y blanco (b).

^cTriangular (T), circular (C), triángulo invertido (Ti), rectangular (R), cuadrado (Ca), diamante (D), indicación de curva (Ic) y restricción (Rs).

^dPeligro y Prohibición, respectivamente

^eProhibición, Mandatorias y Peligro, respectivamente

aledañas a las intersecciones viales, bajo el siguiente esquema.

- Estudio del espacio de color (predominio del rojo) para la detección de señales de tráfico y generación de ROIs
- Estudio de las características de forma para la clasificación de señales de tráfico

Debido a la naturaleza del problema, la detección debe realizarse a una distancia suficiente para que el conductor pueda reaccionar y frenar anticipadamente. En la figura 1 se observa el escenario de actuación de esta propuesta, donde el disco **Pare** y el triángulo invertido de **Ceda el Paso** son los protagonistas.

3.2. Detección de señales de tráfico de intersecciones viales

3.2.1. Análisis y elección del espacio de color para detectar señales de tráfico de intersecciones viales

El color, en las señales de tráfico, es una característica intrínseca que puede ser utilizada para generar ROIs. Los espacios de color utilizados son: *HSV*, *YCbCr*, *RGB* y *ErEgEb* (Jain and Li, 2005; Flores et al., 2007).

Para estudiar y comparar el desempeño de los espacios de color se consideraron dos grupos de imágenes, uno con señales de tráfico y otro sin ellas. Sobre cada uno de éstos se calculó el valor promedio, y posteriormente, se construyeron los histogramas por cada canal de color. Los resultados, para el caso del disco PARE, se presentan en la figura 2.

El espacio de color que manifiesta menor traslape entre los histogramas de positivos y negativos es el espacio *RGB* normalizado, donde particularmente el canal *Er* muestra una clara separación entre ambas clases, ver figura 2 (j-l). *Eg* tiene una pequeña intersección, pero de igual manera sirve para separar las clases involucradas. Finalmente, *Eb* no brinda información considerable, pero será utilizado como parte de la estrategia de clasificación.

De manera análoga se realizó este procedimiento para el caso de la señal Ceda el Paso obteniéndose el mismo espacio de color.

3.2.2. Filtro por cromaticidad para la generación de ROIs

Para fijar los umbrales del filtro selector por cromaticidad, se asume una distribución normal y se fijan los umbrales de acuerdo a la media (μ) y la desviación estándar (σ_μ). Para calcular estos parámetros se usa (1) y (2):

$$\mu_i = \frac{\sum_{px} I[px]}{N_p},$$

$$\mu = \frac{\sum_i \mu_i}{N_I}, \quad (1)$$

$$\sigma_\mu^2 = \frac{\sum_i (\mu_i - \mu)^2}{N_I}. \quad (2)$$

donde px es la posición del píxel en la imagen, N_I es el número de imágenes positivas, N_p es el número de píxeles de la imagen i y μ_i es el promedio de un canal para todos sus píxeles, $i = 1, \dots, N_I$.

De este modo se desechan todos los candidatos cuyos valores promedio no se encuentran entre los rangos $\mu \pm \alpha\sigma_\mu$, por lo que de α dependerá cuantos positivos y falsas alarmas pasarán a la siguiente etapa.

Posteriormente, utilizando ventanas de tamaño fijo se realiza un barrido sobre toda la imagen buscando aquellas zonas que cumplan el umbral de cromaticidad entre $\mu - 2\sigma_\mu$ y $\mu + 2\sigma_\mu$, donde se ha fijado $\alpha = 2$. Para acelerar el cómputo se ha utilizado la imagen integral, propuesta por Viola & Jones (Viola and Jones, 2001a,b). En la figura 3 se muestran las ROIs en los casos 50×50 píxeles y 10×10 píxeles.

El último paso para la preselección es eliminar la sobre-población de candidatos. Primero se descartan todos los candidatos que están contenidos o sean una subventana de otro candidato, para que sólo permanezcan las ventanas más grandes preseleccionadas. Luego de esto, en cada vecindad, se busca la ventana cuyo promedio se encuentre más cercano a μ , con lo que se conservan sólo un candidato por vecindad.

En la figura 4 se muestra una preselección con 10 ventanas de distintos tamaños, desde 10×10 hasta 50×50 en progresión geométrica, es decir, con un factor de escalamiento fijo entre cada tamaño, con lo que de 32849 candidatos que cumplen con los umbrales de media, quedaron sólo 9 para ser procesados con mayor detalle en una siguiente etapa.

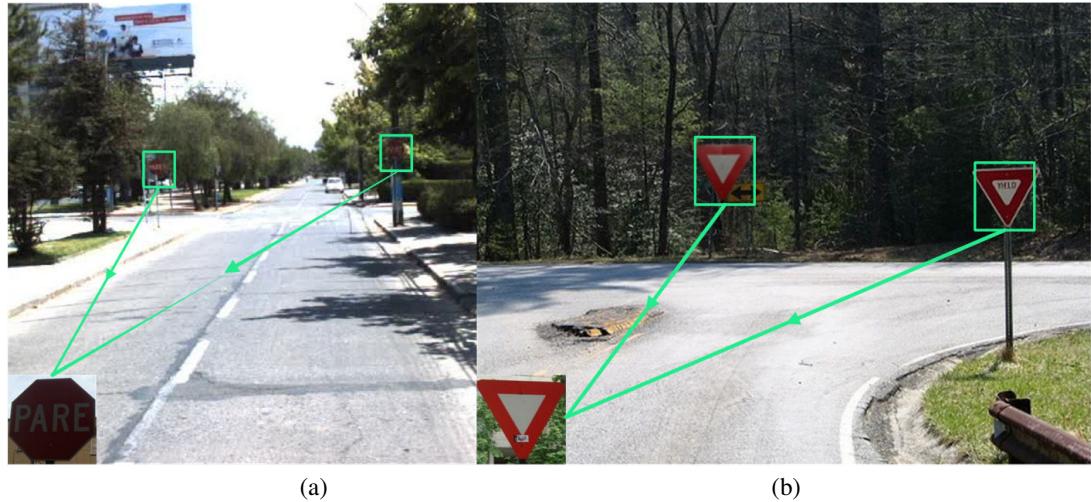


Figura 1: Ejemplo del sistema de detección y reconocimiento de señales de tráfico en intersecciones viales propuesto en esta investigación.



Figura 4: ROIs obtenidas luego del filtro de pre-selección.

3.3. Reconocimiento de señales de tráfico en intersecciones viales y rotundas

La etapa de reconocimiento es utilizada para confirmar que una ROI es verdaderamente una señal de tráfico y posteriormente verificar su tipo, disco Pare o Ceda el Paso.

El método propuesto sigue el algoritmo 1. A partir de las imágenes de entrenamiento se crean dos modelos estadísticos, uno con los promedios y otro con la desviación estándar, que definen los valores máximos y mínimos para cada píxel en cada posición. Luego, si un porcentaje alto de estos píxeles se encuentran dentro de este umbral, el método arroja una detección verdadera.

Además, es importante distinguir el área que ocupa el objeto buscado y el área que ocupa el fondo dentro de la ventana, de modo de no tomar en cuenta estos últimos píxeles. Esto se logra utilizando un umbral de desviación estándar en los modelos

Algoritmo 1: Detección por comparación con plantilla estadística.

Input: I_c : Sub-imagen candidata,
 α : parámetro de amplitud de aceptación por píxel,
 λ_σ : umbral de descarte de píxeles de fondo,
 λ_{det} : umbral de razón mínima de píxeles aceptados para detección.

Output: Z_{det} : Coeficiente binario de detección.

```

// Carga de máscaras previamente entrenadas
 $\bar{A}_M$  = LoadAverageMask();
 $\sigma_M$  = LoadStandardDeviationMask();
// Máscara de descarte de píxeles
// correspondientes al fondo
 $B_M$  =  $\sigma_M \cdot Y < \lambda_\sigma$ ;
// Máscaras de umbrales mínimos y máximos
// aceptados
 $MAX = \bar{A}_M + \alpha * \sigma_M$ ;
 $MIN = \bar{A}_M - \alpha * \sigma_M$ ;
// Máscara de píxeles aceptados
 $P_M = (MIN \leq I_c \leq MAX) * B_M$ ;
// Decisión final
if SumPixels( $P_M$ )/SumPixels( $B_M$ )  $\geq \lambda_{det}$  then
  |  $Z_{det} = true$ ;
else
  |  $Z_{det} = false$ ;
end

```

entrenados, permaneciendo sólo aquellos píxeles que mostraron baja variabilidad de sus niveles en el entrenamiento.

Normalizando el tamaño de todas las imágenes positivas se analiza el histograma de píxeles representativos que se muestra en la figura 5 para el caso del disco PARE de tamaño 200×200 píxeles.

Nuevamente el espacio $ErgEb$ logra ser más robusto para

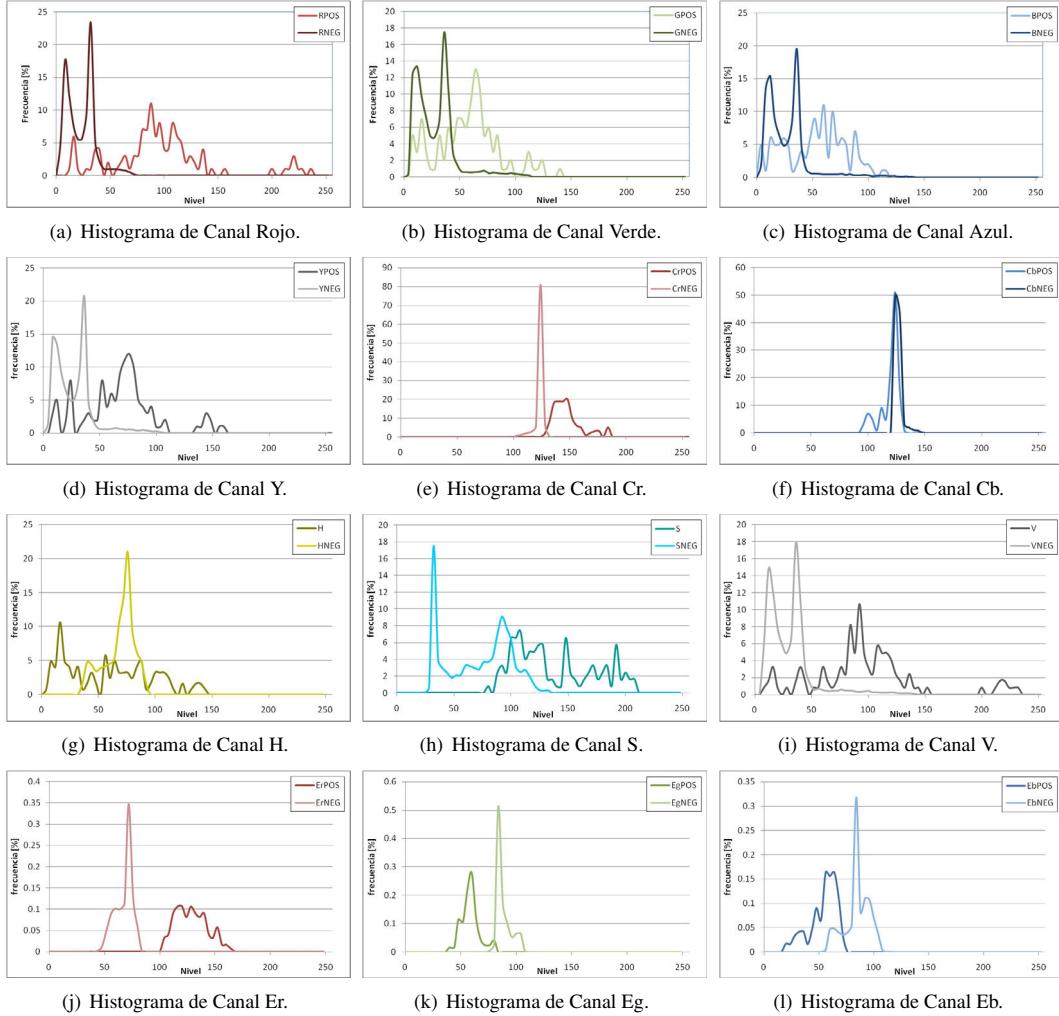


Figura 2: Histogramas de los promedios de las imágenes del disco PARE y del fondo.

Figura 5: Puntos representativos seleccionados sobre una plantilla de imágenes de disco PARE promediadas en el espacio $ErEgEb$.

distinguir entre un píxel del objeto y del fondo (o no parte del objeto). Dado que las operaciones son píxel a píxel, no es necesario contar con los 3 canales, ya que el tercero se puede deducir de $Er + Eg + Eb = 1$.

Se desea detectar ventanas dentro de una imagen que tengan alta probabilidad de ser el objeto buscado, para lo cuál se define:

- $I \triangleq \{I[k], k = 1, \dots, n\}$ subventana definido de acuerdo al valor de sus píxeles $I[k]$,
- $I[k]$: vector con nivel de $ErEgEb$ de píxel $k = 1, \dots, n$,
- $O \triangleq \{O[k], k = 1, \dots, n\}$ corresponde a la clase de objeto de la subventana I ,
- $O[k]$: indica la clase a la cual pertenece el píxel $k = 1, \dots, n$.

La probabilidad de que la ventana corresponda a un determinado objeto es:

$$\begin{aligned}
 P(O | I) &= 1 - P(\bar{O} | I[1], I[2], \dots, I[n]), \\
 &= 1 - \prod_{k=1}^n P(\bar{O}[k] | I[k]), \\
 &= 1 - \prod_{k=1}^n [1 - P(O[k] | I[k])], \tag{3}
 \end{aligned}$$



Figura 3: Generación de ROIs, usando ventanas de tamaño 50×50 , original (a), $\pm 2\sigma$ (1.507 candidatos) (b), $\pm 3\sigma$ (7.843 candidatos) (c), usando ventanas de tamaño 10×10 , original (d), $\pm 2\sigma$ (768 candidatos) (e), $\pm 3\sigma$ (33.589 candidatos) (f)

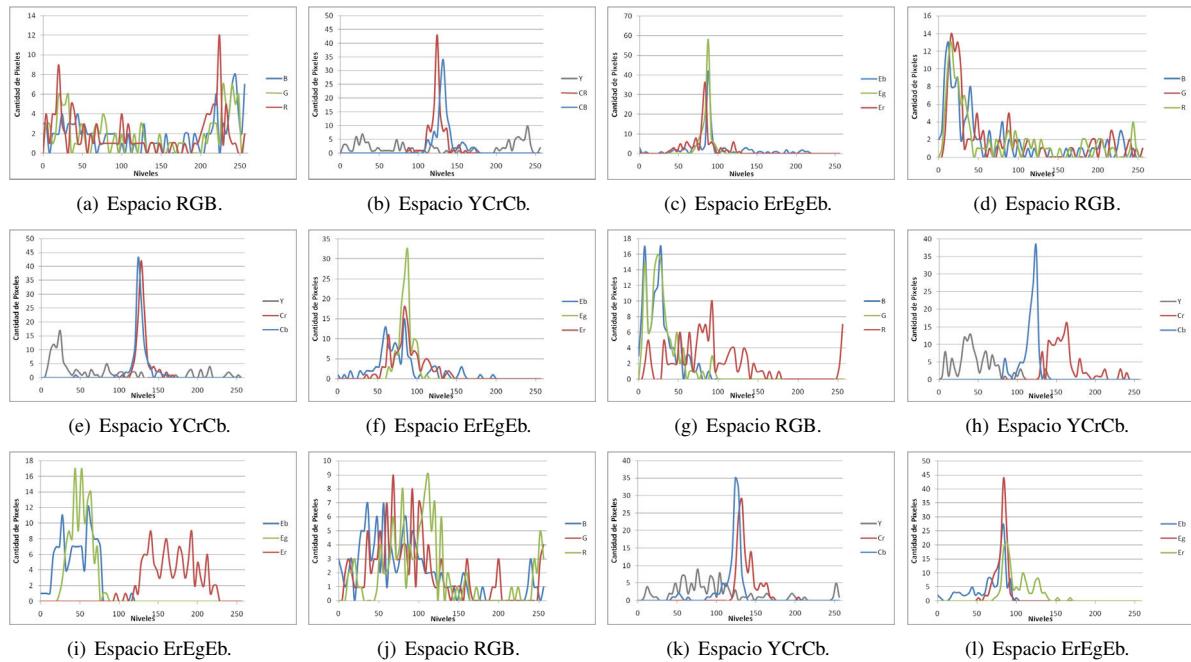


Figura 6: Histogramas en varios espacios de color del píxel correspondiente: al fondo superior en la posición (20, 20) (a-c), al fondo inferior en la posición (180, 180) (d-f), a una zona roja del objeto en la posición (50, 50) (g-i), a una zona blanca del objeto en la posición (22, 84) (j-l).

donde $P(\bar{O}[k] | I[1], \dots, I[n]) = P(\bar{O}[k] | I[k])$, asumiendo independencia entre $\bar{O}[k]$ e $I[j]$, para todo $j \neq k$, es decir la probabilidad de un píxel a no pertenecer a un objeto, solo depende del valor del píxel y no de su vecindad, lo cuál simplifica el cálculo.

Asumiendo una distribución normal en los niveles de cromatidad de los píxeles dentro de una imagen correspondiente

a un positivo, por lo que se crea una plantilla con las medias y otra con las desviaciones estándar para cada posición de píxel, con lo que se cuenta con la probabilidad de que cada píxel tenga un nivel determinado, siendo que pertenece a un objeto positivo:

$$\begin{aligned}
P(I[k] = (er, eb) | O[k]) &= \\
&= \prod_{x=r,b} \frac{1}{\sqrt{\pi}} \left[\int_0^{\frac{e_x - \mu_{Ex}[k] + 0.5}{\sigma_{Ex}[k] \sqrt{2}}} e^{-t^2} dt - \int_0^{\frac{e_x - \mu_{Ex}[k] - 0.5}{\sigma_{Ex}[k] \sqrt{2}}} e^{-t^2} dt \right] \\
&= \prod_{x=r,b} \left[\operatorname{erf}\left(\frac{e_x - \mu_{Ex}[k] + 0.5}{\sigma_{Ex}[k] \sqrt{2}}\right) - \operatorname{erf}\left(\frac{e_x - \mu_{Ex}[k] - 0.5}{\sigma_{Ex}[k] \sqrt{2}}\right) \right]
\end{aligned} \tag{4}$$

donde $\mu_{Ex}[k]$ y $\sigma_{Ex}[k]$ son respectivamente el valor medio y la desviación estándar entrenados para el píxel k en el canal Ex y erf es la función error (Abramowitz and Stegun, 1972).

Utilizando probabilidad condicional es posible obtener la probabilidad de que un píxel, de acuerdo a sus niveles, pertenezca o no al objeto buscado según:

$$P(O[k] | I[k]) = P(I[k] | O[k]) \frac{P(O[k])}{P(I[k])} \tag{5}$$

donde $P(O[k])$ corresponde a la probabilidad de encontrar el objeto en una imagen. Debido a que se cuenta con un clasificador previo a la detección, dicha probabilidad de puede calcular como:

$$P(O[k]) = \frac{TP}{TP + FP}, \tag{6}$$

donde TP y FP son los verdaderos y falsos positivos, respectivamente, y se pueden obtener empíricamente.

La probabilidad $P(I[k] = (er, eb))$ corresponde a la probabilidad de los niveles (er, eb) de aparecer en una imagen cualquiera. Esta probabilidad se puede obtener de dos formas, empírica o analítica. Empírica realizando un histograma de un conjunto representativo de capturas en el espacio $ErEgEb$, donde equivale a la razón entre la cantidad de píxeles con dicho nivel y el total de píxeles. Analíticamente se puede deducir la función de probabilidad acumulada dada por (7) (ver el anexo para el detalle de los cálculos):

$$F(y) = \begin{cases} 0 & y \leq 0 \\ \frac{y}{1-y} & 0 < y \leq \frac{1}{3} \\ \frac{-21y^3 + 27y^2 - 9y + 1}{6y^2(1-y)} & \frac{1}{3} < y \leq \frac{1}{2} \\ \frac{5y^2 + 2y - 1}{6y^2} & \frac{1}{2} < y < 1 \\ 1 & y \geq 1 \end{cases}, \tag{7}$$

A partir de (7) se deriva la función de densidad de probabilidad $f(y)$, y se obtienen las curvas de la figura 7. En la curva $f(y)$ se aprecia un máximo en ~ 0.36 , que es el nivel más probable de encontrar un fondo con colores RGB distribuidos uniformemente, por lo que es deseable que los objetos buscados tengan niveles de cromaticidad alejados de este valor.

Luego de este análisis se cuenta con las herramientas suficientes para calcular la probabilidad de que una ventana corresponda al objeto buscado según (3), en base a plantillas estadísticas de valores medios y desviaciones estándar entrenadas con anterioridad. Sin embargo estas plantillas consideran tanto píxeles del objeto como porciones del fondo presentes en la ventana de entrenamiento, por lo que se requiere descartar el área correspondiente al fondo.

Al comparar los histogramas de los píxeles correspondientes al fondo y los correspondientes al objeto, en la figura 6, se observa que el

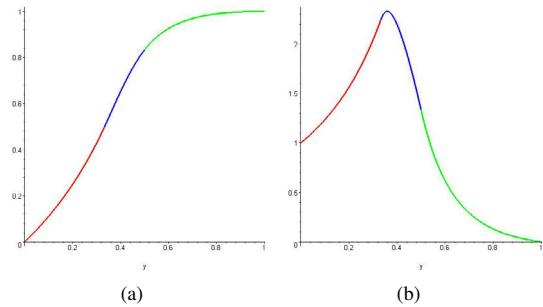


Figura 7: Función de probabilidad acumulada $F(y)$ (a), y función de densidad de probabilidad $f(y)$ (b), para $ErEgEb$, considerando que RGB sigue una distribución uniforme.

canal de luminancia Y , en el caso del fondo, abarca un rango considerablemente mayor al objeto, lo que ocupa prácticamente todo el rango de tonalidades. De aquí se deduce que la desviación estándar del canal luminancia es un buen indicador para determinar si el píxel en cuestión se trata del objeto o del fondo.



Figura 8: Imágenes, Promedio y Desviación Estándar, en la componente Y .



Figura 9: Eliminación del fondo con umbrales de 55, 60 y 65, de izquierda a derecha, respectivamente.

En la figura 9, se ve que 60 niveles de color resulta un buen valor de umbral ya que conserva la gran mayoría de los píxeles del objeto y descarta prácticamente todo el fondo. Asumiendo una distribución gaussiana, una desviación mayor que 60 implicaría que el 95% de las muestras se encontrarían más allá de ± 120 niveles, lo que abarcaría un rango de hasta 240 niveles, que serían la gran mayoría de estos.

Por lo anterior los canales a utilizar en la etapa de detección son la luminancia para eliminar el fondo de la plantilla, y las cromaticidades roja y azul para comparar sus valores con los candidatos a ser objeto, siendo robusto ante cambios de iluminación y sin considerar la información del fondo.

4. Resultados experimentales

4.1. Sistemas de percepción y procesamiento

El sistema de percepción está compuesto por una cámara y un GPS. Ambos elementos se ubican sobre un vehículo, ver figura 10(a), el mismo que se convierte en un aparato destinado al desarrollo de experimentos de sistemas de asistencia a la conducción en condiciones reales.

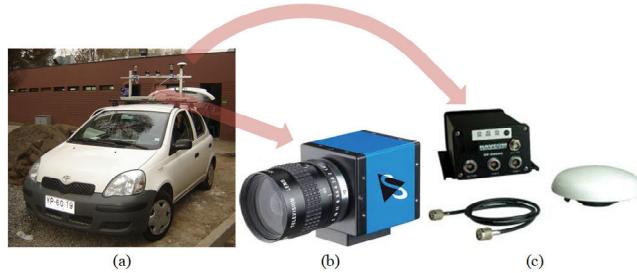


Figura 10: Vehículo de prueba (a), cámara (b) y GPS (c).

La cámara utilizada es la Imaging Source, modelo *DFK31BF03*, ver figura 10(b). El GPS es un Navcom Technology modelo *SF-2050*, ver figura 10(c).

El sistema de procesamiento está compuesto por un computador portátil con procesador Intel Core 2 Duo con una frecuencia de 2.0 GHz y 3.5 GB de memoria RAM.

4.2. Base de datos de entrenamiento y validación

Para los procedimientos de entrenamiento y validación, se tienen los siguientes ejemplos positivos: 122 discos PARE by 80 Ceda el Paso. A partir de lo cual se generaron más positivos modificando ligeramente la rotación, el tamaño y añadiendo ruido, para generar 7000 ejemplos positivos de entrenamiento, ver figura 11. En los ejemplos negativos se tienen 2567 imágenes.

La base de datos para desarrollar estos experimentos está disponible en (RAL, 2012). Esta información podrá contribuir significativamente para que la comunidad científica disponga de más herramientas para el estudio y construcción de este tipo de sistemas.

4.3. Experimentos con los detectores plantillas estadísticas y Viola-Jones

En esta etapa se pusieron a prueba dos métodos de detección de señales de tráfico, el detector de plantilla estadística y el de Viola-Jones (Viola et al., 2005).

A partir de las capturas obtenidas se obtuvieron secuencias de imágenes indexadas según su posición gracias a los datos obtenidos del GPS. Se ejecutaron los métodos sobre 6 capturas de acercamiento a la intersección, 2 para cada tipo de señal. En total se cuenta con 273 cuadros con acercamiento a un disco PARE, 447 cuadros con acercamiento a un Ceda el Paso.

1. Viola & Jones: Para la detección de discos PARE se logró una detección de 100% en proximidades inferiores a 20 metros, y se disminuyó en las capturas más lejanas hasta llegar a 0% para distancias sobre los 35 metros.

Para los Ceda el Paso no se logró detectar la señal en ninguna de las imágenes, lo que se le atribuye a que la señal de tránsito tiene un rayado en su esquina inferior, y posiblemente por

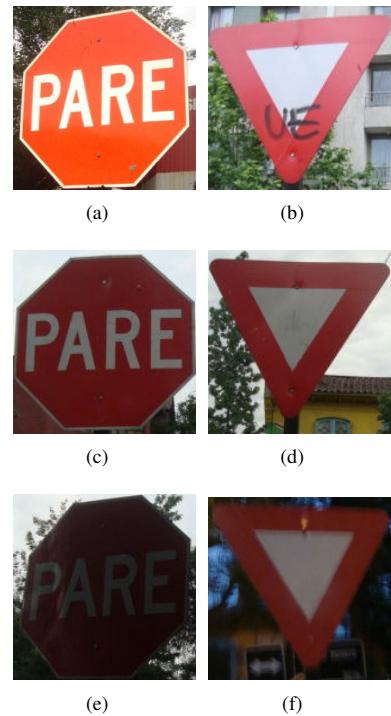


Figura 11: Señales de tráfico típicas de las intersecciones viales y rotondas en Chile: disco Pare (a,c,e) y Ceda el Paso (b,d,f); en diferentes condiciones de iluminación (soleado (a,b), normal (c,d) y oscuro (e,f)) y posición entre la cámara y la señal.

pertenecer a otra comuna de la que se obtuvieron las imágenes de entrenamiento presentaba una leve diferencia en el grosor del borde rojo. Esto muestra una alta sensibilidad a posibles modificaciones en los objetos buscados.

Las tasas de falsas alarmas se mantuvieron bajas, siendo cero en la secuencia de disco PARE y Ceda el Paso.

2. Plantillas estadísticas: Se ejecutó el algoritmo con 8 tamaños de plantillas, desde 14×14 hasta 50×50 en proyección geométrica. Esto debido a que las plantillas de menor tamaño generan una gran cantidad de falsos negativos y no aportan significativamente a la detección.

En la secuencia con disco PARE se logró un alto índice de detección hasta los 37 metros de distancia y el último reconocimiento de la señal fue a los 52 metros. A diferencia del método

Tabla 2: Tasas de detección utilizando Viola & Jones en función de la distancia.

Distancias a la Intersección [mt]	Ceda el Paso %	Disco PARE %
> 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%

anterior, se obtuvieron 10 falsos negativos durante estas capturas.

Tabla 3: Tasas de detección utilizando plantilla estadística en función de la distancia.

Distancias a la Intesección [mt]	Ceda el Paso %	Disco PARE %
> 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%

Para el cartel de Ceda el Paso se lograron detecciones del 100% hasta 35 metros de distancia, y disminuye después hasta lograr detectar por última vez a 48 metros. En esta oportunidad se presentaron 6 falsos positivos. Los desempeños de este método con Viola-Jones se puede apreciar en la figura 12.

Tabla 4: Tasa de falsas alarmas por cuadro.

Método	Ceda el Paso	Disco PARE
Viola & Jones	0.006	0
Plantilla estadística	0.036	0.069

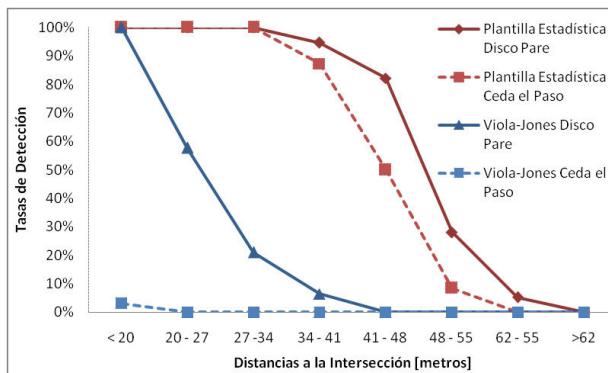


Figura 12: Comparación de desempeños de los métodos de Viola & Jones y plantilla estadística en las señales Disco Pare y Ceda el Paso según su distancia.

Los tiempos de cómputo fueron de 450 mili-segundos para Viola & Jones, y 950 mili-segundos para plantillas estadísticas.

Para culminar esta sección se presenta un resumen de las ventajas y desventajas de la propuesta.

Las principales ventajas son:

- Inmune a los cambios de iluminación. Al incorporar la media y varianza y la interacción simultánea de varios espacios de color se cubre un amplio espectro del color para la generación de las ROIs.
- Alta tasa de clasificación en función de la distancia, en el rango 30 – 50 metros.

- Sistema especializado en señales de tráfico de intersecciones viales.

Las desventajas son:

- Mediano coste computacional.
- Reducido número de casos de señales de tráfico.

5. Conclusiones

En este trabajo se ha presentado un completo procedimiento, basado en el color, para detectar las señales de tráfico próximas a las intersecciones viales, como son el Disco Pare y Ceda el Paso (Triángulo Invertido).

La fase de segmentación esta basada en el estudio del color de dichas señales, en varios espacios de color, a partir de lo cual se ha comprobado que el espacio $ErEgEb$ presenta el mejor desempeño. Luego se ha construido un método para la generación de ROIs basada en el modelos que utilizan el promedio, la varianza como umbrales para posteriormente barrer toda la imagen.

En la etapa de reconocimiento se ha construido un método basado en las plantillas estadísticas (template matching) en el espacio $ErEgEb$ y la componente Y de $YCrcb$, para lo cual, se ha modelado el fondo mediante una función de distribución de probabilidad.

En el estudio relacionado con la detección versus la distancia se han logrado resultados superiores al utilizar la técnica de plantillas estadísticas comparada con Viola & Jones. Así, para distancias menores a 48 metros se ha obteniendo una tasa de detección del 87.5% para el caso del triángulo de Ceda el Paso, y del 95.4% para el caso del disco PARE.

Para el desarrollo de los experimentos se ha construido una base de datos de señales de tráfico de intersecciones viales de Chile, esta información está disponible para el uso abierto de la comunidad científica.

Finalmente, a futuro se desarrollarán dos líneas de trabajo: la primera consiste en utilizar este método para detectar la luz roja del semáforo y el resto de las señales de reglamentación (tonalidades rojas); y la segunda concierne en extender el sistema para detectar las señales preventivas (colores amarillas) e informativas (tonalidades azules).

English Summary

Traffic sign detection system for locating road intersections and braking advance

Abstract

This paper describes a system for traffic sign detection from surrounding traffic intersections and analysis for the ability of detection depending on the distance is presented. The method is based on segmentation by color on the RGB-normalized space ($ErEgEb$) for generating regions of interest (ROIs), the classification of the type of signal using $YCrcb$ and $ErEgEb$ spaces to build a statistical template, where to remove the background has proposed a probability distribution function that models the color of the objects of interest and its background. This system is specialized in Stop and Give way signs.

After that when compared to the Viola & Jones method where the detection capability is verified at different distances is then analyzed. For distances less than 48 meters, this method reaches a detection rate of 87.5%, in the case of Give way sign and 95.4% regarding the Stop sign, while for distances less than 30 meters the detection rate is 100%. These results are higher than those in the state of the art. Experiments

have been developed over a traffic signs database, generated from images taken in several streets of the Santiago, Metropolitan Region, Chile, using an experimental vehicle designed to develop intelligent systems.

Keywords:

Distance, color, road intersection, accidents, traffic signs, statistics templates.

Agradecimientos

Este proyecto ha sido financiado por la Comisión Nacional de Ciencia y Tecnología de Chile (Conicyt) a través del proyecto Fondecyt No. 11060251 y por la Universidad de las Fuerzas Armadas-ESPE, a través de su Plan de Movilidad con Fines de Investigación 2015.

Bibliografía

- Abramowitz, M., Stegun, I. A., 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, 1st Edition. New York: Dover.
- Agencia Nacional de Tránsito del Ecuador, 2015. Siniestros octubre 2015. URL <http://www.ant.gob.ec/index.php/descargable/file/3265-siniestros-octubre-2015>
- Baya, H., Essa, A., Tuytelaars, T., Van Gool, L., June 2008. Speeded-Up Robust Features (SURF) 110 (3), 346–359.
- Carrasco, J., 2009. 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.
- CONASET, 2014. Observatorio de datos de accidentes. URL <https://estadconaset.mtt.gob.cl/>
- Dalal, N., 2006. Finding people in images and videos. Ph.D. Thesis, Institut National Polytechnique de Grenoble.
- Fleyeh, H., Biswas, R., Davami, E., July 2013. Traffic sign detection based on adaboost color segmentation and svm classification. pp. 2005–2010.
- Flores, M., 2009. 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.
- Flores, M., Armingol, M., Escalera de la, A., Oct 2007. New probability models for face detection and tracking in color images. IEEE International Symposium on Intelligent Signal Processing, WISP 2007., 1–6.
- Fraser, B., August-September 2005. Traffic accidents scar Latin America's roads 366, 703–704.
- Greenhalgh, J., Mirmehdi, M., December 2012. Real-time detection and recognition of road traffic signs. IEEE Transactions on Intelligent Transportation Systems 13 (4), 1498–1506.
- Han, Y., Virupakshappa, K., Oruklu, E., 2015. Robust traffic sign recognition with feature extraction and k-NN classification methods. In: 2015 IEEE International Conference on Electro/Information Technology (EIT). pp. 484–488.
- Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning, 2nd Edition. Springer.
- Horgan, J., Hughes, C., McDonald, J., Yogamani, S., 2015. Vision-Based Driver Assistance Systems: Survey, Taxonomy and Advances. In: IEEE 18th International Conference on Intelligent Transportation Systems (ITSC). pp. 2032–2039.
- Huang, Z., Yu, Y., Gu, J., 2014. A Novel Method for Traffic Sign Recognition based on Extreme Learning Machine, 1451–1456.
- Jain, A. K., Li, S. Z., 2005. Handbook of Face Recognition. Springer-Verlag New York, Inc., Secaucus, NJ, USA.
- Lau, M. M., Lim, K. H., Gopalai, A. A., 2015. Malaysia Traffic Sign Recognition with Convolutional Neural Network, 1006–1010.
- Li, H., Sun, F., Liu, L., Wang, L., May 2015. A novel traffic sign detection method via color segmentation and robust shape matching. Neurocomputing, 169, 77–88.
- Lillo, J., Mora, I., Figuera, C., Rojo, J. L., November 2015. Traffic sign segmentation and classification using statistical learning methods. Neurocomputing 1 (153), 286–299.
- Mesriani Law Group, 2015. Accidents caused by dangerous intersections. URL <http://www.hg.org/article.asp?id=7652>
- Mogelmose, A., Trivedi, M. M., Moeslund, T. B., December 2012. Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey. IEEE Transactions on Intelligent Transportation Systems 13 (4), 1484–1497.
- National Highway Traffic Safety Administration, 2015. URL <http://www.nhtsa.gov/>
- Nguyen, B. T., Shim, J., Kim, J. K., 1 2014. Fast traffic sign detection under challenging conditions. pp. 749–752.
- Nie, Y., Chen, Q., Chen, T., Sun, Z., Dai, B., September 2012. Camera and lidar fusion for road intersection detection. pp. 273–276.
- Perez-Perez, S. E., Gonzalez-Reyna, S. E., Ledesma-Orozco, S. E., Avina-Cervantes, J. G., 2013. Principal component analysis for speed limit Traffic Sign Recognition. In: 2013 IEEE International Autumn Meeting on Power Electronics and Computing (ROPEC). pp. 1–5.
- RAL, 2012. Driver assistance system. Robotics and Automation Laboratory, School of Engineering, Pontificia Universidad Católica de Chile. URL <http://ral.ing.puc.cl/datasets/intersection/index.htm>
- Salvi, S., Petrelli, A., Tombari, F., Fioraio, N., DiStefano, L., June 2015. Traffic sign detection via interest region extraction. Pattern Recognition 48, 1039–1049.
- Viola, P., Jones, M., 2001a. Rapid object detection using a boosted cascade of simple features. In: Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. pp. I-511–I-518 vol.1.
- Viola, P., Jones, M., 2001b. Robust real-time face detection. In: Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on. Vol. 2. pp. 747–747.
- Viola, P. A., Jones, M. J. D., S., July 2005. Detecting pedestrians using patterns of motion and appearance. International Journal of Computer Vision 63 (2), 153–161.
- World Health Organization WHO, 2015a. La oms y la fia aúnán esfuerzos para mejorar la seguridad vial. URL <http://www.who.int/mediacentre/news/releases/2003/pr11/es/>
- World Health Organization WHO, 2015b. Lesiones causadas por el tránsito. URL <http://www.who.int/mediacentre/factsheets/fs358/es/>
- World Health Organization WHO, 2015c. Road traffic injuries.
- Zaklouta, F., Stanciuescu, B., December 2012. Real-time traffic-sign recognition using tree classifiers. IEEE Transactions on Intelligent Transportation Systems 13 (4), 1507–1514.
- Zaklouta, F., Stanciuescu, B., 2014. Real-time traffic sign recognition in three stages. Robotics and Autonomous Systems 62, 16–24.

Apéndice A.

Deducción de la función de distribución para el color

Para derivar $F(y)$, la CDF del modelo analítico se tiene que considerar que cada canal del espacio RGB se distribuye uniformemente en su dominio. Para esto se definen 3 variables aleatorias, X_1, X_2, X_3 según la siguiente distribución:

$$X_1, X_2, X_3 \sim U(0, S) \Rightarrow f(x) = \begin{cases} 0 & x < 0 \vee x \geq S \\ \frac{1}{S} & 0 \leq x < S \end{cases}, \quad (A.1)$$

con lo que la función de probabilidad acumulada de un canal de cromaticidad se calcula mediante la probabilidad

$$F(y) = P(Y \leq y) = P\left(\frac{X_1}{X_1 + X_2 + X_3} \leq y\right), \quad (A.2)$$

Para resolver (A.2) se presentan los siguientes casos:
para $y = 0$,

$$\begin{aligned} F(0) &= P\left(\frac{X_1}{X_1 + X_2 + X_3} \leq 0\right) \\ &= P(X_1 \leq 0) = 0, \end{aligned} \quad (A.3)$$

para $y = 1$,

$$\begin{aligned} F(1) &= P\left(\frac{X_1}{X_1 + X_2 + X_3} \leq 1\right) \\ &= P(X_2 + X_3 \geq 0) = 1, \end{aligned} \quad (\text{A.4})$$

y para $0 < y < 1$,

$$\begin{aligned} F(y) &= P\left(\frac{X_1}{X_1 + X_2 + X_3} \leq y\right) \\ &= P\left(X_1 \leq \frac{y}{1-y}(X_2 + X_3)\right), \end{aligned} \quad (\text{A.5})$$

con lo que se obtiene la siguiente integral

$$F(y) = \int_0^S f(x_3) \int_0^S f(x_2) \int_0^{\frac{y}{1-y}(x_2+x_3)} f(x_1) dx_1 dx_2 dx_3. \quad (\text{A.6})$$

La integral en $f(x_1)$ tiene 2 comportamientos en función del valor de $\frac{y}{1-y}(x_2 + x_3)$ ya que si es mayor que S , se integra la función de densidad $\frac{1}{s}$ hasta S , y luego una función nula entre S y $\frac{y}{1-y}(x_2 + x_3)$, mientras que si es menor que S , se integra entre 0 y dicho valor. Para que este último se cumpla en todo el dominio no nulo de $f(x_2)$ y $f(x_3)$, se debe tener

$$\frac{y}{1-y}(x_2 + x_3) \leq S \quad \forall x_2, x_3 \in [0, S] \Rightarrow y \leq \frac{1}{3} \quad (\text{A.7})$$

Del mismo modo, al analizar los otros límites de integración, se separa en tres casos: $(0, \frac{1}{3}]$, $(\frac{1}{3}, \frac{1}{2}]$ y $(\frac{1}{2}, 1)$, de modo que las funciones tomen valores distintos de cero y sean continuas en el rango de integración.

Para $0 < y \leq \frac{1}{3}$

$$F(y) = \int_0^S \frac{1}{s} \int_0^S \frac{1}{s} \int_0^{\frac{y}{1-y}(x_2+x_3)} \frac{1}{s} dx_1 dx_2 dx_3 = \frac{y}{1-y}, \quad (\text{A.8})$$

para $\frac{1}{3} < y \leq \frac{1}{2}$ se separa en 3 integrales

$$\begin{aligned} F(y) &= \int_0^{\frac{1-y}{y}S-S} \frac{1}{s} \int_0^S \frac{1}{s} \int_0^{\frac{y}{1-y}(x_2+x_3)} \frac{1}{s} dx_1 dx_2 dx_3 \\ &+ \int_{\frac{1-y}{y}S-S}^S \frac{1}{s} \int_0^{\frac{1-y}{y}S-x_3} \frac{1}{s} \int_0^{\frac{y}{1-y}(x_2+x_3)} \frac{1}{s} dx_1 dx_2 dx_3 \\ &+ \int_{\frac{1-y}{y}S-S}^S \frac{1}{s} \int_{\frac{1-y}{y}S-x_3}^S \frac{1}{s} \int_0^S \frac{1}{s} dx_1 dx_2 dx_3 \\ &= \frac{-21y^3 + 27y^2 - 9y + 1}{6y^2(1-y)}, \end{aligned} \quad (\text{A.9})$$

$$\begin{aligned} F(y) &= \int_0^{\frac{1-y}{y}S} \frac{1}{s} \int_0^{\frac{1-y}{y}S-x_3} \frac{1}{s} \int_0^{\frac{y}{1-y}(x_2+x_3)} \frac{1}{s} dx_1 dx_2 dx_3 \\ &+ \int_0^{\frac{1-y}{y}S} \frac{1}{s} \int_{\frac{1-y}{y}S-x_3}^S \frac{1}{s} \int_0^S \frac{1}{s} dx_1 dx_2 dx_3 \\ &+ \int_{\frac{1-y}{y}S}^S \frac{1}{s} \int_0^S \frac{1}{s} \int_0^S \frac{1}{s} dx_1 dx_2 dx_3 \\ &= \frac{5y^2 + 2y - 1}{6y^2}. \end{aligned} \quad (\text{A.10})$$

y por último para $\frac{1}{2} < y < 1$ también se separa en 3 casos

Con lo que la función final es la dada por (7).

Apéndice B.

Detección de señales de tráfico usando el método de Viola & Jones

El detector de objetos de Viola & Jones (Viola and Jones, 2001a,b) ha sido utilizado en múltiples aplicaciones (Flores, 2009), (Viola et al., 2005). En este trabajo se lo ha utilizado para construir un detector de señales de tráfico, en los casos de Disco Pare y Ceda el Paso. En la figura B.13 se observan las primeras características de tipo Haar para cada uno de los casos en cuestión.

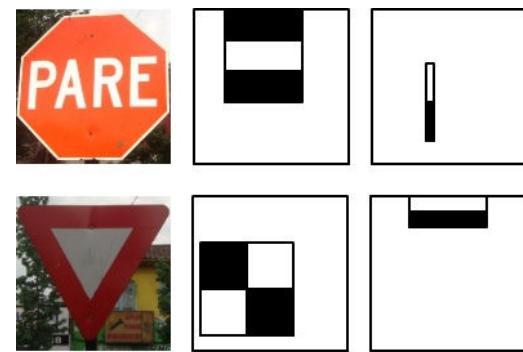


Figura B.13: Primera y segunda característica tipo Haar del clasificador de Viola & Jones para el Disco Pare y Ceda el Paso.