

---

# Detección, Categorización y Reconocimiento de Señales de Tránsito

Informe de Practica Profesional Supervisada

por

Sebastián Prenna

Supervisor: Professor Reinhard Klette

Tāmaki Innovation Campus  
Department of Computer Science  
The University of Auckland  
New Zealand  
Febrero, 2012



---

## Resumen

**Palabras clave:** Sistemas de ayuda al conductor, reconocimiento de señales de tránsito, scale-invariant feature transform, SIFT, template matching.

Los sistemas de reconocimiento de señales de tránsito son una parte importante en los sistemas de ayuda al conductor basados en visión. Estos mantienen al conductor informado acerca de diferentes condiciones del camino, y proveen información adicional útil para mejorar la seguridad vial.

El sistema de reconocimiento de señales de tránsito propuesto aquí, fue diseñado con el fin de que sea capaz de reconocer una amplia variedad de señales de tránsito, y para operar tanto con imágenes a color como en escala de grises.

El sistema presentado se basa en la extracción de regiones de interés, las cuales son áreas de la imagen que pueden llegar a contener señales de tránsito. Estas regiones son seleccionadas buscando determinadas formas en la imagen de entrada, clasificándolas posteriormente mediante un método basado en la comparación de características descriptivas, contra una base de datos de señales de tránsito.

Los experimentos realizados hasta el momento muestran que el método propuesto puede ser plenamente implementado y utilizado en situaciones reales, luego de algunas modificaciones que consideran situaciones particulares en las que el rendimiento se ve afectado.



## **Agradecimientos**

Me gustaría expresar mi especial agradecimiento a mi supervisor Prof. Reinhard Klette, quien me dio la oportunidad de hacer este proyecto, y por su constante apoyo y motivación, que me llevaron a la realización del presente proyecto. También estoy agradecido a mi familia y amigos, quienes me dieron apoyo emocional durante este tiempo. Por último, me gustaría dar las gracias a las personas que trabajan en el Departamento de Ciencias de la Computación del Tāmaki Innovation Campus, por ofrecerme su amistad y por hacerme sentir tan cómodo aquí.

Sebastián Prenna  
Auckland  
17 de Febrero, 2012



---

# Contenido

<b>Resumen</b>	<b>iii</b>
<b>Agradecimientos</b>	<b>v</b>
<b>1 Introducción</b>	<b>3</b>
1.1 Esquema del proyecto . . . . .	3
1.2 Señales de tránsito existentes . . . . .	4
1.3 Sistemas de ayuda al conductor . . . . .	4
<b>2 Conceptos fundamentales</b>	<b>7</b>
2.1 Espacios de Color . . . . .	7
2.2 Morfología matemática . . . . .	8
2.3 Scale-invariant feature transform . . . . .	11
2.4 Métodos de reconocimiento de señales de tránsito . . . . .	14
<b>3 Sistema propuesto para el reconocimiento de señales de tránsito</b>	<b>15</b>
3.1 Perspectiva general del método . . . . .	15
3.2 Conversión del espacio de colores . . . . .	18
3.3 Detección de bordes y cálculo de gradiente . . . . .	19
3.4 Detección de formas . . . . .	20
3.5 Extracción de características simples - Ordenamiento jerárquico . . . . .	21
3.6 Extracción y comparación de descriptores . . . . .	22
<b>4 Experimentos y resultados</b>	<b>25</b>
4.1 Extracción de regiones de interés . . . . .	25
4.2 Extracción de características y clasificación . . . . .	28

---

	1
<b>5 Conclusiones</b>	<b>37</b>
<b>Bibliografía</b>	<b>39</b>
<b>Indice</b>	<b>41</b>



## Capítulo 1

---

# Introducción

*Este capítulo introductorio detalla la estructura del informe. El proyecto es acerca del reconocimiento de señales de tránsito, y el capítulo esboza la diversidad de señales existentes. También se explica brevemente el contexto de los sistemas de ayuda al conductor basados en visión.*

### 1.1 Esquema del proyecto

El objetivo de este proyecto es diseñar e implementar un sistema de reconocimiento de señales de tránsito, para ser utilizado en sistemas de ayuda al conductor basados en visión. El método propuesto fue diseñado de manera de que fuese capaz de reconocer una amplia variedad de señales de tránsito.

A manera de introducción al reconocimiento de señales de tránsito, se expone en ese capítulo una idea general acerca de los diferentes tipos de señales de tránsito, y se abordan también brevemente los sistemas de ayuda al conductor.

En el Capítulo 2 se presentan los fundamentos de diferentes temas que son utilizados luego en el método propuesto, y se analizan también distintos sistemas de reconocimiento de señales de tránsito existentes, dando una idea acerca de diferentes soluciones que fueron propuestas en este campo.

Al comienzo del Capítulo 3, se da una descripción general de las características requeridas para el sistema, junto con una explicación del mismo desde un punto de vista funcional, dando así una idea general de su funcionamiento. Luego se divide el proceso de reconocimiento en bloques funcionales, para ser analizados cada uno por separado.

El Capítulo 4 muestra los resultados de la implementación del método propuesto, los cuales son también analizados, arribando de esta manera al Capítulo 5, el cual contiene las conclusiones finales acerca del proyecto y donde se exponen también algunas ideas para trabajos futuros.

## 1.2 Señales de tránsito existentes

El sistema propuesto para el reconocimiento de señales de tránsito está diseñado para funcionar en las calles y carreteras de Nueva Zelanda. El sistema vial de Nueva Zelanda es construido, regulado y mantenido por la Agencia de Transporte de Nueva Zelanda (*New Zealand Transport Agency* - NZTA) [14]. Esta agencia es la encargada de regular también el sistema de señalización.

La mayoría de las señales de tránsito que pueden ser vistas en las carreteras de Nueva Zelanda son señales internacionales. Esto significa que ellos utilizan como señales de tránsito las mismas formas y símbolos que son utilizadas internacionalmente. Esto hace posible la implementación del sistema propuesto en el presente trabajo, en aquellos países que hagan uso de estos estándares internacionales.

La *New Zealand Transport Agency* clasifica las señales de tránsito en 3 grupos principales:

Un grupo de señales de tránsito es el que conforman las *señales obligatorias*, las cuales indican al conductor que debe o no debe hacer. Usualmente son rojas o azules.

Otro grupo de señales está compuesto por las *señales de advertencia*, que alertan acerca de algún peligro particular que se encuentra más adelante en el camino. Estas advierten al conductor para que sea cuidadoso por su propia seguridad, la de otros conductores, o la seguridad de trabajadores que estén llevando a cabo tareas de mantenimiento. Hay dos tipos de señales de advertencia: aquellas que advierten acerca de algún peligro permanente, y aquellas que advierten acerca de un peligro temporal. Ambos tipos de señales son generalmente de forma romboidal, pero las señales permanentes son de color amarillo y negro, mientras que las temporales son de color naranja y negro.

El tercer grupo es el de las *señales de información*. Estas proveen información útil, como por ejemplo la distancia hasta la próxima ciudad. Todas poseen forma rectangular, pero pueden ser de varios colores y tamaños.

En la figura 1.1 se pueden ver ejemplos de cada una de las diferentes clases de señales de tránsito.

## 1.3 Sistemas de ayuda al conductor

Los sistemas de ayuda al conductor (*Driver Assistance Systems* - DAS) son sistemas que hacen uso de diversos sensores instalados en el auto, como pueden ser radares, sensores laser, cámaras de video y otros, para comprender el entorno que los rodea, como así también el comportamiento del conductor. Con toda esta información

es posible proveer al conductor de información adicional, advertir acerca de situaciones peligrosas, o incluso asistir al conductor o tomar el control del vehículo con el fin de evitar o minimizar las consecuencias de un accidente, prediciendo situaciones de peligro.

Dado que las marcas de la carretera, señales de tránsito y otros tipos de información relevante son posibles de detectar solamente mediante la utilización de información visual, muchos vehículos existentes en el mercado poseen cámaras instaladas.

Se pueden encontrar diversos módulos que componen sistemas de ayuda al conductor en modelos de vehículos en producción comercial. Ejemplos de estos módulos son: Sistema de información de puntos ciegos (*Blind Spot Information System* - BLIS), Advertencia de colisión frontal (*Forward Collision Warning* - FCW), Sistema de visión nocturna (*Night Vision System* - NVS), Sistema de alerta al conductor (*Driver Alert System* - DAS), Control crucero adaptativo (*Adaptive Cruise Control* - ACC), Advertencia de desvío de carril (*Lane Departure Warning* - LDW)

La figura 1.2 muestra las áreas frecuentemente cubiertas por los diferentes tipos de sensores utilizados comúnmente en sistemas de ayuda al conductor, y dos ejemplos de sistemas de ayuda al conductor basados en visión implementados en vehículos en producción, como son el sistema de visión nocturna (*Night Vision System*) de BMW, y la asistencia de límite de velocidad (*Speed Limit Assist*) de Mercedes-Benz.



Figura 1.1: Ejemplos de señales de tránsito. De izquierda a derecha: señales obligatorias (primeras dos), señal de advertencia (permanente), señal de advertencia (temporal) y señal de información.

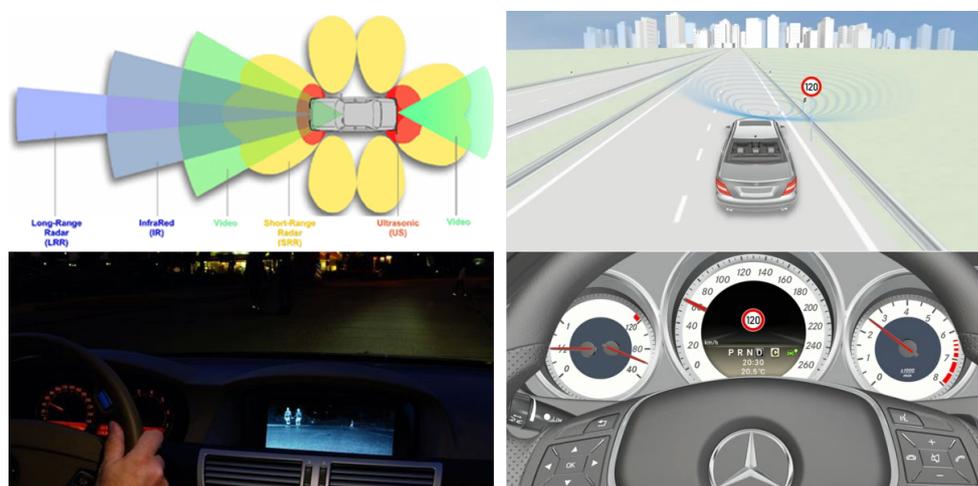


Figura 1.2: Imágenes de sistemas de ayuda al conductor. Izquierda arriba: áreas comúnmente cubiertas por diferentes tipos de sensores y sus alcances. Izquierda abajo: imagen del Night Vision System de BMW. Derecha: imágenes del sistema Speed Limit Assist de Mercedes-Benz

## Capítulo 2

---

# Conceptos fundamentales

*Este capítulo se desarrollan los fundamentos de conceptos utilizados en el resto del informe, y se analizan algunos sistemas existentes de reconocimiento de señales de tránsito, en el contexto particular de sistemas de ayuda al conductor basados en visión.*

## 2.1 Espacios de Color

Un modelo de color es el modelo matemático que describe cómo los colores son representados. Generalmente esta representación es en tuplas de tres o cuatro elementos. Por si mismos, estos números no tienen ningún significado. La función que mapea este modelo con un espacio de color absoluto, da sentido a cada una de estas tuplas numéricas, y un determinado modelo de color con su función de mapeo, define el espacio de color.

Ejemplos de modelos de color son RGB (Red, Green, Blue / Rojo, Verde, Azul), CMYK (Cyan, Magenta, Yellow, Black / Cian, Magenta, Amarillo, Negro) o HSV (Hue, Saturation, Value / Matiz, Saturación, Valor). En esta sección se explicara brevemente el modelo de color HSV y sus diferencias con el modelo RGB en aplicaciones de visión computarizada.

Los colores representados utilizando el modelo de color RGB se especifican mediante tuplas de tres elementos, donde cada uno de ellos representa la cantidad de luz de cada color (Rojo, Verde, Azul) que debe ser añadida para reproducir un color determinado. Esta es una manera sencilla de representar un color, pero no es como el ojo humano percibe el concepto de color. Cuando los humanos perciben un color, no lo describen en términos de sus componentes RGB sino que lo describen mediante un modelo que está basado en las componentes de Matiz, Saturación y Valor.

Definiciones (extraídas de [11]):

- Matiz (en inglés, Hue) es el atributo de la sensación visual en que una superficie parece ser similar a uno de los colores percibidos: rojo, amarillo, verde, azul, o a una combinación de dos de ellos.

- El brillo es el atributo de la sensación visual en que una superficie parece emitir más o menos luz.
- La luminosidad (en inglés, Lightness o Value) es el brillo relativo al brillo del color blanco, tomado como referencia.
- La cromaticidad es el atributo de una sensación visual según el cual una superficie parece presentar un color percibido más o menos cromático.
- La saturación es la cromaticidad de un estímulo relativo a su propio brillo.

Como se puede interpretar de las definiciones dadas, en el modelo de color HSV el color percibido es independiente de la intensidad, mientras que en el modelo RGB las componentes de rojo verde y azul presentan cambios frente a cambios en la iluminación, y esta es la principal ventaja por la que el modelo HSV es ampliamente utilizado en visión computarizada.

En la figura 2.1 se pueden ver las diferencias entre las representaciones del color a través de sus componentes. Es posible apreciar aquí que en el modelo de color HSV, la componente de Matiz (color percibido) y la Saturación (que representa la pureza del color), son ambas independientes de la componente de Valor (cantidad de luz).

## 2.2 Morfología matemática

De acuerdo con [12], la morfología matemática es una herramienta para la extracción de componentes de la imagen, que son útiles en la representación y descripción de una forma, como son sus bordes, estructura, cubierta convexa, etc.

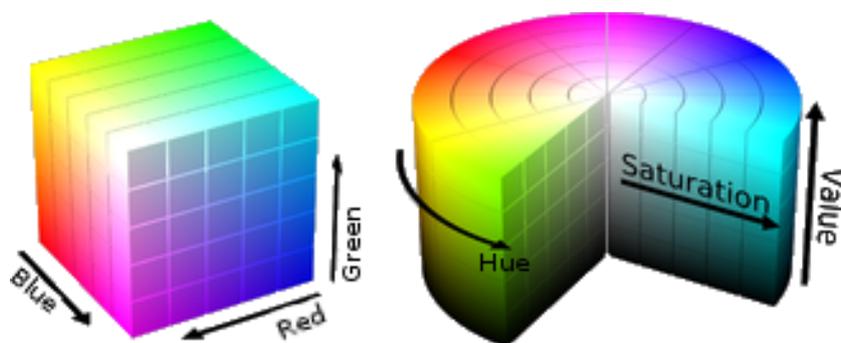


Figura 2.1: Representaciones de los modelos de color RGB y HSV.

Muchas operaciones morfológicas están definidas en términos de espacios Euclidianos  $n$ -dimensionales  $\mathbb{E}^n$ . En este trabajo serán explicadas solo las operaciones de erosión y dilatación, e interesan solo sus aplicaciones sobre imágenes binarias, en las cuales sus elementos están en una cuadrícula entera  $\mathbb{Z}^2$ , y donde cada elemento es un par ordenado cuyas coordenadas  $(x, y)$  son las coordenadas correspondientes a un pixel blanco (o negro, dependiendo de la convención utilizada) de la imagen.

La erosión y la dilatación son las operaciones primitivas de muchos algoritmos morfológicos. En este trabajo estas dos operaciones son utilizadas básicamente para remover ruidos en la imagen (como ruido pimienta) y rellenar huecos en imágenes binarias.

Para entender las operaciones morfológicas, debemos ver algunos conceptos de teoría de conjuntos, donde los conjuntos representan objetos en una imagen.

Sea  $A$  un conjunto en  $\mathbb{Z}^2$ . Si  $a = (a_1, a_2)$  es un elemento de  $A$ , entonces escribimos esto como (2.1).

$$a \in A \quad (2.1)$$

La *unión* de dos conjuntos  $A$  y  $B$  es el conjunto de todos los elementos que pertenecen al conjunto  $A$ , al  $B$ , o ambos, y se denota como (2.2).

$$C = A \cup B \quad (2.2)$$

La *intersección* de dos conjuntos  $A$  y  $B$ , denotada por (2.3), es el conjunto de todos los elementos pertenecientes a ambos conjuntos,  $A$  y  $B$ .

$$C = A \cap B \quad (2.3)$$

La *traslación* del conjunto  $A$  por el punto  $z = (z_1, z_2)$ , denotada  $(A)_z$ , es definida como (2.4).

$$(A)_z = \{c | c = a + z, \text{ para } a \in A\} \quad (2.4)$$

Las operaciones básicas en morfología matemática operan en dos conjuntos: el primero ( $A$ ) es la imagen, y el segundo ( $B$ ) es comúnmente llamado *elemento estructurante*. El elemento estructurante utilizado en la práctica es generalmente de un tamaño mucho menor que la imagen.

Erosión:

Con  $A$  y  $B$  como conjuntos en  $\mathbb{Z}^2$ , la *erosión* de  $A$  por  $B$  se denota  $A \ominus B$ , y está definida como (2.5).

$$A \ominus B = \{c | c = a + b \in A \text{ para todo } b \in B\} \quad (2.5)$$

Que puede ser expresado también como en (2.6).

$$A \ominus B = \bigcap_{b \in B} A_{-b} \quad (2.6)$$

Esto puede ser interpretado como el conjunto de todos los puntos  $c$ , tal que  $B$  trasladado por  $c$  esté contenido en  $A$ .

El efecto de esta operación es remover cualquier pixel del primer plano que no esté totalmente rodeado por otros pixeles del primer plano.

En la figura 2.2 se puede ver un ejemplo de erosión sobre una imagen binaria.

Dilatación:

Con  $A$  y  $B$  como conjuntos en  $\mathbb{Z}^2$ , la *dilatación* de  $A$  por  $B$ , se denota  $A \oplus B$ , y está definida por (2.7).

$$A \oplus B = \{c | c = a + b, a \in A \text{ y } b \in B\} \quad (2.7)$$

Que puede ser dado también por (2.8), y es una operación conmutativa, como queda expresado en (2.9).

$$A \oplus B = \bigcup_{b \in B} A_b \quad (2.8)$$

$$A \oplus B = B \oplus A = \bigcup_{a \in A} B_a \quad (2.9)$$

Si  $B$  tiene centro en el origen, como antes, entonces la dilatación de  $A$  por  $B$  se puede entender como el lugar de todos los puntos cubiertos por  $B$  cuando el centro

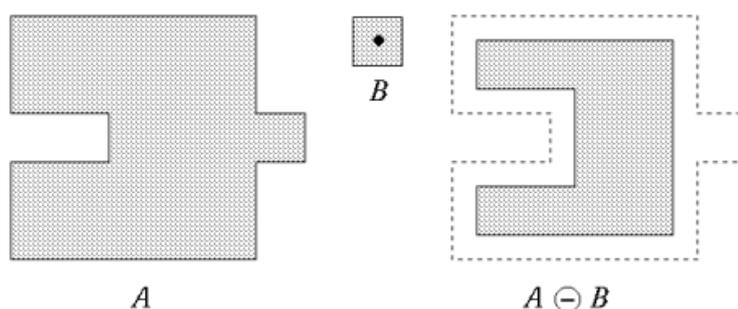


Figura 2.2: Ejemplo de erosión. De izquierda a derecha: imagen original  $A$ , elemento estructurante  $B$ , y el resultado de la erosión de  $A$  por  $B$ . En la imagen resultante se puede ver en líneas punteadas el contorno original de  $A$ .

de  $B$  se mueve dentro de  $A$ , o puede ser visto también como el conjunto de todos los desplazamientos  $b$  tal que  $B$  y  $A$  se superponen por al menos un elemento distinto de cero.

En la figura 2.3 se puede ver un ejemplo de dilatación sobre una imagen binaria.

## 2.3 Scale-invariant feature transform

El método de "Scale-invariant feature transform" (SIFT) fue desarrollado por Lowe en [15]. Este es un método ampliamente utilizado para la extracción de características relevantes (descriptores) en imágenes, las cuales son necesarias en aplicaciones como en reconocimiento de patrones o correspondencia estéreo.

Los descriptores encontrados a través de este método tienen la particularidad de ser invariantes a escalado y rotaciones, robustos frente a cambios en el punto de vista, iluminación, y con distorsión afín. Otra particularidad de estos descriptores es que son muy distintivos, lo que significa que un determinado descriptor puede ser comparado satisfactoriamente en una gran base de datos de otros descriptores, haciéndolos adecuados para aplicaciones como en reconocimiento de patrones (por citar un solo ejemplo).

El proceso para extraer estas características se puede dividir en cuatro etapas principales:

- Detección de extremos en el espacio escala
- Localización de puntos clave

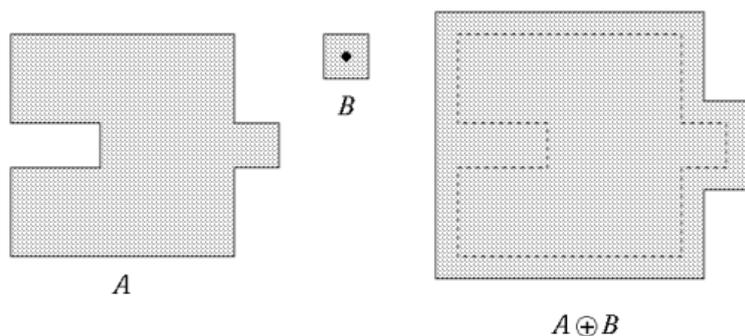


Figura 2.3: Ejemplo de dilatación. De izquierda a derecha: imagen original  $A$ , elemento estructurante  $B$ , y el resultado de la dilatación de  $A$  por  $B$ . En la imagen resultante se puede ver en líneas punteadas el contorno original de  $A$ .

- Asignación de orientación
- Cálculo del descriptor

La detección de extremos en el espacio escala, en la función de diferencia de Gaussianas  $D(x, y, \sigma)$ , definida en la ecuación (2.10), arroja puntos que son invariantes a escala. La función de diferencia de Gaussianas se calcula como la diferencia de dos espacios de escala  $L(x, y, \sigma)$  en escalas cercanas, separadas por una constante  $k$ .  $L(x, y, \sigma)$  es definida como (2.11), donde  $I(x, y)$  es la imagen, y  $G(x, y, \sigma)$  es la variable de escala Gaussiana (2.12).

$$D(x, y, \sigma) = (G(x, y, \sigma) - G(x, y, k, \sigma)) * I(x, y) \quad (2.10)$$

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2.11)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (2.12)$$

Luego, para que sea más rápido computacionalmente, la diferencia de Gaussianas (2.10) se calcula utilizando la ecuación (2.13).

$$D(x, y, \sigma) = L(x, y, \sigma) - L(x, y, k, \sigma) \quad (2.13)$$

Los extremos locales en  $D(x, y, \sigma)$  son detectados por comparación contra sus ocho vecinos en la misma escala, y sus nueve vecinos de las escalas superior e inferior. Si el punto analizado es mayor o menor que todos sus vecinos, es seleccionado como un extremo local.

Habiendo encontrado estos extremos locales, tenemos una idea acerca del radio de influencia de ese punto clave, dado que se conoce la escala  $k$ .

Examinando ahora las vecindades de cada punto clave, son descartados aquellos que están pobremente localizados a lo largo de un borde, o los que poseen un bajo contraste.

Luego, para cada muestra de la imagen en su escala, se calcula la magnitud y orientación del gradiente, mediante las ecuaciones (2.14) y (2.15) respectivamente.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (2.14)$$

$$\theta(x, y) = \tan^{-1} (L(x+1, y) - L(x-1, y)) + (L(x, y+1) - L(x, y-1)) \quad (2.15)$$

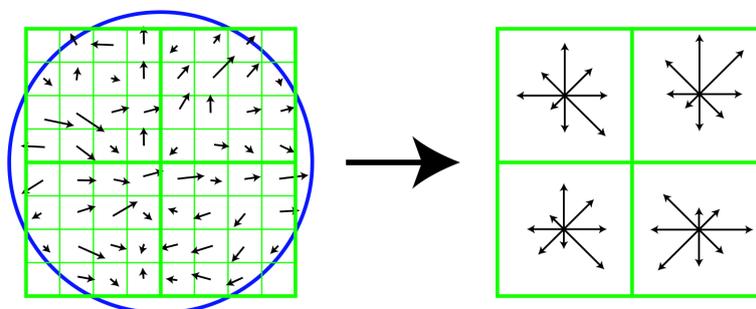


Figura 2.4: Cálculo del descriptor del punto clave. Izquierda: gradientes de la imagen en una ventana de 8x8 centrada en el punto clave. En azul, la ventana Gaussiana utilizada para ponderar la magnitud del gradiente. Derecha: descriptor del punto clave, de 2x2, compuesto por los histogramas de orientación de los gradientes en cada sub ventana de la imagen de gradientes original.

Con el objetivo de tener descriptores invariantes a rotaciones, se asigna una orientación principal a cada punto clave. En [15], se calcula la orientación principal del punto clave tomando el pico más alto del histograma calculado con las orientaciones de los gradientes. Los pixeles que son utilizados para calcular el histograma se encuentran en una ventana centrada en el punto clave, y cada muestra es ponderada respecto a la magnitud del gradiente en ese punto, y también con una ventana Gaussiana centrada en el punto clave. La orientación principal puede ser también asignada utilizando criterios diferentes, como por ejemplo utilizando directamente la orientación del gradiente en el punto clave.

Con las etapas previas se localizaron puntos clave en la imagen, y se extrajo información acerca de la escala y orientación. Ahora, para hacer distintivo a cada punto clave de otros, se debe calcular su descriptor. Este descriptor se calcula haciendo uso de la magnitud y orientación del gradiente en una ventana centrada en el punto clave. Primero, se rotan todos los gradientes empleando la orientación principal asignada al punto clave, con el fin de que el descriptor resultante sea invariante a rotaciones, y se ponderan los valores de magnitud mediante una ventana Gaussiana. Luego, la ventana es subdividida en regiones, y por cada región se calcula un histograma de orientación del gradiente, utilizando como peso para cada muestra la correspondiente magnitud del gradiente. El vector de características se forma entonces con los valores de los histogramas de orientación de cada región, el cual es normalizado a la unidad, se limitan sus valores a un valor máximo determinado y se vuelve a normalizar a la unidad.

La Figura 2.4 muestra la región utilizada para calcular el descriptor, con la mag-

nitud y orientación del gradiente en cada pixel, y el histograma resultante en cada sub-región.

La normalización del vector de características provee de robustez frente a variaciones en la iluminación promedio, dado que la magnitud relativa de sus componentes permanece constante. El umbral aplicado a los valores del vector previo a normalizar por segunda vez, previene que una determinada dirección del vector de características llegue a ser altamente dominante, y provee robustez por ejemplo ante cambios en la iluminación de objetos 3D, los cuales pueden ser no lineales.

## 2.4 Métodos de reconocimiento de señales de tránsito

Existen muchos sistemas de reconocimiento de señales de tránsito, los cuales utilizan distintos métodos para alcanzar su objetivo. La mayoría de ellos hacen uso de dos etapas bien diferenciadas, una dedicada a la extracción de áreas de interés, y la otra para clasificarlas. Algunos sistemas utilizan imágenes a color, y realizan una segmentación de la imagen realizando un filtrado por color [1, 2], o agrupamiento de colores [3], y solo algunos pocos hacen uso de imágenes en escala de grises [4]. Muchos sistemas son diseñados para ser capaces de detectar solo algunas señales específicas, como por ejemplo solo señales circulares [5], y algunos tienen el objetivo de reconocer una amplia variedad de señales diferentes [10]. Para clasificar las áreas de interés resultantes de la etapa previamente mencionada, se utilizan diversos algoritmos de clasificación, ya sea mediante redes neuronales [6], support vector machines (SVM) [7], AdaBoost o alguna variante del mismo [8, 9], o comparación de vectores de características utilizando descriptores SIFT [10], SURF o técnicas similares.

Este trabajo tiene como objetivo reconocer una amplia variedad de señales de tránsito, utilizando imágenes en escala de grises. En la etapa de clasificación se utilizan descriptores SIFT, con el fin de aprovechar las ventajas de ser invariantes a escala, rotación, transformación afín, y robustos frente a cambios en la iluminación, etc. Otra ventaja en la utilización de descriptores SIFT es que se pueden añadir nuevas señales de tránsito a la base de datos sin que sea necesario realizar un proceso de entrenamiento del sistema luego de la modificación de la base de datos. El sistema desarrollado puede también trabajar con imágenes a color, situación en la cual la performance del sistema resulta incrementada.

## Capítulo 3

---

# Sistema propuesto para el reconocimiento de señales de tránsito

*Este capítulo detalla mi método propuesto para el reconocimiento de señales de tránsito en imágenes de video.*

### 3.1 Perspectiva general del método

El sistema presentado está diseñado para ser utilizado en sistemas de ayuda al conductor, los cuales deben operar en tiempo real, por lo que es importante cualquier mejora que pueda ser realizada en virtud de aumentar la velocidad del sistema.

Como para el reconocimiento y clasificación de las señales se utiliza la extracción y comparación de descriptores utilizando SIFT, el sistema podría constar de simplemente la extracción de descriptores sobre toda la imagen, y una posterior comparación con cada una de las señales almacenadas en la base de datos.

Esta solución funciona, pero es demasiado lenta como para ser aplicable en sistemas de tiempo real. Primero que nada, buscar descriptores en toda la imagen no es una buena aproximación en términos de velocidad. El segundo problema acerca de este enfoque es que la etapa de clasificación se vuelve demasiado lenta cuando la base de datos contiene un gran número de señales de tránsito.

Para acelerar los procesos de extracción y comparación de los vectores de características, el método propuesto hace uso de dos técnicas: Una es la selección de regiones de interés (ROI) en la imagen, las cuales son zonas que pueden llegar a contener señales de tránsito, y realizar la extracción de características solo en estas regiones. La otra técnica es utilizada en la etapa de clasificación, donde en vez de simplemente comparar los vectores de características extraídos de cada región de interés con todas las señales almacenadas en la base de datos, se realiza un ordenamiento previo de esta última, utilizando descriptores más simples.

Este proceso de ordenamiento se realiza en virtud de prevenir comparaciones sin propósito, aumentando la probabilidad de obtener una coincidencia satisfactoria en

las primeras comparaciones, ahorrando tiempo de procesamiento.

Una de las ventajas añadidas mediante la utilización de descriptores SIFT para la clasificación de las señales de tránsito, es que no es necesario realizar un entrenamiento del sistema, posibilitando la adición on-the-fly de nuevas señales para ser reconocidas, en contraste con otros sistemas de clasificación como AdaBoost y redes neuronales.

Existen muchos sistemas de reconocimiento de señales de tránsito que utilizan información de color para realizar la selección de regiones de interés, filtrando colores que comúnmente contienen las señales de tránsito (por ejemplo, las señales con indicación de velocidad máxima contienen un círculo rojo en el borde) u otros colores distintivos. Esta propuesta restringe, a un grupo determinado, las señales que el sistema es capaz de detectar.

En el método propuesto se utilizan imágenes en escala de grises, pero se contempla también el caso de imágenes a color. En el caso de que se disponga de imágenes a color, se utiliza esta información adicional para mejorar el desempeño del sistema.

Muchas veces no es posible distinguir las señales de tránsito entre el fondo de la imagen si se utilizan imágenes en escala de grises, pero estas mismas resaltan cuando se analizan algunas características de la imagen a color. Utilizando esto podemos encontrar algunas señales en imágenes en que, si se utilizasen imágenes en escala de grises, pasarían desapercibidas. Otro uso de la información de color, es que se puede aprovechar esta información para la extracción de nuevas características, mejorando también el desempeño de la etapa de clasificación.

En este sistema no se hace uso de ningún filtrado por color, en vista de mantener un esquema general para cualquier clase de señales de tránsito.

La figura 3.1 muestra el diagrama de flujos del sistema cuando trabaja con imágenes en escala de grises, y la figura 3.2 cuando trabaja con imágenes a color. Sin importar la naturaleza de la imagen de entrada, se puede explicar el funcionamiento del sistema en términos de tres módulos, los cuales son:

- Pre procesamiento, detección de bordes, reconocimiento de formas y extracción de regiones de interés.
- Extracción de características simples y ordenamiento jerárquico de la base de datos.
- Extracción de vectores de características, comparación y reconocimiento.

El objetivo de la primera etapa es extraer todas las señales de tránsito que puedan llegar a estar presentes en la imagen. Para esto, se realiza primeramente una detección de bordes en la imagen, identificando luego formas particulares que consti-

tuyen las regiones de interés, las cuales son candidatos a señales para analizar en la etapa de reconocimiento.

Luego, cada una de las regiones de interés es analizada por separado en las etapas subsiguientes. En la segunda etapa, se extraen vectores de características simples, y con estos se realiza un ordenamiento jerárquico los templates presentes en la base de datos, con el fin de aumentar la probabilidad de encontrar una coincidencia satisfactoria en las primeras comparaciones, y con el consecuente ahorro de tiempo de procesamiento. Finalmente, en la tercera etapa se calculan los vectores de características sobre la región de interés utilizando SIFT, y se comparan con los almacenados en la base de datos, lo que da un resultado del proceso de clasificación y reconocimiento.

En las siguientes secciones, se analizarán por separado cada uno de los bloques del diagrama de flujos.

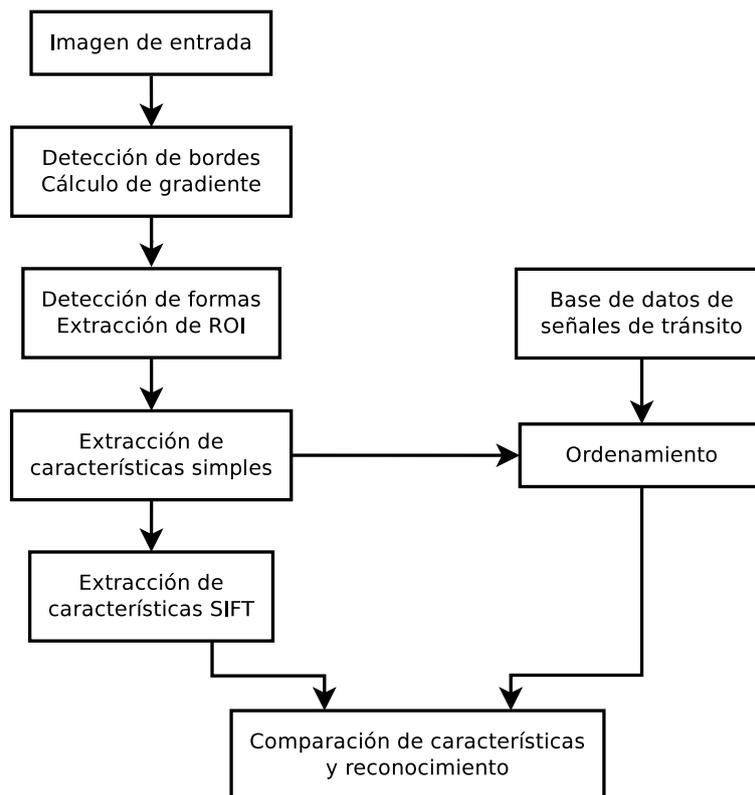


Figura 3.1: Diagrama de flujos para imágenes en escala de grises

## 3.2 Conversión del espacio de colores

Si las imágenes de entrada son a color, la información adicional se utiliza para mejorar el rendimiento respecto de la utilización de imágenes solo en escala de grises. En este caso se utilizará también una imagen en escala de grises, de la misma forma que si no estuviese disponible la imagen a color.

En el espacio de colores RGB, las componentes de color de un píxel son fuertemente dependientes del brillo, razón por la cual no es este el mejor modelo para aplicaciones en exteriores, donde las condiciones lumínicas son impredecibles y suelen cambiar rápidamente.

En contraste, en el espacio de colores HSV el color es representado mediante dos canales (Hue y Saturation), y el brillo es representado en un canal separado (Value).

En el espacio HSV, los colores son representados de manera más consistente con

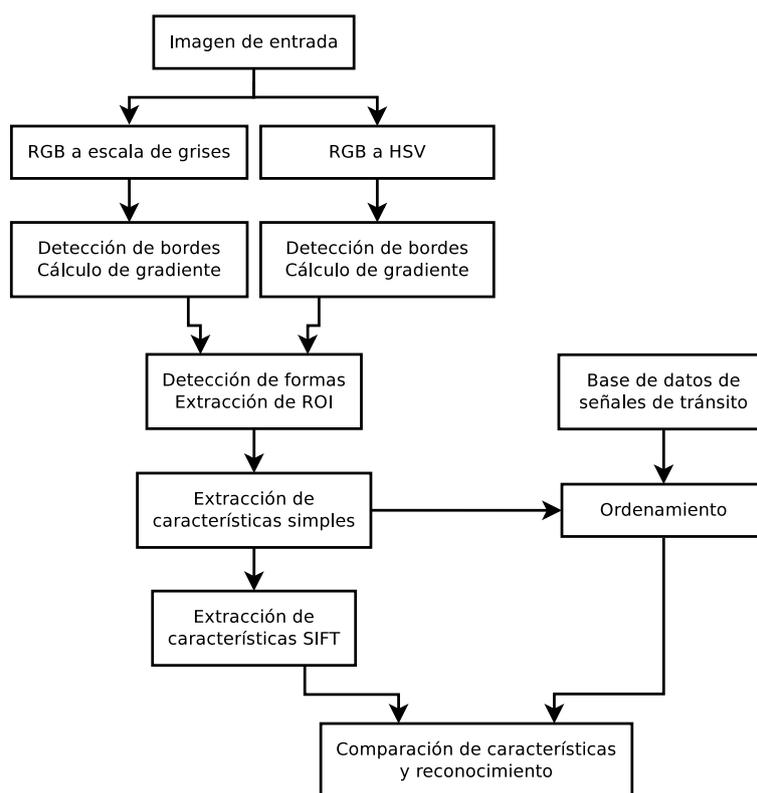


Figura 3.2: Diagrama de flujos para imágenes a color

la percepción humana del color que en el espacio RGB. Existen otros espacios de colores, como CIELAB, que representan los colores de forma más similar a la percepción humana, pero son computacionalmente más costosos que la conversión de RGB a HSV. Teniendo en cuenta todo esto, cuando se pueda mejorar alguna etapa del sistema se utilizara la información de color representada en el espacio de colores HSV.

Considerando que la mayoría de las imágenes provistas por cámaras de video están en formato RGB, será necesaria una conversión del espacio de colores. Esto no es muy costoso computacionalmente, y se realiza utilizando las ecuaciones (3.1) a (3.3).

$$V = \max \{R, G, B\} \quad (3.1)$$

$$S = \begin{cases} \frac{V - \min\{R, G, B\}}{V} & \text{si } V \neq 0 \\ 0 & \text{otros casos} \end{cases} \quad (3.2)$$

$$H = \begin{cases} \frac{60(G-B)}{S} & \text{si } V = R \\ \frac{120+60(B-R)}{S} & \text{si } V = G \\ \frac{240+60(R-G)}{S} & \text{si } V = B \end{cases} \quad (3.3)$$

### 3.3 Detección de bordes y cálculo de gradiente

Para reducir el ruido de la imagen, previo al detector de borde se aplica sobre la misma una operación de suavizado, utilizando un filtro Gaussiano. Después de esto, se utiliza el detector de bordes de Canny para generar una imagen binaria, compuesta por todos los bordes de la imagen original.

Luego de muchas observaciones realizadas sobre imágenes reales se observo que, para las señales de tránsito, el canal de saturación generalmente presenta valores constantes y muy altos en contraste con la imagen de fondo. Esta característica hace posible la detección de señales de tránsito en áreas en las que la imagen en escala de grises no, dada la similitud entre los valores de grises de la señal de tránsito y los objetos que la rodean. Por lo tanto, si se dispone de la información de color, se añaden a la imagen de bordes proveniente de la imagen en escala de grises, los bordes encontrados sobre el canal de saturación de la imagen HSV.

En la etapa de reconocimiento de formas, será necesario conocer también los valores de dirección del gradiente en los puntos que rodean a los bordes. Estos valores son calculados aplicando una operación de dilatación sobre la imagen de bordes, para obtener una máscara que se utiliza para el cálculo de la orientación del gradiente en los puntos de interés. En el caso de imágenes a color, se realiza

este proceso tanto sobre la imagen de escala de grises como la imagen del canal de saturación, y las imágenes resultantes son combinadas en una única imagen de un solo canal.

Con este procedimiento, se obtienen tres imágenes: una que contiene los bordes, una con la ubicación de los puntos de interés del gradiente (la cual es una máscara binaria), y otra con la orientación del gradiente en las ubicaciones indicadas por la máscara. Al finalizar el proceso, se obtienen estas 3 imágenes, sin importar el tipo de imagen de entrada (color o escala de grises).

### 3.4 Detección de formas

Las señales de tránsito están diseñadas con formas fijas, como círculos, rectángulos, triángulos y octágonos. Teniendo en cuenta esto para la selección de regiones de interés de la imagen original, se implementa un algoritmo que busca estas formas particulares.

Para reconocer las formas circulares se utiliza la simetría radial que estos poseen, en forma similar a la mencionada en [13].

A través de un sistema de votación, y para cada pixel de la máscara de gradiente, se afectan positivamente en un acumulador los pixeles que se encuentran en la dirección indicada por la orientación del gradiente. Ya que no se sabe si la señal de tránsito posee mayor o menor brillo que su fondo (en un caso el gradiente apunta hacia el interior del círculo, y en el otro hacia afuera), los pixeles en la dirección opuesta a la indicada por el gradiente también son afectados positivamente.

Como se están buscando señales de tránsito, existe un tamaño máximo que estas pueden llegar a tener en la imagen, y un tamaño mínimo por debajo del cual no es posible reconocer su contenido. Teniendo en cuenta esto, los pixeles afectados están restringidos a una determinada vecindad centrada en el punto que se toma en cuenta. Esto se expresa matemáticamente mediante la ecuación (3.4).

$$\bar{p}_i = \bar{p}_e \pm \text{int} \left( \frac{\nabla I(\bar{p}_e)}{|\nabla I(\bar{p}_e)|} k \right) \quad ; \quad 0 \leq k \leq R_{max} \quad (3.4)$$

Después de realizada la votación, los centros de los círculos que contiene la imagen original se encuentran localizando máximos locales mayores que un determinado umbral y que se encuentren por encima del valor medio de una ventana centrada en el punto en consideración. Esto se lleva a cabo substrayendo del acumulador una versión suavizada de sí mismo, para eliminar regiones que posean bordes ruidosos que no provengan de formas simétricas, las cuales resultan en valores de acumulador elevados pero no poseen máximos distintivos.

Para encontrar el radio de cada círculo, se toma una ventana centrada en cada punto encontrado como centro, y se seleccionan aquellos pixeles que poseen un valor de dirección de gradiente que apunte hacia ese centro (ver ecuación (3.5)), puntos con los cuales se procede al cálculo del radio del círculo, mediante la ecuación (3.6).

$$f_{\bar{p}_e} = \begin{cases} 1 & \text{si } \frac{|(\bar{p}_c - \bar{p}_e) \cdot \nabla I_{(\bar{p}_e)}|}{|(\bar{p}_c - \bar{p}_e)| \cdot |\nabla I_{(\bar{p}_e)}|} \geq 1 - \varepsilon \\ 0 & \text{otros casos} \end{cases} \quad (3.5)$$

$$R = \frac{\sum |\bar{p}_e| \cdot f_{(\bar{p}_e)}}{\sum f_{(\bar{p}_e)}} \quad (3.6)$$

Una vez detectados los círculos y calculado el radio de los mismos, el área dentro de cada uno de ellos constituye una de las regiones de interés, que será analizada más adelante en la etapa de clasificación. Para pasar las regiones de interés a la etapa de clasificación, se crea una máscara con todos los círculos encontrados en la imagen, y se aplica luego sobre esta una operación de dilatación, para asegurar que la totalidad de la señal de tránsito se encuentre contenida dentro de la máscara.

### 3.5 Extracción de características simples - Ordenamiento jerárquico

Cuando se utilizan métodos basados en comparación de vectores de características con el objetivo de clasificar muchas señales (como es el caso analizado aquí), el proceso de comparación se vuelve lento, y más lento aun cuantas más señales diferentes se encuentran en la base de datos.

El método propuesto intenta mejorar esto valiéndose de otros vectores de características, los cuales son computacionalmente rápidos para calcular y comparar, para realizar un pre ordenamiento de la base de datos.

Utilizando estos vectores de características simples, se ordena la base de datos en orden creciente de probabilidad de que sea la señal de tránsito buscada en cada momento. Este enfoque acelera el proceso de comparación, ya que cuando se compara luego con los vectores de características complejos (descriptores arrojados por el método SIFT), se tiene una mayor probabilidad de encontrar una coincidencia satisfactoria en los primeros intentos.

En el sistema implementado, se utilizan como vectores de características simples los histogramas de la región de interés analizada. Si se trata de imágenes a color, el histograma es calculado sobre el canal Hue de la imagen HSV, ya que estos histogramas presentan patrones más distintivos, puesto que cada grupo de señales de tránsito presenta colores diferentes.

Luego de equalizar y normalizar, el histograma perteneciente a la región de interés es comparado con todos los histogramas de las imágenes en la base de datos, obteniendo de cada comparación un único valor numérico que es utilizado para ordenar la base de datos.

Las métricas utilizadas para comparar los histogramas son la correlación (3.7) y  $\chi^2$  (3.8).

$$d(H_1, H_2) = \frac{\sum_i (H_1(i) - \bar{H}_1) (H_2(i) - \bar{H}_2)}{\sqrt{\sum_i (H_1(i) - \bar{H}_1)^2 \sum_i (H_2(i) - \bar{H}_2)^2}} \quad (3.7)$$

$$d(H_1, H_2) = \sum_i \frac{(H_1(i) - H_2(i))^2}{H_1(i) + H_2(i)} \quad (3.8)$$

Se pueden utilizar otras características de la imagen para realizar el pre ordenamiento de la base de datos, como por ejemplo la forma, tamaño o el número de objetos que contiene la sección extraída. Estas características adicionales deben ser rápidas de calcular, de manera de que se mantenga la idea general de este proceso, el cual no tiene el objetivo de clasificar la señal sino ordenar la base de datos.

Si se utilizan varias características diferentes como vectores de características simples, la distancia total (como un único número para efectuar el proceso de ordenamiento) se puede calcular con (3.9). Esta es una suma ponderada de las distancias individuales, donde los pesos  $\alpha_k$  son estimados de acuerdo a la importancia que se da a cada característica (histograma, forma, tamaño, etc.) frente a las otras.

$$d = \sum_{k=1}^N \alpha_k d_k \quad (3.9)$$

### 3.6 Extracción y comparación de descriptores

Luego de la etapa de ordenamiento, y utilizando SIFT, se extraen puntos clave de cada una de las regiones de interés, y se calculan los descriptores para ser comparados luego con los almacenados en la base de datos. Para encontrar los puntos clave y calcular sus descriptores asociados, se hace uso de los métodos provistos por las librerías OpenCV. También esta implementado en las librerías OpenCV el método "Speed-up robust features" (SURF), el cual es otro método para encontrar descriptores basado en wavelets, más rápido que SIFT.

Para la comparación de los descriptores de la región de interés con los de la base de datos, se utilizan también las librerías OpenCV, que tienen implementados diversos métodos de comparación. Entre ellos, se pueden usar el método *brute force*

(fuerza bruta), o el método *FLANN*, que utiliza una búsqueda por el vecino más cercano.

Para la realización de las pruebas de desempeño, se utilizaron descriptores tanto SIFT como SURF, y el proceso de comparación se realizó mediante el comparador por fuerza bruta, utilizando norma  $L_2$ .



## Capítulo 4

# Experimentos y resultados

*Este capítulo muestra los resultados del sistema implementado, y analiza los datos obtenidos.*

### 4.1 Extracción de regiones de interés

Como fue explicado en el Capítulo 3, la primer etapa del sistema consiste en una etapa de detección de bordes, que funciona tanto con imágenes en escala de grises como con imágenes a color.

En la figura 4.1 podemos observar la salida de esta etapa cuando está trabajando con imágenes en escala de grises, y se puede ver que los bordes de la señal de tránsito aparecen también en la imagen de bordes, lo que resultará muy probablemente en una correcta selección de este área como una de las regiones de interés.

Por otro lado, se puede ver en la figura 4.2 que los bordes de la señal de tránsito que aparece en esta imagen no figuran en la imagen de bordes. La razón de esto es que los valores de grises de fondo son muy similares a los valores de grises de la señal de tránsito.

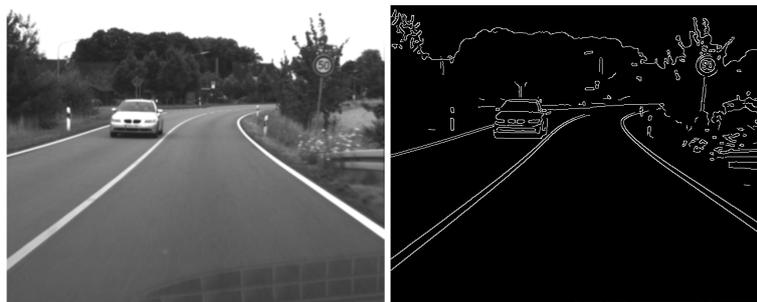


Figura 4.1: Imagen de entrada y bordes detectados por la etapa de detección de bordes, con imágenes de entrada en escala de grises.

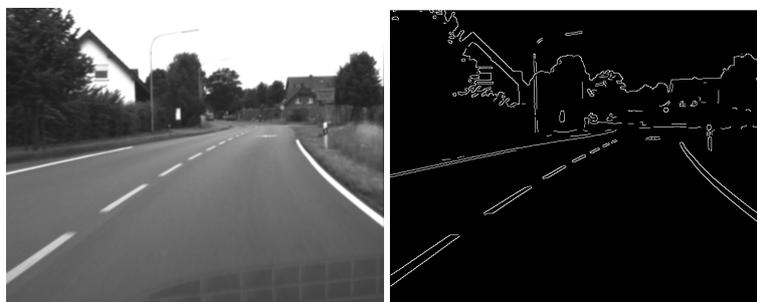


Figura 4.2: Imagen de entrada y bordes detectados por la etapa de detección de bordes, con imágenes de entrada en escala de grises.

Como una forma de mejorar el desempeño del sistema ante estas situaciones, y como fue explicado en el Capítulo 3, el método propuesto provee una mejora para la detección de señales en estas situaciones, utilizando imágenes a color pero sin la implementación de filtrado por color, lo que restringiría el uso del sistema a señales que contengan solo algunos colores predefinidos.

En las figuras 4.3 y 4.4 se puede ver que, como pasa para la mayoría de las señales de tránsito, el canal de saturación posee valores constantes y relativamente altos en contraste con el fondo que se encuentra en situaciones de la vida real. En estas dos imágenes se puede observar también que (como situación extrema) la señal podrá ser detectada únicamente a través de la imagen de bordes proveniente del canal de saturación. También es posible el caso inverso, es decir, que la señal de tránsito este enmascarada por el fondo en el canal de saturación, y que sea posible extraer bordes útiles de la imagen en escala de grises.

Entonces, si las imágenes de entrada son a color, el sistema se vuelve más robusto frente a su funcionamiento con imágenes en escala de grises en lo que refiere a la detección de regiones de interés.

Luego de calcular la imagen de bordes, y teniendo también la orientación del gradiente en los puntos que rodean a los bordes, el siguiente paso es la extracción de formas. Se ha implementado solo la detección de círculos, utilizando el algoritmo detallado en el Capítulo 3. Las imágenes correspondientes a este proceso pueden ser vistas en las figuras 4.5, 4.6 y 4.7.

Las figuras 4.5 y 4.7 muestran la imagen de entrada, la imagen que contiene la orientación del gradiente en los puntos vecinos a los bordes (como fue explicado), y el estado del acumulador luego del proceso de votación, en donde se puede observar que las áreas más blancas corresponden a los centros de los círculos presentes en la imagen de entrada. Se muestran también las imágenes del procedimiento para el

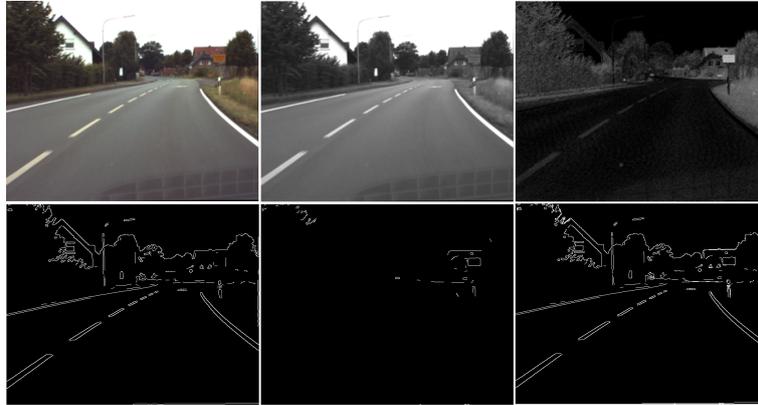


Figura 4.3: Imágenes de la etapa de detección de bordes, con imágenes a color como entrada. De izquierda a derecha y de arriba hacia abajo: imagen de entrada a color, imagen en escala de grises, canal de saturación, bordes extraídos utilizando la imagen en escala de grises, bordes extraídos utilizando el canal de saturación, y las dos imágenes de bordes fusionadas en una sola.



Figura 4.4: Imágenes de la etapa de detección de bordes, con imágenes a color como entrada. De izquierda a derecha y de arriba hacia abajo: imagen de entrada a color, imagen en escala de grises, canal de saturación, bordes extraídos utilizando la imagen en escala de grises, bordes extraídos utilizando el canal de saturación, y las dos imágenes de bordes fusionadas en una sola.

cálculo del radio del círculo, y por último, la imagen de entrada con los círculos detectados y sus centros dibujados en rojo.

Para una mejor comprensión del cálculo del radio del círculo, nos referiremos a la figura 4.6. Se puede ver a la izquierda de la figura una sub-imagen, la cual es un recorte de la imagen de orientación del gradiente, utilizando los centros de círculos encontrados, y una cierta ventana al rededor de estos. Aquí se puede ver que en éste área recortada, hay muchos pixeles distintos de cero que no pertenecen al círculo, y que no deberían intervenir en el cálculo del radio. Se seleccionan entonces todos los pixeles distintos de cero que contengan un valor de dirección del gradiente que apunte hacia el centro, los cuales serán utilizados en el cálculo del radio. Estos puntos seleccionados se consideran como pertenecientes a la circunferencia, y se muestran en la máscara resultante (imagen al centro de la figura 4.6), donde se aprecia también que hay todavía algunos pixeles que no pertenecen al círculo que se intenta detectar, pero que por su reducido número no introducen mayor error en el cálculo. La figura 4.6 muestra también el contenido de la imagen de entrada en el área que se encuentra dentro del círculo detectado (que constituye ahora una de las regiones de interés), que en este caso es una señal de tránsito.

En la figura 4.7 puede verse que círculos que no contienen señales de tránsito también son detectados, como es el caso de la lámpara de alumbrado público presente en esta imagen, que presenta forma semicircular en su sección inferior.

## 4.2 Extracción de características y clasificación

Utilizando las sub-imágenes extraídas, se realiza el proceso de comparación como fue explicado en el Capítulo 3.

Las figuras 4.8, 4.9 y 4.10 muestran el resultado de las comparaciones, en las cuales se utilizó una secuencia de imágenes a color como entrada. En las figuras 4.11, 4.12 y 4.13 se hizo uso de la misma secuencia, pero en escala de grises. En ambos casos se obtienen resultados similares. Se analizan tres situaciones diferentes, en las cuales se selecciona una imagen de la base de datos, y se la compara con las regiones de interés extraídas de la imagen de entrada. En las figuras 4.8 y 4.11, el template de la base de datos utilizado para las comparaciones es el que debería ser clasificado como si se tratase de la misma señal de tránsito. En las figuras 4.9 y 4.12, la señal de tránsito que se encuentra en la imagen de entrada es similar a la utilizada para las comparaciones (pero no se trata de la misma señal de tránsito), y en las figuras 4.10 y 4.13, la región de interés extraída de la entrada es muy diferente del template utilizado para las comparaciones. Los pares de puntos clave que se consideran coincidencias satisfactorias aparecen dibujados con líneas rectas entre los correspondientes puntos clave del template utilizado y de la región de interés extraída de la imagen de entrada.

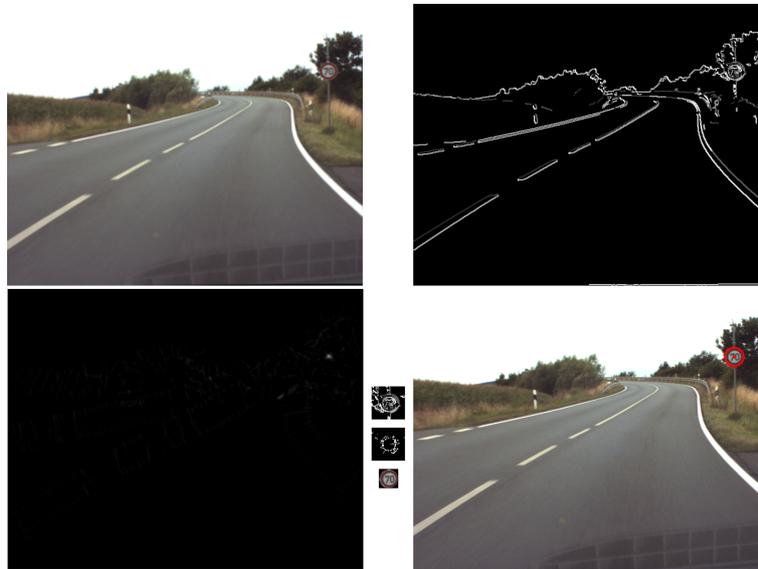


Figura 4.5: Imágenes de la etapa de detección de círculos. Arriba izquierda: imagen de entrada. Arriba derecha: imagen de orientación del gradiente. Abajo izquierda: estado del acumulador luego del proceso de votación. Abajo medio: imágenes del procedimiento de cálculo del radio. Abajo derecha: imagen de entrada con los círculos detectados dibujados en rojo.



Figura 4.6: Imágenes del procedimiento de cálculo del radio. De izquierda a derecha: área recortada de la imagen de orientación del gradiente, máscara de los píxeles que serán utilizados para el cálculo del radio, y el área circular extraída de la imagen de entrada.

En estos ejemplos se puede observar que, cuando el recorte extraído de la imagen de entrada contiene la misma señal de tránsito que el template utilizado, se encuen-

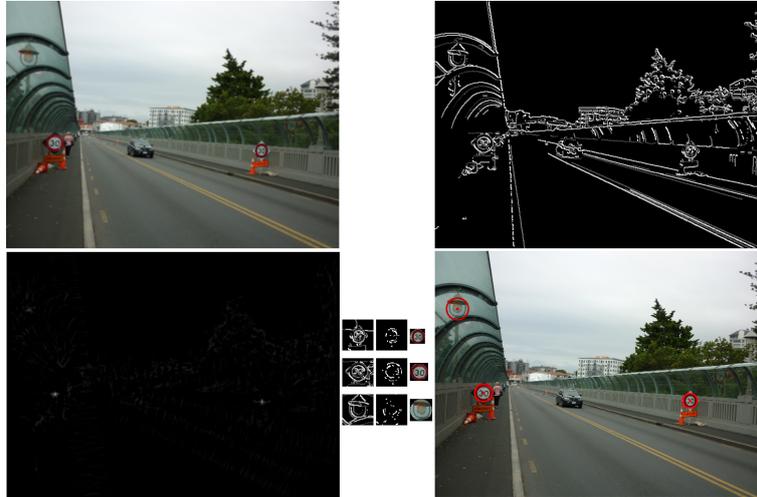


Figura 4.7: Imágenes de la etapa de detección de círculos. Arriba izquierda: imagen de entrada. Arriba derecha: imagen de orientación del gradiente. Abajo izquierda: estado del acumulador luego del proceso de votación. Abajo medio: imágenes del procedimiento de cálculo del radio. Abajo derecha: imagen de entrada con los círculos detectados dibujados en rojo.

tran muchas coincidencias entre puntos clave, mientras que si el template utilizado es muy diferente, no se encuentra coincidencia alguna. En el caso de que el recorte extraído y el template utilizado sean similares, se encuentran algunas coincidencias pero en una cantidad sensiblemente menor que si fuese comparado contra el template que contiene la misma señal de tránsito. El comportamiento del clasificador en este último caso es el esperado, ya que señales de tránsito similares comparten algunas características, por lo que se puede llegar a dar la coincidencia de algunos puntos clave.

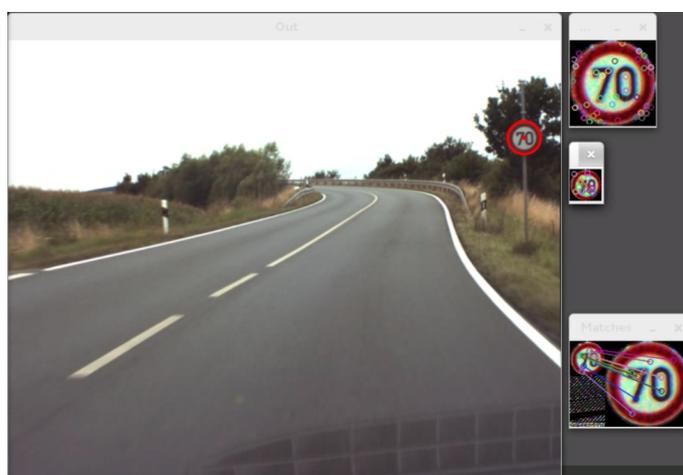


Figura 4.8: Comparación entre la imagen extraída y su correspondiente señal de tránsito de la base de datos. Izquierda: Imagen de entrada, con la región de interés extraída dibujada en rojo. Derecha: de arriba hacia abajo, template utilizado para las comparaciones con sus puntos clave marcados con círculos, región de interés extraída con sus puntos clave, y coincidencias encontradas entre el template y la región de interés extraída.

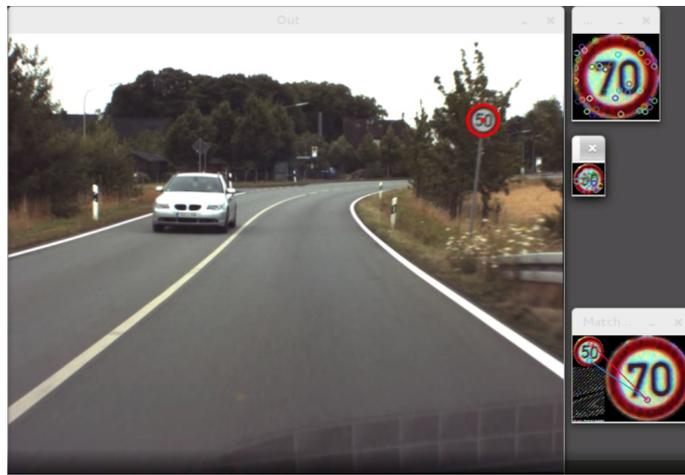


Figura 4.9: Comparación entre la imagen extraída y una señal de tránsito similar de la base de datos. Izquierda: Imagen de entrada, con la región de interés extraída dibujada en rojo. Derecha: de arriba hacia abajo, template utilizado para las comparaciones con sus puntos clave marcados con círculos, región de interés extraída con sus puntos clave, y coincidencias encontradas entre el template y la región de interés extraída.

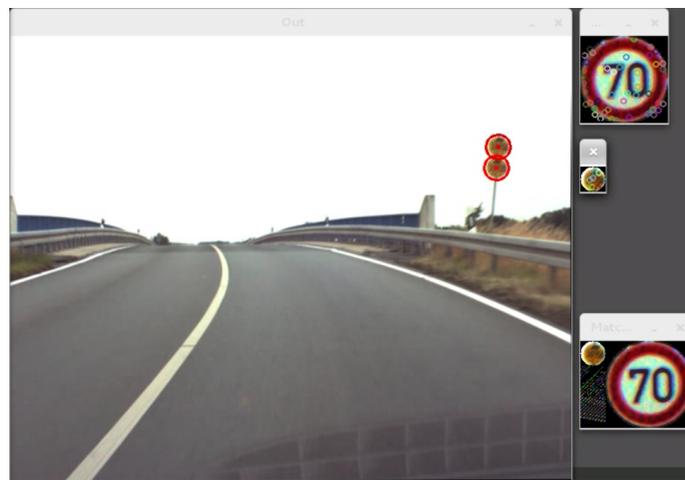


Figura 4.10: Comparación entre la imagen extraída y una señal de tránsito diferente de la base de datos. Izquierda: Imagen de entrada, con la región de interés extraída dibujada en rojo. Derecha: de arriba hacia abajo, template utilizado para las comparaciones con sus puntos clave marcados con círculos, región de interés extraída con sus puntos clave, y coincidencias encontradas entre el template y la región de interés extraída.



Figura 4.11: Comparación entre la imagen extraída y su correspondiente señal de tránsito de la base de datos. Izquierda: Imagen de entrada, con la región de interés extraída dibujada en rojo. Derecha: de arriba hacia abajo, template utilizado para las comparaciones con sus puntos clave marcados con círculos, región de interés extraída con sus puntos clave, y coincidencias encontradas entre el template y la región de interés extraída.



Figura 4.12: Comparación entre la imagen extraída y una señal de tránsito similar de la base de datos. Izquierda: Imagen de entrada, con la región de interés extraída dibujada en rojo. Derecha: de arriba hacia abajo, template utilizado para las comparaciones con sus puntos clave marcados con círculos, región de interés extraída con sus puntos clave, y coincidencias encontradas entre el template y la región de interés extraída.



Figura 4.13: Comparación entre la imagen extraída y una señal de tránsito diferente de la base de datos. Izquierda: Imagen de entrada, con la región de interés extraída dibujada en rojo. Derecha: de arriba hacia abajo, template utilizado para las comparaciones con sus puntos clave marcados con círculos, región de interés extraída con sus puntos clave, y coincidencias encontradas entre el template y la región de interés extraída.

## Capítulo 5

---

# Conclusiones

*Este capítulo expone las conclusiones acerca del método propuesto, analizando los resultados obtenidos y presentando nuevas ideas para mejorar el sistema implementado.*

En este trabajo, se propuso un método para detectar y reconocer señales de tránsito. Este método fue diseñado basado en parte en sistemas existentes, tratando de mejorarlos en lo que hace a robustez y velocidad, sin perder la flexibilidad que poseen algunos sistemas de permitir el agregado dinámico de nuevas señales de tránsito a reconocer. El otro objetivo fue mantener una generalización, en el sentido de que el sistema debe ser capaz de detectar y reconocer una amplia variedad de señales de tránsito. Quedan todavía funciones y características por implementar, como funciones para el manejo de la base de datos o la detección de otras formas en la imagen, pero los resultados obtenidos hasta el momento son prometedores. Otras tareas que deben ser realizadas en un futuro son un análisis completo de la velocidad del algoritmo en diferentes situaciones, y una evaluación de desempeño de la etapa de clasificación.

Como fue visto en el Capítulo 4, hay algunas señales que arrojan coincidencias en diferentes puntos cuando se las compara con la misma señal de tránsito de la base de datos, y también (aunque en menor medida) con algunas señales de similares características. Este no es un problema mayor, dado que si se encuentran algunas coincidencias entre la señal de tránsito extraída y alguna otra similar de la base de datos, cuando esta misma señal a comparar sea contrastada con su correspondiente señal de la base de datos el número de puntos que coincidan será mayor, y la señal será clasificada correctamente. Un análisis más profundo de esta situación puede llevarnos a una nueva forma de agrupar las señales de la base de datos, previo a ser comparadas con cada nueva región de interés extraída de la imagen de entrada, dando una guía acerca de métodos más eficaces para realizar el ordenamiento de la base de datos.

Hay algunas consideraciones y observaciones acerca de la utilización de descrip-

tores SIFT en la etapa de clasificación. Una de ellas, es que este método es sensible a diferentes variaciones de la misma señal. Si la base de datos solo consiste en los templates extraídos del sitio de la NZTA (u otro conjunto de señales de tránsito ideales), la cantidad de puntos coincidentes entre el parche extraído de la imagen de entrada y la señal de la base de datos será demasiado baja. Este problema se puede resolver teniendo no una única muestra de cada señal para conformar la base de datos, sino añadiendo a la misma imágenes de señales de tránsito reales. Esto puede ser implementado dentro del mismo sistema, guardando las imágenes que no pueden ser clasificadas e indicando en forma manual a que señal de tránsito pertenecen, para añadirlas luego a la base de datos.

El otro problema que se observo, es que existen señales de tránsito que no poseen una cantidad apreciable de puntos clave, y en consecuencia no pueden ser clasificadas apropiadamente a causa de la falta de buenos descriptores. Como propuesta para resolver esto, se pueden utilizar otros métodos de clasificación en los casos en que no se encuentren muchos puntos clave en la región de interés. Incluso si este método de clasificación es más lento que el implementado mediante la utilización de SIFT, no sería utilizado en todo momento, por lo que no afectaría significativamente la velocidad media del sistema.

---

## Bibliografía

- [1] Fleyeh H. and Dougherty M.: Traffic Sign Classification Using Invariant Features and Support Vector Machines *IEEE Intelligent Vehicles Symposium*, pages 530-535, 2008.
- [2] Kuş M. C., Gökmen M. and Etaner-Uyar Ş.: Traffic Sign Recognition using Scale Invariant Feature Transform and Color Classification In Proc. *Int. Symp. Computer Information Sciences*, 2008
- [3] Khan J. F., Bhuiyan S. M. A., and Adhami R. R.: Image Segmentation and Shape Analysis for Road-Sign Detection *IEEE Trans. Intelligent Transportation Systems* Vol. 12, No. 1, March 2011
- [4] Weijie Liu and Maruya K.: Detection and Recognition of Traffic Signs in Adverse Conditions In Proc. *IEEE Intelligent Vehicles Symp.*, pages 335-340 (2009)
- [5] Yanlei Gu, Yendo T., Tehrani M. P., Fujii T., and Tanimoto M.: A New Vision System for Traffic Sign Recognition In Proc. *IEEE Intelligent Vehicles Symp.*, pages 7-12 (2010)
- [6] Yok-Yen Nguwi and Kouzani A. Z.: Automatic Road Sign Recognition Using Neural Networks *International Joint Conference on Neural Networks*, pages 3955-3962, 2006
- [7] Pazhoumand-Dar H. and Yaghobi M.: DTBSVMs: a New Approach for Road Sign Recognition *2nd International Conference on Computational Intelligence, Communication Systems and Networks*, pages 314-319, 2010
- [8] Bahlmann C., Zhu Y., Visvanathan R., Pellkofer M. and Koehler T.: A System for Traffic Sign Detection, Tracking, and Recognition Using Color, Shape, and Motion Information In Proc. *IEEE Intelligent Vehicles Symposium*, pages 255-260, 2005
- [9] Ruta A., Porikli F., Watanabe S. and Li Y.: In-vehicle Camera Traffic Sign Detection and Recognition *Machine Vision and Applications*, Vol. 22 Iss. 2, pages 359-375, March 2011
- [10] Ren F., Huang J., Jiang R. and Klette R.: General Traffic Sign Recognition by Feature Matching In Proc. *IVCNZ*, pages 409-414, 2009
- [11] HSL and HSV, [http://en.wikipedia.org/wiki/HSL\\_and\\_HSV](http://en.wikipedia.org/wiki/HSL_and_HSV). Last visit: Feb. 2012

- [12] Gonzalez R. C., Woods R. E.: *Digital Image Processing* 2nd ed., Prentice Hall (2002)
- [13] Loy G. and Zelinsky A.: A Fast Radial Symmetry Transform for Detecting Points of Interest *IEEE Trans Pattern Analysis and Machine Intelligence*, Vol. 25, No. 8, pages 959-973, Aug. 2003
- [14] New Zealand Transport Agency Website, [www.nzta.govt.nz](http://www.nzta.govt.nz). Last visit: Feb. 2012
- [15] Lowe, D. G.: Distinctive Image Features from Scale-Invariant Keypoints *International Journal of Computer Vision*, Vol. 60, No. 2, pages 91-110, Nov. 2004
- [16] Smith D. and Harvey R.: Document Retrieval Using SIFT Image Features. *Journal of Universal Computer Science*, Vol. 17, No. 1, pages 3-15, 2011

---

## Indice

- Agencia de Transporte de Nueva Zelanda, 4
- color
  - definiciones
    - brillo, 7
    - cromacidad, 8
    - hue, 7
    - luminosidad, 8
    - matiz, 7
    - saturación, 8
    - value, 8
  - espacios de, 7
  - modelos de, 7
    - HSV, 7
    - RGB, 7
- conjuntos
  - intersección, 9
  - traslación, 9
  - unión, 9
- conversión RGB a HSV, 19
- detector de bordes Canny, 19
- Driver Assistance Systems, 4
- histogramas
  - comparación, 21
    - $\chi^2$ , 22
  - correlación, 22
- morfología matemática, 8
  - elemento estructurante, 9
  - operaciones
    - dilatación, 10
    - erosión, 9
- New Zealand Transport Agency, 4
- reconocimiento de señales de tránsito, 14
- scale-invariant feature transform, 11
- señales de tránsito
  - de advertencia, 4
  - de información, 4
  - obligatorias, 4
- SIFT, 11
  - asignación de orientación, 13
  - comparación, 22
  - descriptor del punto clave, 13
  - detección de extremos en el espacio es-  
cala, 12
  - localización de puntos clave, 12
- sistemas de ayuda al conductor, 4