

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR



TRABAJO FIN DE GRADO

Modelado y Detección de Elementos de Interés en Secuencias de Vídeo de Carreteras mediante Técnicas de Visión Artificial

Adrián Tomé Alonso

Tutor: Luis Salgado Álvarez de Sotomayor

Julio 2014

Agradecimientos

En primer lugar quiero dar las gracias a mi tutor Luis Salgado por haber conseguido que diera lo mejor de mí, y sobre todo por ayudarme hasta el último día. Quiero agradecer también a mis compañeros de teleco todos esos inolvidables días de risas entre clases y estudio, en especial a vosotros, Erik, Ángel y Paula, porque habéis demostrado ser grandes compañeros y mejores amigos. No puedo olvidarme de mi familia, que siempre me ha apoyado, en los buenos y en los malos momentos. Sin vosotros jamás hubiera llegado hasta aquí. Por último, gracias a ti Noelia, porque me has dado fuerza durante todos estos años para seguir adelante.

De nuevo, gracias.

Resumen

El objetivo de este trabajo de fin de grado es desarrollar diferentes estrategias de modelado y segmentación de secuencias de carretera en autopistas que permitan su uso en sistemas de ayuda a la conducción. Para ello, se modelarán los elementos de interés que deseemos identificar, empleando dichos modelos para la clasificación Bayesiana de sus píxeles.

Se han estudiado e implementado un total de cuatro estrategias, de forma que se añaden características a cada una para obtener la siguiente y mejorar los resultados. Estas estrategias utilizan técnicas usadas en diferentes trabajos dentro del área y otras técnicas desarrolladas para este trabajo en concreto, de forma que se usen todas en conjunción con el fin de obtener los mejores resultados.

Todos los resultados se han analizado metódicamente estudiando las ventajas y las desventajas de cada una de las estrategias con el fin de poder avanzar hacia la siguiente mejorando aspectos de la anterior. Así, se concluye en una estrategia final, de la que se analizarán los resultados validando su funcionamiento a través de pruebas tanto objetivas como subjetivas basadas en secuencias reales y anotaciones que constituyen segmentaciones de referencia.

Palabras clave

Modelado de carretera, segmentación, elementos de interés, región de interés, ayuda a la conducción, clasificación, Bayes, visión artificial, pavimento, líneas de carretera, vehículos, detección.

Abstract

The aim of this final degree project is to develop different strategies of modeling and segmentation of road sequences on highways that allow its use on driving assistance systems. To do so, the elements of interest we expect to identify will be modeled using them for the Bayesian classification of the pixels.

There have been studied and implemented a total of four strategies, in such a way that features are added to each of them in order to obtain the next one and improve outcomes. These strategies apply techniques previously used in other projects in the field as well as techniques specifically developed for this project, in such a way that all of them are used in conjunction to obtain the best results.

All results have been analyzed methodically, studying the advantages and disadvantages of each strategy, in order to move forward to the next one by improving aspects of the previous. Thus, it is concluded in a final strategy, whose results will be analyzed validating its performance through subjective and objective proofs based in real sequences and annotations that constitute reference segmentations.

Keywords

Road modeling, segmentation, elements of interest, region of interest, driving assistance, classification, Bayes, artificial vision, pavement, lane marking, vehicles, detection.

Índice general

Índice de Figuras	VI
Índice de Tablas.....	VII
Índice de Ecuaciones y Fórmulas	VIII
Glosario	VIII
1. Introducción.....	1
1.1. Motivación	1
1.2. Objetivos.....	1
1.3. Organización de la memoria	2
2. Trabajos relacionados	3
2.1. Líneas de carretera (Lane Marking).....	3
2.2. Detección de Vehículos	4
2.3. Modelado del conjunto de elementos de interés.....	4
3. Marco teórico.....	5
3.1. Criterios de clasificación	5
3.2. Modelado de las clases	6
4. Estrategias de modelado y segmentación	7
4.1. Primera aproximación basada en luminancia	7
4.1.1. Criterio de clasificación ML.....	14
4.1.2. Criterio de clasificación MAP	15
4.2. Segunda aproximación añadiendo la clase objetos.....	16
4.2.1. Criterio de clasificación ML.....	17
4.2.2. Criterio de clasificación MAP	18
4.3. Tercera aproximación	18
4.3.1. Criterio de clasificación ML.....	22
4.3.2. Criterio de clasificación MAP	22
4.4. Cuarta aproximación. Actualización de los modelos.....	23
4.4.1. Verificación de problemas en los modelos.....	24
4.4.2. Criterios de actualización de los modelos	26
5. Pruebas y resultados	29
5.1. Pruebas subjetivas.....	31
5.2. Pruebas objetivas	32
5.3. Análisis de las pruebas.....	35
6. Conclusiones y Trabajo futuro	37
Referencias	39

Índice de Figuras

Figura 4.1 Diagrama de flujo de la estrategia general de clasificación de los píxeles	7
Figura 4.2 Diagrama de flujo de la primera estrategia general de segmentación, haciendo uso de la luminancia de los píxeles de la imagen	7
Figura 4.3 Imagen original de la secuencia de vídeo con la región de interés y la región de modelado	8
Figura 4.4 Diagrama funcional del modelado basado en luminancia	9
Figura 4.5 Conjunto de celdas $c_{nr,nc}$. Cada una de estas celdas contiene el histograma del bloque de I_{ROI} correspondiente. El sentido de la flecha indica los histogramas de las celdas que se compararán.	10
Figura 4.6 Distancias entre las celdas y distancias en dos agrupaciones	11
Figura 4.7 Imagen con el pavimento a nivel cero I_{segP} y la misma imagen con umbral adaptativo	12
Figura 4.8 Curva del umbral adaptativo	12
Figura 4.9 Modelo de la clase pavimento e histograma de I_{ROI} con el modelo superpuesto	13
Figura 4.10 Modelos de las clases de la primera estrategia de segmentación (P,L,U).....	13
Figura 4.11 Imagen segmentada mediante la primera aproximación usando ML y mediante MAP	14
Figura 4.12 Primera estrategia de segmentación con la inclusión de información a priori	15
Figura 4.13 Diagrama de flujo de la segunda estrategia de segmentación.....	16
Figura 4.14 Modelos de la estrategia, con el nuevo modelo para objetos.....	17
Figura 4.15 Imagen segmentada con el conjunto de clases $S = P, L, O, U$ clasificando con ML y MAP	17
Figura 4.16 Diagrama de flujo de la tercera aproximación con modelado de respuesta a filtro detector de líneas	18
Figura 4.17 Respuesta del filtro en las filas marcadas en rojo. El filtro reacciona en píxeles de líneas divisorias	20
Figura 4.18 Histograma de la respuesta al filtro detector de líneas.....	20
Figura 4.19 Modelos de las clases L, O y P para la respuesta al filtro detector de líneas.	21
Figura 4.20 Clasificación de los píxeles de I_{ROI} según el criterio de ML y según el criterio MAP	22

Figura 4.21 Diagrama de flujo utilizando los atributos L e I añadiendo información a priori	22
Figura 4.22 Estrategia en la actualización del modelo	23
Figura 4.23 Imágenes de una misma secuencia y sus imágenes clasificadas (I).....	24
Figura 4.24 Histogramas de algunas imágenes de la secuencia con iluminación normal, con aumento de la iluminación y reducción de la iluminación	25
Figura 4.25 Imágenes de una misma secuencia y sus imágenes clasificadas (II)	28
Figura 5.1 Fotogramas de la secuencia de pruebas 1 con imágenes originales y clasificadas con la aprox. 4 y 3	30
Figura 5.2 Fotogramas de la secuencia de pruebas 2 con imágenes originales y clasificadas con la aprox. 4 y 3	31
Figura 5.3 Fotogramas de la secuencia de pruebas 3 con imágenes originales y clasificadas con la aprox. 4 y 3	31

Índice de Tablas

Tabla I Porcentaje de píxeles asignados a cada clase usando los criterios de clasificación ML y MAP. Secuencia 1	26
Tabla II Porcentaje de píxeles asignados a cada clase usando los criterios de clasificación ML y MAP. Secuencia 2	27
Tabla III Porcentaje de píxeles asignados a cada clase usando los criterios de clasificación ML y MAP. Secuencia 3	27
Tabla IV Aciertos y errores de la secuencia de pruebas uno con la aproximación 4	32
Tabla V Aciertos y errores de la secuencia de pruebas dos con la aproximación 4	33
Tabla VI Aciertos y errores de la secuencia de pruebas tres con la aproximación 4	33
Tabla VII Aciertos y errores de la secuencia de pruebas uno con la aproximación 3	34
Tabla VIII Aciertos y errores de la secuencia de pruebas dos con la aproximación 3	34
Tabla IX Aciertos y errores de la secuencia de pruebas tres con la aproximación 3	34

Índice de Ecuaciones y Fórmulas

(1) Teorema de Bayes.....	5
(2) Distribución normal o gaussiana (como verosimilitud).....	6
(3) Distancia de Bhattacharyya.....	9
(4) Función de Umbral Adaptativo.....	12
(5) Clasificación de píxeles de la clase P (I).....	12
(6) Probabilidad de la clase X_i	15
(7) Clasificación de píxeles de la clase P (I).....	16
(8) Verosimilitud como producto de las verosimilitudes de I_{xy} y L_{xy}	18
(9) Filtro detector de líneas de carretera.....	19

Glosario

ADAS: *Advanced Driving Assistance Systems*, conjunto de sistemas instalados en vehículos, encargados de ayudar en la conducción.

ML: *Maximum Likelihood*, criterio de clasificación según máxima verosimilitud.

MAP: *Maximum a Posteriori*, criterio de clasificación que busca la máxima probabilidad a posteriori.

IPM: *Inverse Perspective Mapping*, técnica que permite cambiar la perspectiva de una imagen.

ROI: *Region of Interest*, región de una imagen en la que se encuentran los elementos de interés en las imágenes de carretera.

1. Introducción

1.1. *Motivación*

Actualmente, los llamados sistemas de ayuda a la conducción (*Advanced Driver Assistance Systems*; ADAS) han cobrado gran importancia en el mundo de la automoción. Estos métodos obtienen información que hasta hace relativamente poco tiempo no era factible obtener del entorno de los vehículos, debido en gran medida al coste y la complejidad de los sensores a instalar. La reducción de los costes de estos sistemas, así como los avances en las técnicas y algoritmos usados en los mismos, han permitido usarlos de manera cada vez más frecuente.

Los ADAS utilizan la información que se obtiene a través de diversos sensores con el objetivo de mejorar aspectos de la conducción, tales como la seguridad o la eficiencia del vehículo y en definitiva, hacer más sencilla la tarea del conductor a la par que salvar vidas.

Algunos de estos sistemas obtienen información a partir del análisis de secuencias de vídeo obtenidas de cámaras instaladas en el vehículo. Existen diferentes configuraciones en la posición de las cámaras, si bien la fundamental orienta el campo de visión hacia la carretera en el sentido del desplazamiento del vehículo.

Las secuencias capturadas se analizan por sistemas específicos que son capaces de extraer de manera automática información de los elementos de interés. Estos elementos de interés pueden ser muy variados dependiendo del objetivo concreto del sistema en cuestión, si bien algunos son fundamentales en la mayor parte de los ADAS. Concretamente, la carretera, así como los vehículos que circulan por la misma, centran gran parte de los trabajos de investigación. A su vez, la carretera contendrá varios elementos de interés, como son el pavimento, y las líneas divisorias de los carriles.

La región de interés (*Region of Interest*, ROI), es el conjunto de píxeles de las imágenes de la secuencia que contiene los elementos cuyos píxeles se busca clasificar en clases. Detectarlos de forma precisa y eficiente mediante técnicas de segmentación de imágenes será crucial, puesto que el sistema interpretará la clasificación realizada y será capaz de actuar en consecuencia, por ejemplo, evitando una colisión o manteniendo la distancia de seguridad con los demás vehículos y en el caso de las líneas divisorias, evitar cambios involuntarios de carril o salidas de la vía.

1.2. *Objetivos*

El objetivo de este trabajo de fin de grado consiste realizar el análisis de distintas estrategias para la detección de los elementos de interés en la carretera, concretamente el pavimento, las líneas divisorias de los carriles y vehículos. El resultado será una imagen segmentada donde sus píxeles estarán asociados a las diferentes clases que definan los elementos de interés.

Ya que existen multitud de factores que pueden alterar el funcionamiento de nuestro sistema, va a haber diversas complejidades que se deberán sortear para que el algoritmo funcione de forma efectiva en la mayor parte de las situaciones posibles.

Entre estas dificultades se encuentran:

- Elegir las mejores estrategias de modelado y segmentación de los elementos de interés.
- Especificación de los parámetros óptimos de los algoritmos para obtener los mejores resultados.
- Conseguir que el sistema permita adaptarse a cambios de iluminación que alteren los modelos desarrollados (i.e. Entrada a túneles, diferentes clases de pavimento, líneas,...).
- Adaptación a diferentes anchuras de líneas divisorias.

Se comenzará por realizar aproximaciones sencillas a la resolución del problema que se irán mejorando y modificando a medida que se observen sus limitaciones, buscando desarrollar un algoritmo robusto que logre con el máximo éxito el objetivo final.

Así pues se analizarán rigurosamente las ventajas y desventajas de cada uno de los procesos implementados de forma justificada y detallada, para ilustrar el desarrollo del trabajo. En última instancia se harán propuestas de mejoras y trabajos futuros en el campo de la segmentación de carreteras.

1.3. Organización de la memoria

La presente memoria se estructura de acuerdo con los siguientes capítulos:

Capítulo 1: Introducción, motivación y objetivos del trabajo.

Capítulo 2: Trabajos relacionados.

Capítulo 3: Marco introductorio de la teoría usada.

Capítulo 4: Desarrollo de estrategias de modelado y segmentación.

Capítulo 5: Pruebas y análisis de los resultados.

Capítulo 6: Conclusiones del trabajo realizado y orientaciones para trabajo futuro.

2. Trabajos relacionados

Son muchos los trabajos que buscan modelar o detectar elementos de interés en carreteras. De hecho existen trabajos específicos para vías urbanas, como hacen [3] y [7] y existen otros que se encargan de la segmentación y localización de elementos de interés en carreteras interurbanas y autopistas, como se hace en [1], [4] y [5], que asumen diferentes características del entorno con respecto a los primeros.

Una aclaración que es necesario hacer, es que mientras algunos trabajos tratan de realizar una segmentación de la carretera buscando localizar y modelar varios elementos de interés que aparezcan en ellas, existen mucho otros que tratan de centrarse sólo en algunos de ellos. Este es el caso de [4], [5], [6], [7] y [8] que se centran en buscar líneas divisorias de carril o líneas de carretera. En cuanto a detección de vehículos encontramos [1], [2], [9] y [12] como trabajos representativos.

2.1. *Líneas de carretera (Lane Marking)*

Sin duda las líneas de carretera son uno de los elementos fundamentales en el análisis. Su detección y seguimiento permite realizar multitud de tareas en sistemas de ayudas a la conducción, como son alertas por cambios involuntarios de carril o salidas de las vía. Pero además, existen otras tareas para las que se necesitan detectar. Una de las más importantes es la búsqueda del punto de fuga de la carretera [8] información que permite, entre otras cosas, estabilizar imágenes con perspectiva corregida.

Para encontrar las líneas de carretera la mayor parte de los trabajos proponen la detección de líneas rectas a partir de puntos de borde extraídos de las imágenes.

Para la búsqueda de bordes en la carretera, se suele usar el algoritmo de Canny [13], como se hace en [4] y [5]. Este algoritmo es muy usado, debido sobre todo a que implementa búsqueda de bordes de diferentes grosores e implementa un filtro gaussiano que reduce notablemente el ruido [4].

En cuanto a la determinación de líneas rectas (como característica de las líneas de carretera), se realiza sobre la imagen con los bordes resaltados, ya que de esta forma se eliminan zonas de la imagen que no interesan. El método que más se usa es la transformada de Hough [13]. Esta transformada se usa en [4] y [5]. Esta transformada se basa en un sistema de votación, donde se prueban diferentes rectas que podrían encontrarse en la imagen como resultados la unión de sus píxeles. Las más votadas suelen ser las líneas divisorias de carril cuando se aplica sobre una imagen de carretera.

También existe la posibilidad de implementar un filtrado de líneas, como se hace en [1] y que se implementará en este trabajo, que al aplicarlo sobre una imagen de carretera, en las zonas donde las líneas aparecen, la respuesta al filtro es alta. Este filtro depende de la anchura de la línea en cuestión, lo que lo hace muy atractivo a la hora de buscar una anchura específica. Por otro lado, en [8] se utiliza un método de búsqueda de bordes alternativo, que hace uso de una máscara sencilla de detección de bordes que permite encontrar fácilmente las líneas de carretera.

Existen otras muchas posibilidades para detectar las líneas en imágenes de carretera. Este es el caso de sistemas basados en gradientes [6], de predicciones que buscan el lugar

de aparición más probable de las líneas de carretera [7], o incluso algún trabajo que modela una función geométrica para las líneas [14].

2.2. *Detección de Vehículos*

Los trabajos dedicados a ello son numerosos y como también ocurría con las líneas de carretera, hay multitud de métodos que se usan para detectar los vehículos. Algunos basados en descriptores y filtros como [3], [15] y [16]. En este trabajo no se va a realizar una detección tan exhaustiva, de hecho, se usará la sombra y la parte bajos los mismos para su detección. La característica más importante de la sombra es su nivel bajo de luminancia en cualquier circunstancia de iluminación, y por ello es un método bastante robusto. [15] la utiliza para la detección de vehículos aunque su uso está combinado con otras características, como pueden ser simetrías o bordes. Trabajos que tan sólo usan las sombras y las partes bajas de los vehículos son [1] (se implementa en este trabajo) y [9]. Este último utiliza las sombras para realizar un reconocimiento de patrones que ayuden a detectar el vehículo.

2.3. *Modelado del conjunto de elementos de interés*

Como ya se ha dicho, varias partes de este trabajo están basadas en el trabajo realizado de [1]. En él, se realiza una segmentación bayesiana de los elementos de interés de la carretera, que utiliza verosimilitudes para modelar cada clase y un criterio de clasificación MAP (*maximum a posteriori*) para segmentar en función de los modelos y la probabilidad de aparición de cada clase. Este tipo de aproximación se emplea también en otros trabajos [6] debido a la flexibilidad que ofrece al poder tanto introducir información a priori en la clasificación como adaptar los modelos a las características cambiantes del entorno dinámico.

En este trabajo se van a realizar segmentaciones bayesianas, donde además del modelado de las clases usando las verosimilitudes, también se usará MAP y el criterio de máxima verosimilitud (ML) para la segmentación de las imágenes.

3. Marco teórico

3.1. *Criterios de clasificación*

La segmentación de una imagen, como en este trabajo se acomete, es la tarea de clasificación de sus píxeles en un conjunto de clases que tienen unas características que los diferencian de las demás. Por lo tanto, la asignación de un píxel a una clase dependerá de cómo se ajusten las características asociadas a ese píxel al modelo establecido para la clase.

El conjunto $S = \{X_i\}$, contendrá las clases a las que puede pertenecer un píxel de la imagen. Cada una de estas clases tendrá un representante o etiqueta y el resultado de la clasificación será una imagen donde cada píxel tendrá asociada la etiqueta de la clase a la que pertenece.

El marco de clasificación que se va a utilizar en este trabajo es probabilístico y sigue la fórmula de Bayes:

$$p(X_i|z_{xy}) = \frac{p(z_{xy}|X_i) \cdot p(X_i)}{p(z_{xy})} \quad (1)$$

donde $p(X_i|z_{xy})$ representa la probabilidad de que dado el píxel z_{xy} indexado por sus coordenadas (x, y) dentro de la imagen I pertenezca a la clase X_i ; $p(z_{xy}|X_i)$ es la probabilidad de aparición del píxel z_{xy} dada la clase X_i , también conocida como función de verosimilitud y serán los modelos de cada una de nuestras clases; $p(X_i)$ es la probabilidad de aparición de las clases consideradas o información a priori y por último $p(z_{xy})$ es la evidencia, que es igual a $\sum_{i \in S} p(z_{xy}|X_i) \cdot p(X_i)$ y permite que la suma de todas las probabilidades $p(X_i|z_{xy})$ sea igual a la unidad.

Existen diferentes criterios para la clasificación. En este trabajo aplicaremos el criterio de máxima verosimilitud (*Maximum Likelihood*; ML) y el criterio MAP (*Maximum a Posteriori*). El criterio de máxima verosimilitud se basa exclusivamente en los modelos $p(z_{xy}|X_i)$ de las clases. Se selecciona como clase a la que pertenece un píxel, aquella para la que $p(z_{xy}|X_i)$ es máxima. A diferencia de este criterio, MAP permite la introducción de información a priori, representada por $p(X_i)$. Así, en este caso, se clasifica un píxel z_{xy} a la clase X_i si $p(X_i|z_{xy})$ es máxima con respecto a las probabilidades de las demás clases.

3.2. Modelado de las clases

El modelo de cada clase se asumirá que sigue una forma de gaussiana $G_{\mu_{X_i}, \sigma_{X_i}}$, es decir, vendrá caracterizado por su valor medio μ_{X_i} y su desviación típica σ_{X_i} [12].

$$p(z_{xy}|X_i) \propto \exp\left(-\frac{1}{2\sigma_{X_i}^2}(z_{xy} - \mu_{X_i})^2\right) \quad (2)$$

La obtención de los parámetros de cada modelo se realizará identificando en primer lugar los píxeles que pertenezcan a cada clase y calculando el valor medio y desviación típica de sus atributos. Se considerarán como atributos en este trabajo, el valor de luminancia, así como la respuesta a un filtro detector de líneas divisorias de carril que se detallará más adelante.

4. Estrategias de modelado y segmentación

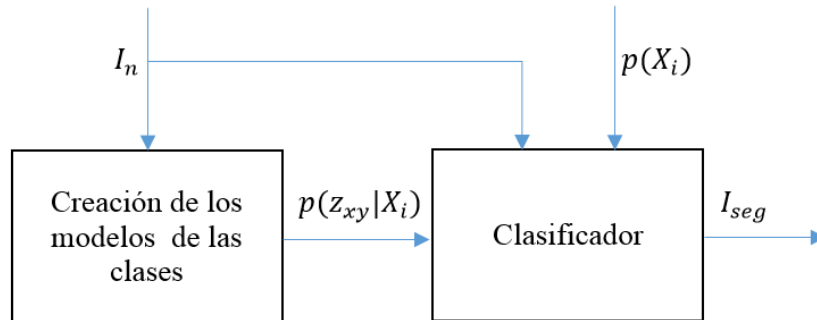


Figura 4.1 Diagrama de flujo de la estrategia general de clasificación de los píxeles

En la Figura 4.1 se muestra el diagrama con la estrategia general de segmentación que se va a seguir en este trabajo. Existen dos tareas principales a las que se hará frente. La primera de ellas es la obtención de los modelos de las clases a las que pueden pertenecer los píxeles de la imagen de entrada I_n de la secuencia.

Los modelos obtenidos entrarán en el segundo bloque, que lleva a cabo la tarea de clasificación de los píxeles de la imagen de entrada asignando a la imagen de salida I_{seg} una etiqueta para cada píxel en función de la clase a la que pertenezca. Para ello, puede usarse el criterio de máxima verosimilitud (ML) y el criterio MAP. En los siguientes apartados, se detallan diferentes métodos para la creación de los modelos de las clases y por cada una de ellos, se hará la clasificación de los píxeles con los dos métodos anteriores, valorando las diferencias entre uno y otro.

4.1. Primera aproximación basada en luminancia

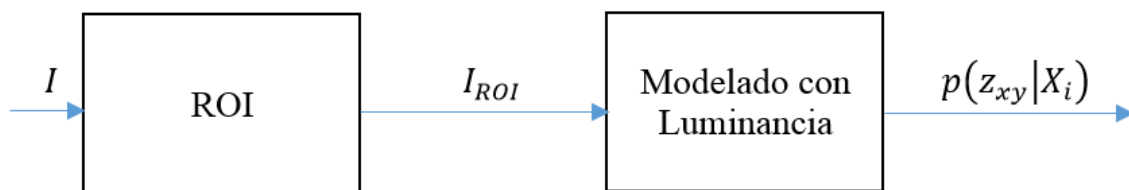


Figura 4.2 Diagrama de flujo de la primera estrategia general de segmentación, haciendo uso de la luminancia de los píxeles de la imagen

En la Figura 4.2 se encuentra el diagrama de esta primera aproximación, donde el atributo que se usará para realizar los modelos es el valor de luminancia ($z_{xy} = I_{xy}$) de los píxeles. El conjunto de clases considerado será $S = \{P, L, U\}$, correspondientes al pavimento (*Pavement*, P), líneas divisorias de carril o líneas de carretera (*Lane Marking*, L) y “Desconocido” (*Unknown*) respectivamente.



Figura 4.3 Imagen original de la secuencia de vídeo con la región de interés y la región de modelado

En las imágenes de entrada, (Figura 4.3) hay multitud de elementos distinguibles. El primero de ellos y el más evidente es la calzada. Prácticamente la mitad del fotograma está formado por píxeles pertenecientes a esta. La calzada contiene a su vez las líneas de carretera. El segundo elemento más importante son los vehículos de la calzada, y por último el resto de elementos que no son necesarios detectar en la segmentación. Estos elementos pueden ser cielo, farolas, árboles,... que se encuentran en la mitad superior de la imagen.

La perspectiva de adquisición permite conocer a priori que la mayor parte de los elementos de interés se encuentran en la mitad inferior de la imagen. Esto hace que para modelar las diferentes clases se pueda restringir el análisis a una región de la imagen, denominada región de interés o ROI (del acrónimo de su denominación en inglés, *Region of Interest*). Por lo tanto, la funcionalidad del primer bloque de la Figura 4.2 es el establecimiento de esta región sobre la que se realizará el modelado y segmentación posterior.

Esta región es configurable, si bien asumiendo que la cámara está orientada al punto de fuga de la carretera, por simplicidad en este trabajo se ha fijado a fija la mitad inferior de la imagen. La imagen de la región de interés está recuadrada en rojo en la Figura 4.3.

La Figura 4.4 muestra en detalle la estrategia de modelado propuesta de las clases P y L usando I_{ROI} , (como a partir de ahora se denominará la imagen con la región de interés de cada imagen) es decir, la estrategia de obtención de $p(I_{xy}|P)$ y $p(I_{xy}|L)$.

El hecho de que los píxeles pertenecientes a la clase pavimento sean tan numerosos será aprovechado para realizar el modelado de $p(I_{xy}|P)$, además del hecho de que estos valores son más bajos que los de las líneas divisorias y se presentan con una varianza baja al ser píxeles de una región de la imagen uniforme. Los píxeles de las líneas divisorias, en cambio, tienen valores de luminancia mucho más elevados y aparecen concentrados en pequeños bloques y a ráfagas.

Hay que resaltar el hecho de que para evitar que aparezcan vehículos al modelar la clase pavimento, es conveniente recortar la imagen I_{ROI} . Limitando la zona de análisis,

debido a que en la mayor parte de las situaciones en conducción en carretera se mantiene una distancia mínima de seguridad, se suele reducir la presencia de vehículos, así como de elementos laterales tales como las medianas, guardarrañes, etc. La reducción en un porcentaje de la altura de I_{ROI} (en nuestro caso un 25%), que se aplica también lateralmente. El resultado es una región que denominamos I'_{ROI} , que aparece recuadrada en amarillo en la Figura 4.3.

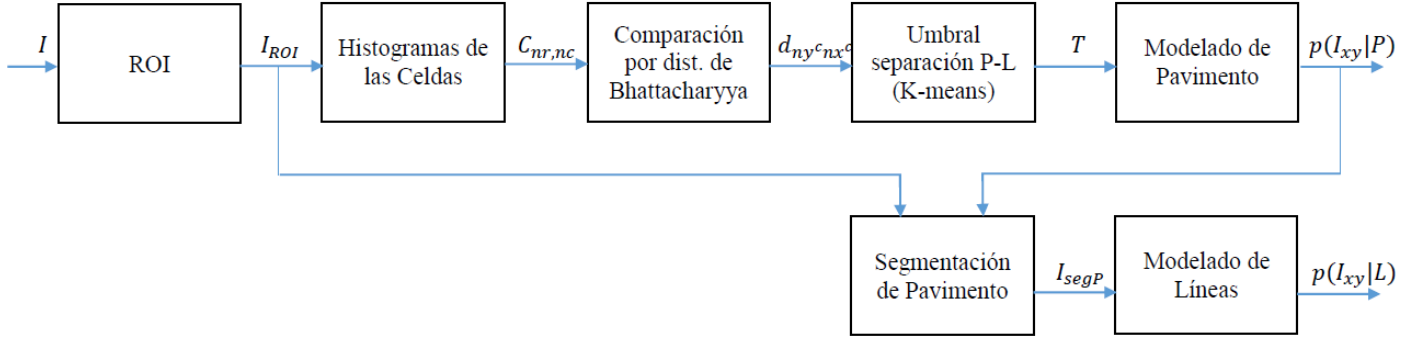


Figura 4.4 Diagrama funcional del modelado basado en luminancia

La dificultad es por lo tanto alimentar el conjunto de los píxeles de la clase pavimento sin que entren píxeles pertenecientes a la clase de las líneas divisorias. Para ello, se divide en celdas la imagen I'_{ROI} y se calculan los histogramas de estos, guardándose en el conjunto de celdas denominado $c_{nr,nc}$, donde nr indica la fila y nc la columna dentro del conjunto. Con el objetivo de que celdas con píxeles de líneas de carretera no entren en el conjunto de la clase P es necesaria alguna medida de comparación entre los histogramas contenidos en $c_{nr,nc}$. Con esta comparación realizada, se podrá saber qué celdas contienen íntegramente píxeles de la clase pavimento y cuáles no.

Las celdas en las que se divide I'_{ROI} deben ser rectangulares para obtener los mejores resultados, con mayor anchura que altura (en la práctica se han usado celdas de tamaño 8 filas por 18 columnas). La razón es que siempre ocurre que las líneas de carretera aparecen en las secuencias como ráfagas prácticamente verticales (no del todo, como sí ocurre con las imágenes IPM, *Inverse Perspective Mapping*), de anchura reducida. Si se usan celdas de una anchura mayor que la anchura de una línea, ningún bloque contendrá de ninguna manera píxeles pertenecientes exclusivamente a la clase L .

La comparación entre las celdas del conjunto $c_{nr,nc}$ sólo va a tener en cuenta celdas de una misma fila y que sean adyacentes, como se puede ver en la Figura 4.5. Las flechas indican los histogramas de las celdas que se comparan. Como medida de comparación se usará la distancia de Bhattacharyya [11], que tiene como entrada dos histogramas y devuelve un valor entre cero y uno, donde cero indica correspondencia perfecta y uno indica histogramas totalmente diferentes:

$$d_{Bhattacharyya}(H_1, H_2) = \sqrt{1 - \frac{\sum_{i=1}^N \sqrt{H_1(i) \cdot H_2(i)}}{\sqrt{\sum_{j=1}^N H_1(j) \cdot \sum_{j=1}^N H_2(j)}}} \quad (3)$$

Distancia de Bhattacharyya entre los histogramas H_1 y H_2

donde H_1 y H_2 representan los histogramas que se van a comparar, y N es el número de intervalos (*bins*) en los histogramas. Analizando la ecuación (3) se puede comprobar que la distancia de Bhattacharyya se basa en realizar una comparación *bin a bin* de los dos histogramas. Este método tiene una mayor carga computacional comparado con otros factores o distancias que se usan para comparar histogramas, como *Chi-Square*, el método de correlación o de intersección, si bien ofrece mejores resultados [11].

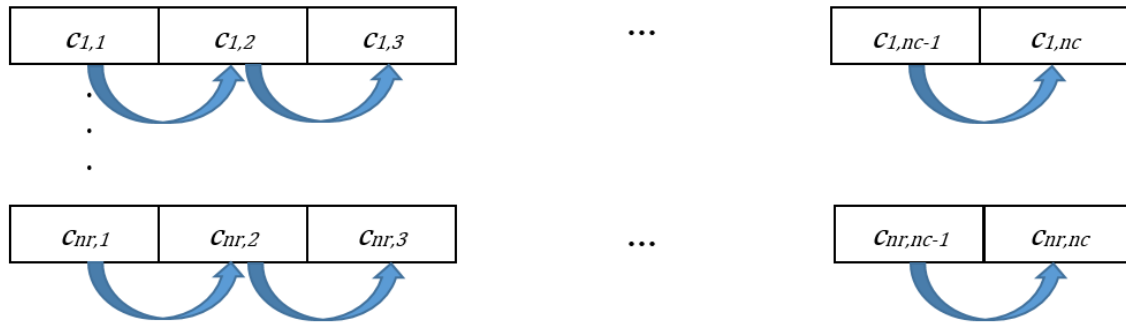


Figura 4.5 Conjunto de celdas $c_{nr,nc}$. Cada una de estas celdas contiene el histograma del bloque de I_{ROI} correspondiente. El sentido de la flecha indica los histogramas de las celdas que se compararán.

Al aplicar esta distancia entre las celdas, aparecerá un grupo de distancias más bajas, correspondientes a la comparación entre celdas con píxeles íntegramente de pavimento (debido a su esperada alta uniformidad) y otro grupo de distancias más altas correspondientes a la comparación de celdas que contienen una mezcla de píxeles de pavimento y líneas de carretera.

Si ocurriera que dos celdas con píxeles exclusivamente de líneas de carretera se compararan, el valor de la distancia sería también muy bajo, añadiendo esos píxeles erróneamente al conjunto de los de pavimento. Sin embargo, debido a que la anchura de la celda se va a seleccionar lo suficientemente grande como para que no contenga sólo píxeles pertenecientes a la clase L , el método será efectivo siempre y cuando se logre un umbral que separe las celdas que contengan sólo píxeles de pavimento de las que contengan píxeles que contengan una mezcla de píxeles de las clases línea y pavimento (bloques mixtos).

Inicialmente se estableció el umbral de distancia para separar los bloques de pavimento de los bloques mixtos, analizando un conjunto significativo de imágenes de diferentes secuencias segmentadas manualmente. Si bien los resultados obtenidos fueron razonablemente correctos, fijar el umbral a priori restaba flexibilidad y capacidad de adaptación a la estrategia. Por lo tanto, se ha decidido aplicar un método que separa de forma automática los grupos de píxeles de la calzada del resto, aplicando una técnica de agrupación conocida con el nombre de *K-Means*, puesto que claramente va a haber dos agrupaciones diferenciadas.

Se buscarán dos centroides, uno por cada agrupación (*cluster*) en el espacio de distancias, aunque el centroide correspondiente a los píxeles de celdas que contengan líneas divisorias representará una agrupación con puntos más bien dispersos, debido a la no uniformidad de las distancias entre celdas con píxeles de las clases pavimento y líneas.

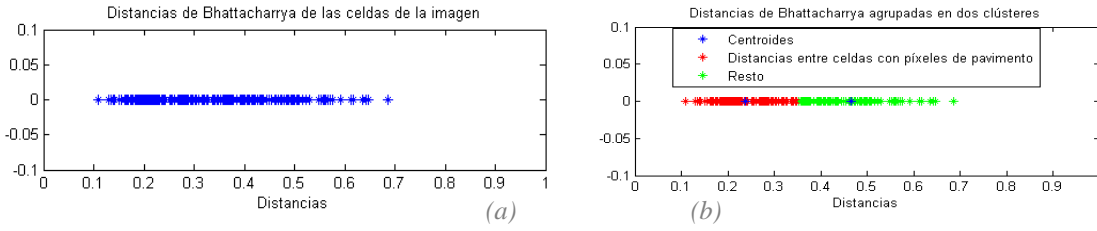


Figura 4.6 Distancias entre las celdas y distancias en dos agrupaciones

La Figura 4.6(a) muestra el conjunto de las distancias de Bhattacharyya obtenidas en la comparación entre las celdas y la Figura 4.6b muestra el resultado de la aplicación del algoritmo *K-Means*, quedando agrupadas en dos *clusters* representados por los centroides. Se puede observar cómo existe una agrupación clara de medidas bajas de distancias mientras que medidas más altas quedan dispersas a la derecha de la distribución.

El umbral que se obtiene de forma automática es la media entre los dos centroides. Este valor garantiza que la mayor parte de las celdas escogidas sirvan para alimentar el modelo de la clase *P*. Sin embargo, aunque este valor de umbral es satisfactorio, no garantiza que todas las celdas escogidas sean de píxeles que pertenezcan exclusivamente a la clase pavimento, por lo tanto, se ha optado por introducir un factor de restricción que el usuario puede ajustar. Este factor, que puede tomar valores entre cero y uno, escala el umbral. Si se introduce el valor uno, se establece como umbral el centroe más bajo (mayor restricción), consiguiendo un menor número de celdas para crear el modelo a cambio de obtener un mayor porcentaje de celdas de píxeles de la clase pavimento para el mismo. Si el parámetro se introduce con valor cero, el umbral se mantiene como se ha calculado de forma automática, es decir, la media entre ambos.

Finalmente, para crear el modelo tan sólo se tendrá que calcular μ_P y σ_P (media y desviación típica de la luminancia de los píxeles seleccionados como pavimento) y se conseguirá $p(I_{xy}|P) \propto G_{\mu_P, \sigma_P}$. Un ejemplo del modelo para la clase *P* se muestra en la Figura 4.9(a).

El siguiente paso en la aproximación es modelar las líneas de carretera, es decir, obtener $p(I_{xy}|L)$. Se necesita un conjunto de píxeles de líneas de carretera para obtener μ_L y σ_L , y para ello un primer paso será etiquetar los píxeles de I'_{ROI} que pertenezcan a la clase pavimento con valor de luminancia cero. A esta imagen se le llamará I_{segP} , que puede verse en la Figura 4.7(a). Efectivamente, los píxeles de pavimento han sido marcados a cero, como esperábamos, y tan sólo quedan las líneas divisorias.

Teóricamente, los píxeles que no se hayan etiquetado deberían pertenecer a las líneas de carretera. Sin embargo, en las transiciones entre pavimento y líneas habrá píxeles de nivel de luminancia intermedio que añadirían ruido al modelo. Por lo tanto, se aplicará una umbralización adaptativa tal y como se propone en [8] para descartar aquellos píxeles que no pertenezcan a la clase *L*. El resultado de hacerlo sobre I_{segP} se muestra en la Figura 4.7(b).

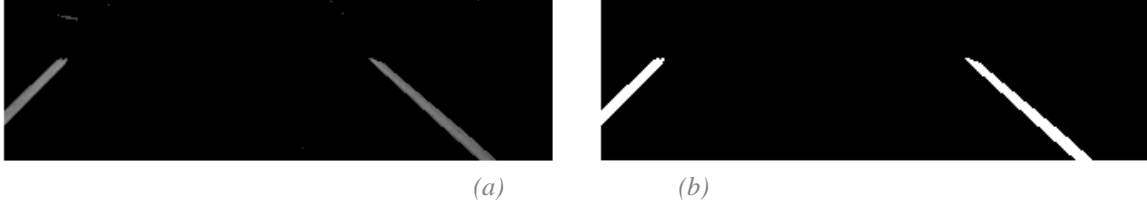


Figura 4.7 Imagen con el pavimento a nivel cero I_{segP} y la misma imagen con umbral adaptativo

El método de umbralización adaptativa píxel a píxel tiene en cuenta los niveles de luminancia. La asociación de un píxel de coordenadas (x, y) a la clase L dependerá de su nivel de luminancia y si este supera un determinado umbral t_{xy} :

$$t_{xy} = A \cdot e^{-\tau \cdot I_{xy}} \quad (4)$$

donde A es un factor de escala (255 para este caso por ser imágenes en escala de grises de 8 bits); $\tau = \ln 255 / 255$ e I_{xy} el valor de luminancia del píxel (x, y) . La Figura 4.8 ilustra la curva del umbral.

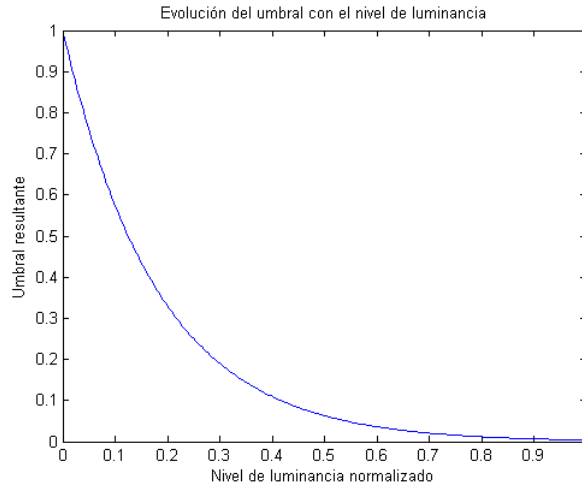


Figura 4.8 Curva del umbral adaptativo

Con este método, un nivel bajo de valor de luminancia del píxel aumentará el umbral para que el píxel sea seleccionado como píxel de la clase L , y un valor alto hará justo lo contrario. De esta forma se eliminan píxeles residuales con valor de luminancia más bajo que los valores que tienen las líneas, obteniendo una imagen binaria en la que los píxeles marcados con valor máximo de luminancia son los que se usarán para crear el modelo, es decir, se escogerá un subconjunto de píxeles con valor de luminancia mayor que $\mu_P + 3\sigma_P$.

Para el etiquetado de los píxeles con valor cero de luminancia, se usará el modelo obtenido del pavimento, es decir, su media μ_P y su desviación σ_P .

$$z_{xy} \in P \Leftrightarrow z_{xy} \in [0, \mu_P + 3\sigma_P], \quad \text{con } z_{xy} = I_{xy} \quad (5)$$

En la Figura 4.9(b) se puede observar el histograma de I'_{ROI} con el modelo de la clase P superpuesto. El conjunto de los píxeles que usarán para alimentar el modelo de la clase

L serán los que tengan un valor de luminancia por encima de los valores que delimite la curva del modelo P en la Figura.

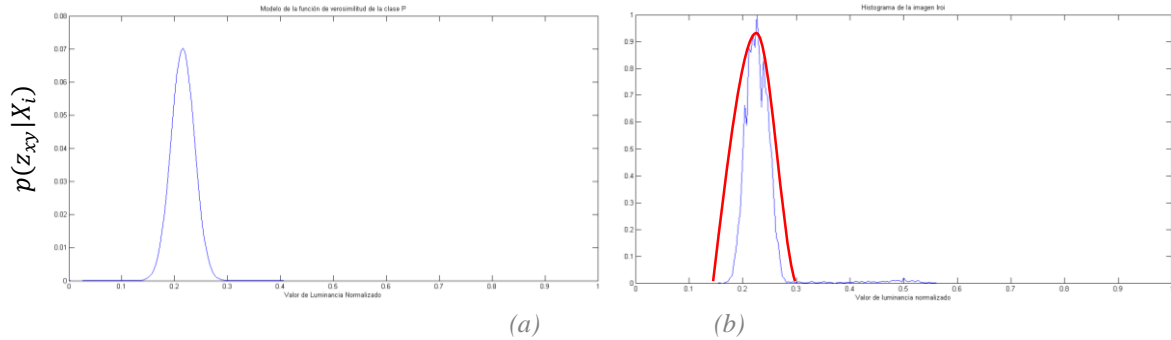


Figura 4.9 Modelo de la clase pavimento e histograma de I_{ROI} con el modelo superpuesto

El modelo $p(I_{xy}|L) \propto G_{\mu_L, \sigma_L}$ se determina calculando la media y desviación (μ_L y σ_L) del conjunto de píxeles seleccionados como perteneciente a la clase de las líneas de carretera.

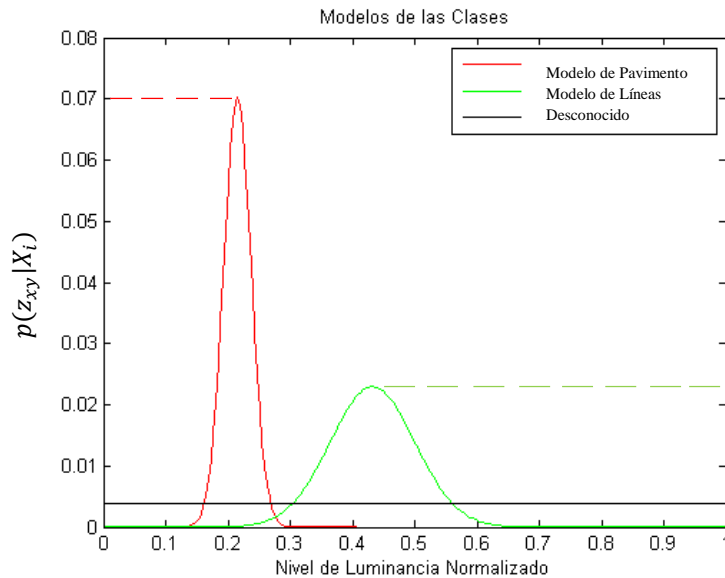


Figura 4.10 Modelos de las clases de la primera estrategia de segmentación (P, L, U)

En la 4.10 se muestran los tres modelos de las clases de la primera aproximación. Como se puede observar, se han modificado las distribuciones gaussianas para que en el caso de las de líneas, el máximo de la probabilidad sea asignado a partir del nivel de luminancia en el que ocurre. Esto tiene sentido, ya que nivel muy altos, más allá de donde se tiene la máxima probabilidad, lo más probable es que sea debido a que sean líneas de carretera, puesto que son los elementos con niveles más altos en I_{ROI} . De manera similar ocurre con el pavimento, ya que no se han tenido en cuenta las clases con niveles más oscuros que los de esta. Hasta el mínimo nivel lo más probable va a ser la clase pavimento.

La tercera clase considerada es la clase U (Desconocido). Modela aquellas situaciones en que la pertenencia a las clases P o L sea dudosa por los bajos valores obtenidos en sus probabilidades, y de esta forma no se corre el riesgo de asignar píxeles a clases con una probabilidad muy cercana a la probabilidad de pertenencia a otras clases.

El modelo de la clase U será una distribución gaussiana con muy alta varianza, semejante a una distribución uniforme de probabilidad de valor $1/MaxNivel$ de la imagen, puesto que a priori todos los píxeles de la imagen tienen la misma probabilidad de pertenencia a la clase U .

En la misma Figura 4.10 se puede comprobar también la utilidad del modelado de la clase desconocido. En torno al nivel 0.3 de luminancia, las probabilidades de pavimento y líneas decaen. En este punto en el que sería complicado clasificar los píxeles con esos niveles, se clasifican como desconocido.

4.1.1. Criterio de clasificación ML

Con todo el conjunto de clases $S = \{P, L, U\}$ modelado se puede proceder a la segmentación según esta primera estrategia. El resultado de la clasificación mediante el criterio de máxima verosimilitud para una de las imágenes de la secuencia puede verse en la Figura 4.11(a).

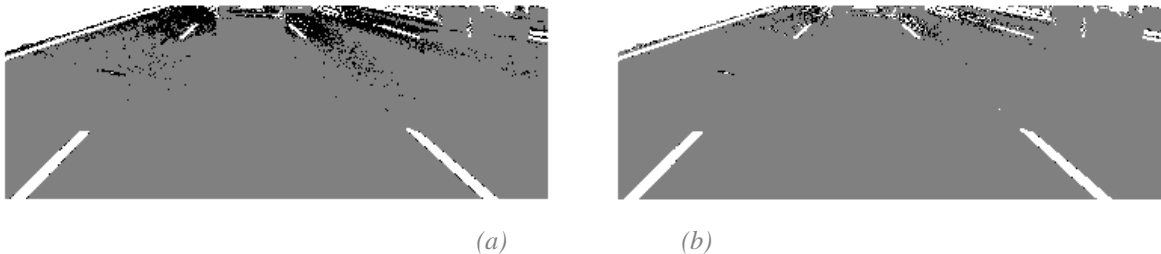


Figura 4.11 Imagen segmentada mediante la primera aproximación usando ML y mediante MAP

Esta Figura va a servirnos para realizar un análisis del funcionamiento de la primera aproximación planteada. El máximo valor de luminancia, en este caso 255, ha sido asignado a los píxeles clasificados como líneas de carretera, el valor intermedio 128 ha sido asignado para los píxeles de pavimento y el valor 0 para aquellos que, al no corresponderse con ninguna otra clase, pertenecen a la clase desconocido. En líneas generales, las líneas se clasifican correctamente a excepción de los bordes, donde la probabilidad de pertenencia a la clase P y L es muy similar. En cuanto al pavimento, aunque será normal que algún píxel de esta clase sea clasificado como perteneciente a la clase desconocido, con esta estrategia existen demasiados píxeles con valor cero. Se puede comprobar la eficacia del modelo en la zona de la imagen de donde se han extraído los píxeles pertenecientes a las clases P y L , si bien pierde robustez según nos alejamos de ella (parte superior de la imagen).

4.1.2. Criterio de clasificación MAP

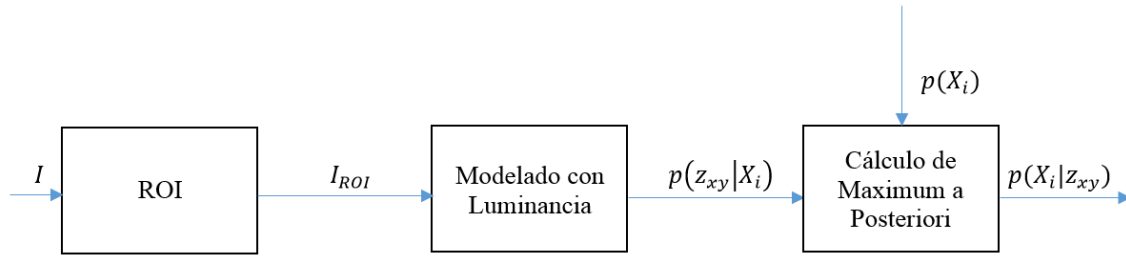


Figura 4.12 Primera estrategia de segmentación con la inclusión de información a priori

En la Figura 4.12 se puede ver el diagrama de flujo de la primera estrategia, utilizando el criterio MAP. Este método permite la consideración de información a priori, $p(X_i)$, que se va a obtener de las segmentaciones de imágenes anteriores, guardando para ello el porcentaje de píxeles que se han asignado a cada una de las clases posibles. Esta información se puede usar debido a que de una imagen a otra, se puede suponer que la escena cambia poco y que la proporción de píxeles asignados a cada clase de esta nueva imagen es el mismo que las proporciones de segmentación anterior.

$$p(X_i) = \frac{N_{X_i}}{N_T} \quad (6)$$

donde N_{X_i} representa el número de píxeles de I_{ROI} etiquetados como pertenecientes a la clase X_i ; y N_T es el número total de píxeles.

En la Figura 4.11(b) se presenta el resultado de la segmentación del mismo fotograma de la secuencia, con la información a priori de la segmentación del anterior. Se puede observar que los resultados son notablemente mejores puesto que se dan pesos a las probabilidades $p(z_{xy}|X_i)$ en función de su probabilidad de aparición. Existen muchos menos píxeles sin clasificar (pertenecientes a la clase U), extendiéndose la región donde la segmentación ofrece resultados correctos.

El método presentado permite obtener unas segmentaciones razonables, si bien presenta la limitación importante de no identificar los vehículos.

4.2. Segunda aproximación añadiendo la clase objetos

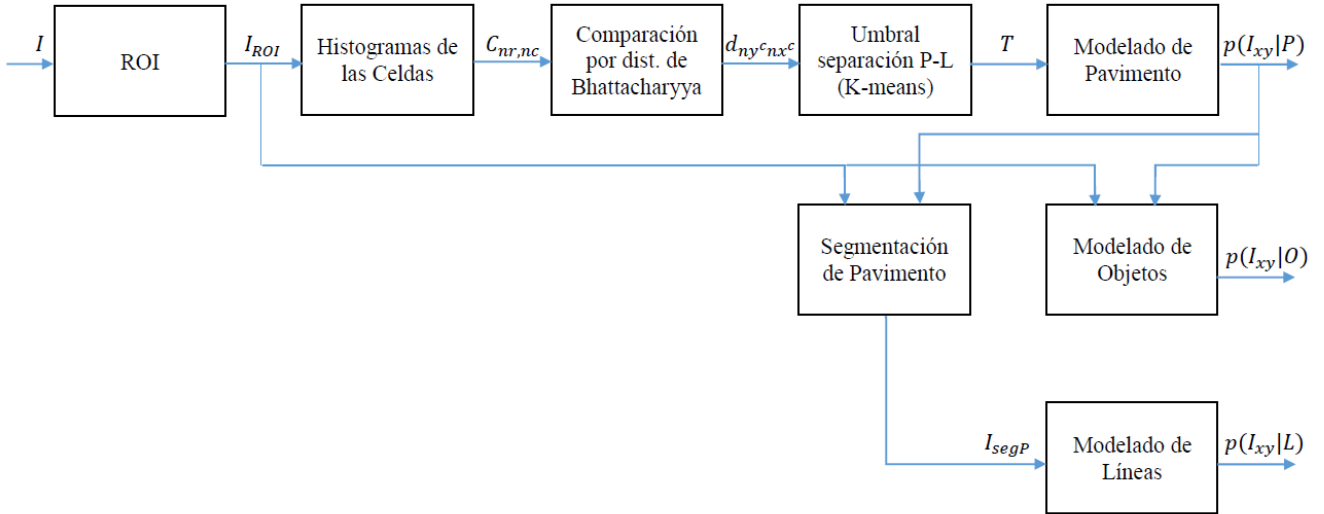


Figura 4.13 Diagrama de flujo de la segunda estrategia de segmentación

La Figura 4.13 presenta el diagrama funcional de la segunda estrategia de segmentación. Este es prácticamente idéntico al de la primera. Si bien en los bloques de modelado se ha añadido una nueva clase correspondiente a los objetos que aparezcan en la calzada, tales como vehículos.

Puesto que identificar los vehículos sería complicado en función de la luminancia debido a sus diferentes diseños y colores, se ha buscado una característica común a los vehículos que pueda aprovecharse. Esta es la sombra que aparece siempre debajo como se puede ver en la Figura 4.3 característica muy empleada en los trabajos relacionados [1], [9] y [12]. En cualquier situación, los vehículos proyectan esta sombra que oscurece notablemente el pavimento debajo de ellos.

Si se caracterizan los píxeles pertenecientes a esta sombra (incluidas las ruedas, que son también oscuras), se localizarán regiones donde probablemente haya vehículos. Para poder modelar la clase objeto (O), se empleará una estrategia similar a la utilizada para asignar píxeles a la clase L . Partiendo del modelo de la clase pavimento se asociarán a esta clase los píxeles que cumplan:

$$z_{xy} \in P \Leftrightarrow z_{xy} \in [\mu_P - 3\sigma_P, \mu_P + 3\sigma_P], \quad \text{con } z_{xy} = I_{xy} \quad (7)$$

Con ello, se garantiza que el 99.999% de los píxeles que podrían ser clasificados como pertenecientes al pavimento efectivamente sean así clasificados, pero con una diferencia con respecto a la anterior estrategia: no se incluyen aquellos píxeles con valores de luminancia muy por debajo del nivel medio de la luminancia del pavimento, correspondiente a las sombras bajo los vehículos.

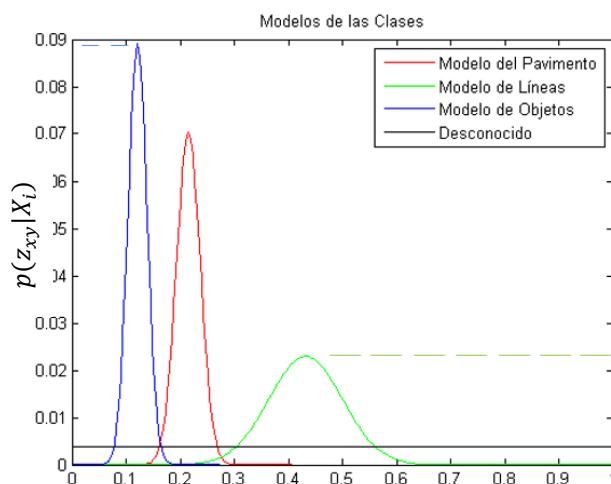


Figura 4.14 Modelos de la estrategia, con el nuevo modelo para objetos

Así pues, los modelos van a cambiar. Con los píxeles con valores de luminancia menores que $\mu_P - 3\sigma_P$ se alimentará el modelo de los objetos $p(z_{xy}|O) = p(I_{xy}|O)$. En la Figura 4.14 aparecen los nuevos modelos. Ahora es el modelo objetos en el que se mantiene hasta nivel cero el máximo del modelo. Con todos los modelos ya creados, se puede proceder a clasificar los píxeles.

4.2.1. Criterio de clasificación ML

El resultado de la segmentación de I_{ROI} (Figura 4.3), usando esta nueva estrategia con $S = \{P, L, O, U\}$ se puede ver en la Figura 4.15(a). Se puede comprobar cómo se han localizado las sombras y parte inferior de los vehículos. Las sombras se van a identificar como un gris oscuro, mientras que el pavimento se va a identificar con un gris claro. La clase L y la clase U se mantienen como en la primera estrategia, es decir, valor de luminancia 255 para L y valor de luminancia 0 para la clase U .

Esta estrategia mejora notablemente los resultados de la anterior por el hecho de albergar las cuatro clases reales que van a aparecer. Sin embargo este criterio sigue mostrando muchos píxeles asociados a la clase desconocido en la parte superior de la imagen.

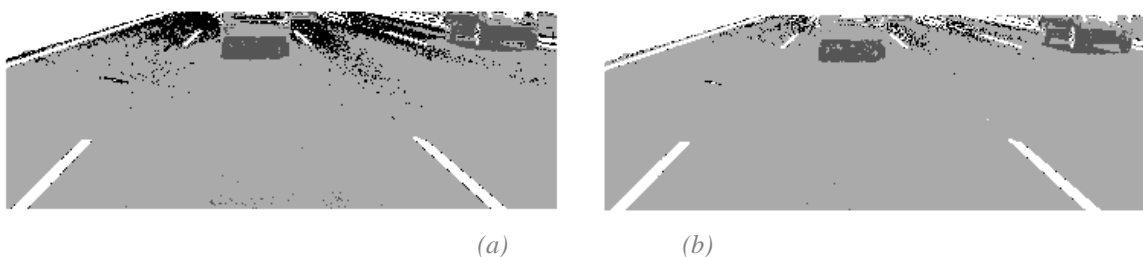


Figura 4.15 Imagen segmentada con el conjunto de clases $S = \{P, L, O, U\}$ clasificando con ML y MAP

4.2.2. Criterio de clasificación MAP

Se seguirá de nuevo el diagrama de la Figura 4.4, con la excepción de que ahora se modela una clase más, la de los objetos, por lo que el conjunto S se ajusta mejor que la primera aproximación a los tipos de elementos de interés de la carretera. Los resultados de segmentar una de las imágenes de la secuencia se puede ver en la Figura 4.15(b). En ella se comprueba que el número de píxeles asignado a la clase U se ha reducido con respecto al uso de máxima verosimilitud, manteniéndose un resultado coherente con la disposición real de los elementos de interés. Se puede también observar cómo se delinean mejor las partes inferiores de los vehículos, particularmente en el vehículo de la derecha, donde se separa muy bien la parte inferior y la sombra del pavimento, que empleando ML no se podía clasificar.

Hay que recordar que se está usando una estrategia que tan sólo usa los valores de luminancia de los píxeles para crear los modelos. En las siguientes estrategias, se añadirá otro tipo de información además de la luminancia que permita dar mayor precisión en la clasificación.

4.3. Tercera aproximación

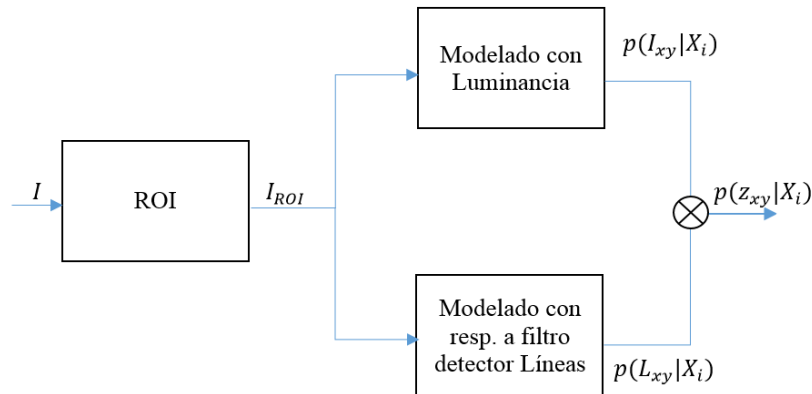


Figura 4.16 Diagrama de flujo de la tercera aproximación con modelado de respuesta a filtro detector de líneas

Hasta ahora se ha usado la luminancia como único atributo para modelar las características de cada una de las clases, esto es, $z_{xy} = I_{xy}$. En esta tercera aproximación, se va a introducir la respuesta (L_{xy}) a un filtro detector de líneas divisorias para modelar $p(L_{xy}|X_i)$. Como se puede ver en la Figura 4.16, se añadirá al modelado con luminancia explicado anteriormente. Asumimos que los modelos de respuesta a este filtro son independientes (probabilísticamente) a los modelos que utilizan los valores de luminancia. Es por ello que las probabilidades $p(I_{xy}|X_i)$ y $p(L_{xy}|X_i)$ pueden multiplicarse y el resultado se usará para clasificar mediante los criterios de ML o MAP. Esto es:

$$p(z_{xy}|X_i) = p(I_{xy}|X_i) \cdot p(L_{xy}|X_i) \quad (8)$$

siendo L_{xy} la respuesta del filtro en el píxel con coordenadas (x, y) , definido como [1]:

$$L_{xy} = 2I_{xy} - (I_{x-\tau,y} + I_{x+\tau,y}) - |I_{x-\tau,y} - I_{x+\tau,y}| \quad (9)$$

El filtro detector de líneas va a ser a efectos prácticos una máscara que se aplica sobre cada píxel de la región analizada. Esta máscara tiene una fila y anchura dependiente de un factor τ . Cuanto más cercano sea τ a la anchura real de las líneas, mejor responderá la máscara a las mismas, puesto que la respuesta será alta en toda la anchura de la línea.

El primer término multiplica por dos la luminancia del píxel con coordenadas (x, y) , mientras que el segundo resta la luminancia de los píxeles a una distancia τ en cada lado. De esta forma, si se aplica el filtro en un píxel de línea divisoria y τ es lo suficientemente grande como para que los píxeles con coordenadas $(x \pm \tau, y)$ queden fuera de esta, la respuesta será alta, mientras que si no se aplica en un píxel de línea la respuesta va a ser reducida. El tercer término provoca que si la diferencia entre los valores $I_{x-\tau,y}$ y $I_{x+\tau,y}$ es alta, la respuesta al filtro en ese punto se reduzca, puesto que si esto ocurre significa que las zonas a los lados de la supuesta línea no son similares y por lo tanto sea menos probable que no encontremos en un píxel de línea de carretera.

Para poder aplicar la máscara en todos los píxeles de I_{ROI} , se expande lateralmente $\tau/2$ columnas a cada lado. Se puede ver la respuesta del filtro en diferentes filas de la imagen en la Figura 4.17, y como se puede comprobar, justo en los píxeles perteneciente a la clase L la respuesta es mucho más alta que con respecto al resto de píxeles.

Considerando este atributo, se pueden modelar las siguientes clases: La clase de línea divisoria, que está caracterizada por una respuesta de valor alto al filtro y las clases pavimento y objetos, que tienen una respuesta de valor bajo.

Por lo tanto, para el modelado de $p(L_{xy}|X_i)$ nuestro conjunto de clases será $S = \{P, L, O, U\}$ de la misma forma que anteriormente. La peculiaridad de modelar estas clases con este atributo es que P y O van a tener el mismo modelo. Es decir que $p(L_{xy}|P) = p(L_{xy}|O)$ ya que en ambos casos se espera una respuesta baja del filtro. Por último, también habrá una clase U (*Unknown*) para aquellos píxeles cuya respuesta al filtro no pueda clasificarse como perteneciente a ninguna de las otras dos, de igual forma a como se ha hecho en aproximaciones anteriores, se modela como una distribución uniforme, en este caso con probabilidad constante $1/(MaxResp)$, con $MaxResp = 2 * MaxNivel$, siendo $MaxNivel = 255$ para imágenes es escala de grises de 8 bits.

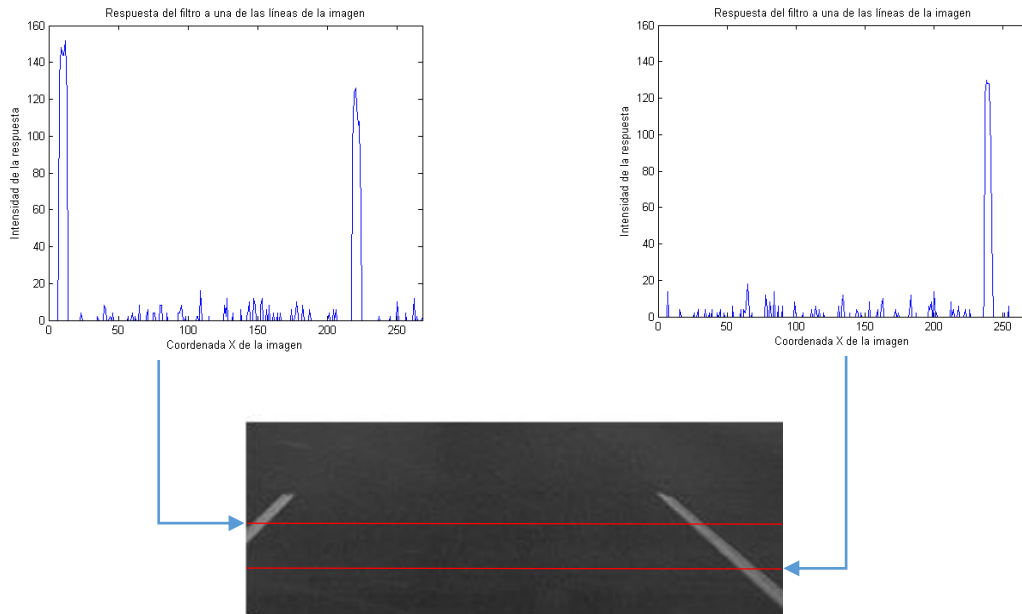


Figura 4.17 Respuesta del filtro en las filas marcadas en rojo. El filtro reacciona en píxeles de líneas divisorias

La Figura 4.18 muestra el histograma H_L de las respuestas al filtro detector de líneas. La frecuencia de las respuestas bajas al detector es muy abundante en comparación con las respuestas de valores altos. Esto es lógico ya que existe una proporción muy superior de píxeles que pertenecen a la clase P y O . Además, es un histograma monomodal similar a una gaussiana con media de valor prácticamente nulo y una desviación típica baja en

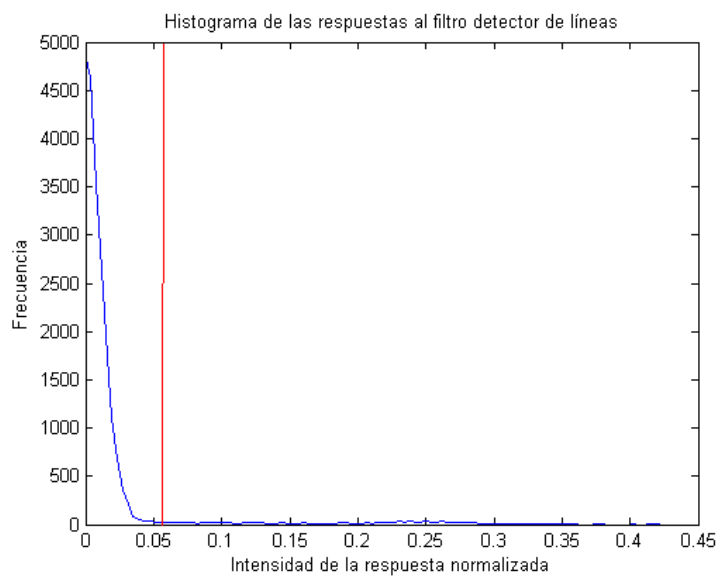


Figura 4.18 Histograma de la respuesta al filtro detector de líneas

comparación con los valores que alcanza la intensidad de la respuesta en los extremos. Esta desviación σ_{H_L} (en rojo en la Figura 4.18) va a ser la que haga de umbral ($T = \sigma_{H_L}$) a la hora de separar los valores que van a modelar $p(L_{xy}|L)$, $p(L_{xy}|P)$ y $p(L_{xy}|O)$, ya que hasta este valor, se concentran aquellas respuestas no pertenecientes a píxeles de la clase L [1].

Mientras que para las respuestas bajas existe una alta uniformidad en sus valores, se observa en la figura 4.18 que para las respuestas por encima del umbral, hay muy poca. Esto es así por la variabilidad en las respuestas de las líneas al filtro detector.

Con la media μ_L y la desviación σ_L del conjunto de respuestas mayores que T , se creará la gaussiana G_{μ_L, σ_L} equivalente a $p(L_{xy}|L)$. Las clases P y O usarán ambos el mismo modelo, con valor medio $\mu_P = \mu_O$ y desviación $\sigma_P = \sigma_O$, y se creará con los píxeles cuya respuesta al filtro quede por debajo de T .

Los modelos obtenidos se muestran en la Figura 4.19, además del modelo de la clase U . De nuevo, las distribuciones creadas, a excepción de la clase U se prolongan desde el valor máximo hasta el mínimo valor de respuesta para las clases O y P y el máximo valor de respuesta para la clase L . De nuevo, puesto que ya se dispone de los modelos usando el atributo L_{xy} como respuesta al filtro detector y de los modelos de la aproximación 2 usando el atributo I_{xy} , se procederá a la clasificación de los píxeles.

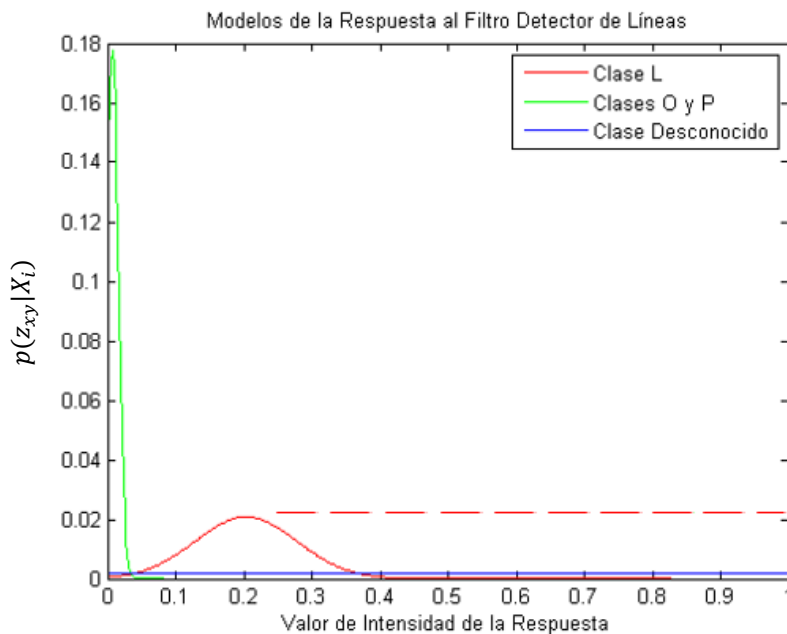


Figura 4.19 Modelos de las clases L , O y P para la respuesta al filtro detector de líneas

4.3.1. Criterio de clasificación ML

En la Figura 4.20(a) se muestra el resultado de la segmentación de la misma imagen utilizada en las aproximaciones anteriores.

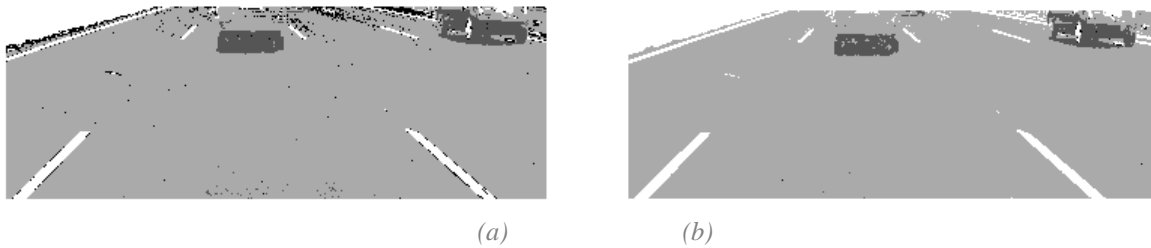


Figura 4.20 Clasificación de los píxeles de I_{ROI} según el criterio de ML y según el criterio MAP

Recordemos que se ha introducido la respuesta al filtro detector. Lo que provoca este filtro es una mayor heterogeneidad a la hora de clasificar, lo que en aquellos casos donde las probabilidades de aparición entre modelos de dos clases como por ejemplo P y L sea muy cercana en valor, una respuesta de alto valor al filtro detector va a permitir clasificar con mayor exactitud.

Con respecto al criterio ML de la anterior aproximación, los resultados han mejorado. Ahora existe un mayor número de píxeles correctos. Es cierto que también aparecen pequeños grupos de píxeles de la clase U en la parte inferior de la imagen pero en la parte superior se mejora el resultado notablemente.

4.3.2. Criterio de clasificación MAP

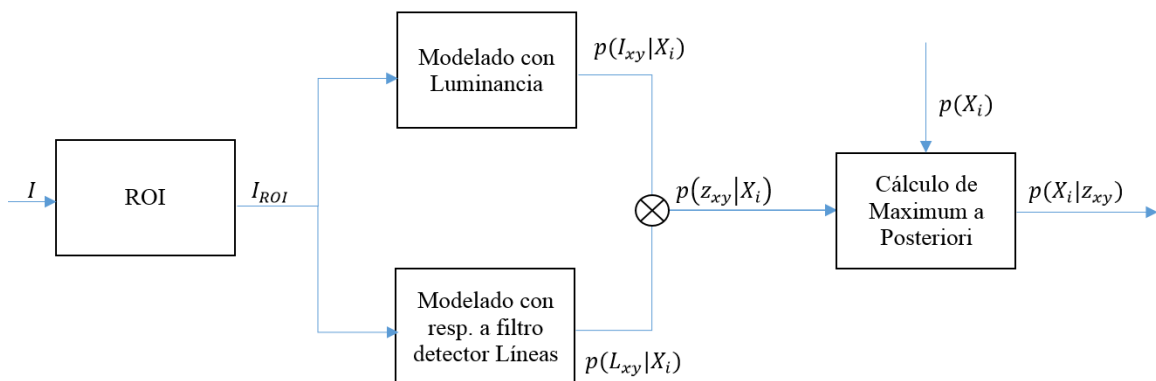


Figura 4.21 Diagrama de flujo utilizando los atributos L e I añadiendo información a priori

La Figura 4.21 muestra el diagrama de la Figura 4.16 que usa los atributos L_{xy} e I_{xy} adaptado para recibir información a priori y poder usar el criterio de clasificación MAP.

En la Figura 4.20(b) se muestra la imagen segmentada con el modelado de esta última aproximación y el criterio de clasificación MAP. Con este criterio de clasificación es con el que se obtienen los mejores resultados mejores que empleando ML y sobretodo mejor que empleando MAP con la aproximación que sólo emplea luminancia.

El número de píxeles pertenecientes a la clase U se ha reducido casi por completo. En la Figura 4.20a aparece cierto número de píxeles pertenecientes a esta clase, incluso en la zona inferior de la imagen, a pesar de que los modelos se han realizado con píxeles de esta zona. La introducción de la información a priori implica que el número de píxeles clasificados aumenta y las asignaciones de esta nueva clasificación son fundamentalmente correctas. Además se delimitan correctamente y con precisión los elementos de interés, consiguiendo una muy buena definición en la segmentación, como se puede ver por ejemplo en la parte baja del vehículo de la derecha de la figura 4.20(b).

La información a priori está formada por las probabilidades de aparición de cada una de las clases, y se basa en las proporciones que se han obtenido de píxeles clasificados en la anterior imagen. Si la anterior clasificación no ha sido correcta, la información a priori va a ser también errónea, lo que conduciría irremediablemente a que las siguientes clasificaciones también lo sean, provocando un mal funcionamiento del algoritmo. Por lo tanto, la adaptación dinámica de los modelos al entorno es fundamental para permitir una buena segmentación. Esto se tratará en la siguiente aproximación, con la que el sistema será capaz de detectar problemas en la adaptación de los modelos.

4.4. Cuarta aproximación. Actualización de los modelos

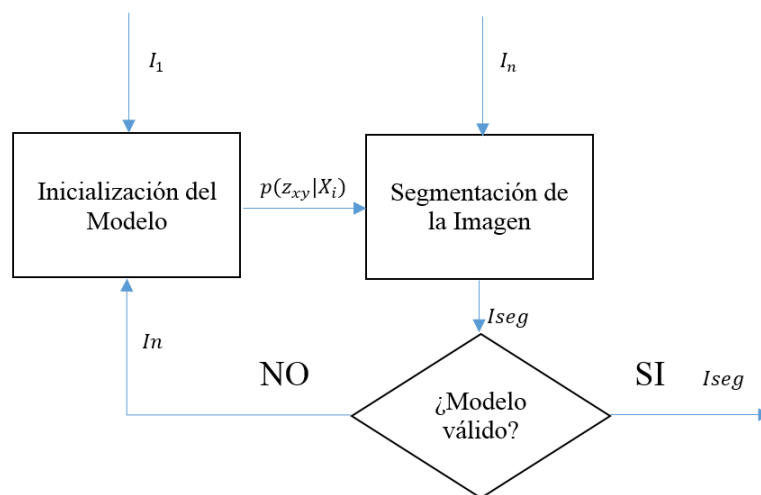


Figura 4.22 Estrategia en la actualización del modelo

Las últimas tres aproximaciones se centraban en la creación de los modelos, habiendo conseguido los mejores resultados en la última. En esta aproximación, por lo tanto, nuestro conjunto de clases es $S = \{P, L, O, U\}$ y los atributos de los píxeles que van a modelar son I_{xy} y L_{xy} (Luminancia y respuesta al filtro detector de líneas). Sin embargo, hay algo que hasta ahora no se ha tenido en cuenta: Cuándo actualizar los modelos.

Realmente, se pueden crear los modelos de las clases cada vez que entre una nueva imagen al sistema pero esto no se lleva a cabo en casos reales, ya que no es nada práctico. Recalcular los modelos en cada imagen de entrada aumentaría considerablemente el coste computacional en el sistema, que rápidamente se vería desbordado al poder entrar hasta 25 imágenes por segundo desde la cámara. Además, podríamos estar recalculando un modelo que prácticamente no haya cambiado en comparación con el anterior, con el que se podría realizar perfectamente la clasificación de los píxeles.

La Figura 4.22 ilustra la nueva aproximación propuesta. En el primer bloque del diagrama, se encuentra el bloque en el que se inicializarán los modelos de las diferentes clases. La primera imagen de las secuencias entrará siempre a este bloque, ya que se necesita crear los modelos iniciales. El siguiente bloque de segmentación se corresponde con el bloque denominado clasificador de la Figura 4.1, y trabaja como hasta ahora se venía haciendo. Se van usar los modelos del conjunto de clases para clasificar los píxeles de la imagen según el criterio ML o MAP.

La diferencia fundamental se encuentra en el bloque de decisión, que verificará cuándo un modelo es válido o no. A continuación se propondrán los criterios que usará este bloque, que permitirá, en función de las imágenes clasificadas reconocer cuando reinicializar los modelos de las clases.

4.4.1. Verificación de problemas en los modelos

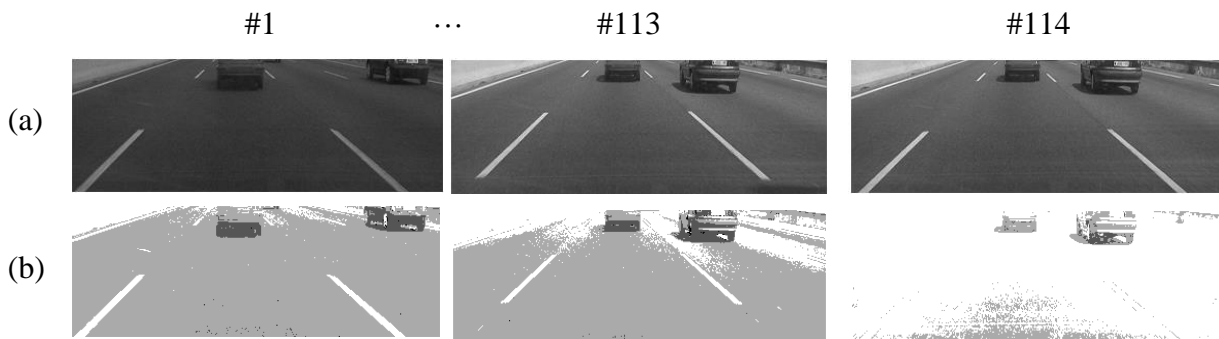


Figura 4.23 Imágenes de una misma secuencia y sus imágenes clasificadas (1)

En la Figura 4.23 se muestra para varias imágenes de una secuencia (a) el resultado de la clasificación (b) empleando siempre el modelo obtenido con la primera imagen #1(a). Como se puede observar, en la imagen #114(b) la mayoría de píxeles pertenecientes a la clase pavimento han sido erróneamente clasificados como pertenecientes a la clase de líneas divisorias. El problema en la actualización ha surgido por un cambio brusco en la iluminación que se produce justo en la imagen #113(a) que ha aumentado los valores de luminancia de los píxeles, provocando lo que ocurre en la imagen #114(b). También que la información a priori va cambiando erróneamente al tener más píxeles clasificados como *Lane Marking*. Por ello, además de actualizar los modelos cuando sea necesario, habrá que reiniciar las probabilidades a priori, es decir, que todas las clases sean igualmente probables.

Esto nos da una pista de qué características debemos verificar a la hora de buscar criterios que permitan saber cuándo actualizar los modelos y si se pueden utilizar las técnicas de modelado propuestas en todas las situaciones.

Como se ha dicho, los cambios de iluminación van a ser la causa principal de los errores en las clasificaciones de los píxeles, y por lo tanto de los modelos. Estos cambios de iluminación ocurren cuando se nubla o se despeja el cielo o cuando entramos y salimos de un túnel, entre otros. Se ha estudiado como cambian los histogramas debido a estos

cambios de iluminación. Se verificará cómo cambian los histogramas de las imágenes cuando aparecen cambios en la iluminación.

En la Figura 4.24(a) se muestra el histograma de una de las primeras I_{ROI} de la secuencia. Como era de esperar, tiene tres modos muy marcados. En general, el modo de luminancia más reducido se corresponde con las sombras de los vehículos, el del medio a los píxeles pertenecientes a la clase pavimento y el último a las líneas de carretera, aunque estos no serán las únicas clase de píxeles que aparecen en el histograma, ya que píxeles de otros elementos de la calzada se encuentran incluidos.

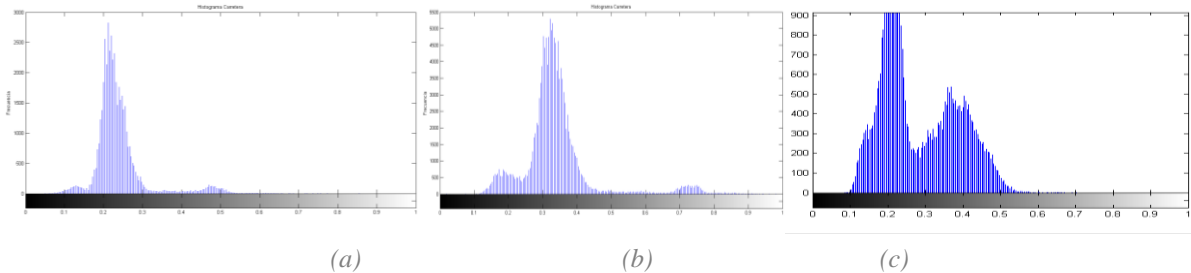


Figura 4.24 Histogramas de algunas imágenes de la secuencia con iluminación normal, con aumento de la iluminación y reducción de la iluminación

En la Figura 4.24(b) se encuentra el histograma de una imagen de la misma secuencia en la que ha habido un aumento significativo en la iluminación, que se corresponde con la Figura 4.23, imagen #1(a). En cambio, en la Figura 4.24(c), aparece el histograma de una imagen en una entrada a un túnel. El aumento significativo de la frecuencia del modo de luminancia más alto se corresponde con que hay flechas en la calzada, así como un bloque de hormigón en uno de los lados. Se puede ver cómo el modo correspondiente a la sombra de los vehículos prácticamente se ha fusionado en su totalidad con el modo perteneciente a los píxeles de pavimento.

Comparando los histogramas, vemos cómo hay un desplazamiento global a cada lado en función de si ha habido un aumento (desplazamiento hacia la derecha) o una disminución (desplazamiento a la izquierda) de la luminancia.

Los errores se producen al desplazarse los valores de luminancia de los píxeles, que acaban siendo asignados a las clases erróneas. El efecto de estos desplazamientos globales pueden controlarse actualizando los modelos, mientras se siga manteniendo diferenciación entre ellas. En el caso de situaciones críticas donde la iluminación sea excesivamente baja (en túneles por ejemplo) se producirán algunos errores en las clases cuyos niveles de intensidad se mezclen.

4.4.2. Criterios de actualización de los modelos

Actualización periódica de los modelos

El primer criterio para actualizar los modelos es un criterio de actualización periódica. Esto es, reinicializar los modelos cada cierto tiempo o cada cierto número de imágenes.

Existen dos problemas claros que van a surgir si utilizamos este método. El primero de ellos es que si se da el caso que nuestro modelo estaba funcionando correctamente, esto es, los píxeles se clasifican como pertenecientes a su clase, lo reinicializaremos cada cierto tiempo invirtiendo un coste computacional innecesario. El segundo problema claro, es que si entre el periodo de actualización de los modelos hay un cambio de iluminación como los explicados anteriormente, la imagen será clasificada erróneamente, asignando píxeles a clases que no son las suyas.

Por lo tanto, este método no es recomendable si se quieren obtener unos resultados satisfactorios.

Actualización en función de la clasificación anterior

Este criterio de actualización de los modelos se postula como el más indicado. Al igual que utilizamos la proporción de píxeles clasificados de la anterior imagen como información a priori para el criterio de clasificación MAP, podemos usar esta proporción para saber cuándo ha habido problemas en las clasificaciones, y por ende, en los modelos.

En las Tablas I, II y III aparecen las proporciones de píxeles clasificados a cada una de las clases con los criterios ML y MAP de tres de las secuencias de prueba. En media, se han obtenido unas proporciones de 91.03%, 6.08%, 2.78% y 0.0763% en las proporciones de píxeles para las clases *P*, *L*, *O* y *U* respectivamente. Las desviaciones típicas son de valores 0.93%, 0.71%, 0.4133% 0.027%. Para realizar los cálculos se han utilizado 50 imágenes clasificadas para cada secuencia y criterio de clasificación. Se han escogido estas secuencias debido a que en este intervalo de imágenes no hay cambios de iluminación significativos y la clasificación es visualmente correcta de forma que se puedan extraer datos fiables sobre la segmentación.

Tabla I Porcentaje de píxeles asignados a cada clase usando los criterios de clasificación ML y MAP. Secuencia 1

Secuencia 1	ML		MAP	
	% Medio y Desv. Típica		% Medio y Desv. Típica	
Pavimento	89.10%	1.49	91.25%	1.54
Líneas	8.08%	1.02	6.56%	1.03
Objetos	2.76%	0.91	2.17%	0.95
Desconocido	0.006%	0.003	0.001%	0.001

Tabla II Porcentaje de píxeles asignados a cada clase usando los criterios de clasificación ML y MAP. Secuencia 2

Secuencia 2	ML		MAP	
	% Medio y Desv. Típica		% Medio y Desv. Típica	
Pavimento	92.80%	0.35	93.62%	0.40
Líneas	5.69%	0.36	5.24%	0.38
Objetos	1.42%	0.07	1.14%	0.05
Desconocido	0.08	0.03	0.001%	0.009

Tabla III Porcentaje de píxeles asignados a cada clase usando los criterios de clasificación ML y MAP. Secuencia 3

Secuencia 3	ML		MAP	
	% Medio y Desv. Típica		% Medio y Desv. Típica	
Pavimento	89.03%	0.89	90.41%	0.91
Líneas	5.73%	0.73	5.22%	0.79
Objetos	4.90%	0.26	4.32%	0.24
Desconocido	0.33	0.08	0.04%	0.039

Estos datos nos indican que con diferencia la clase pavimentos es a la que más son asignados los píxeles, seguidas de las líneas de carretera, objetos y clase desconocido. Si nos fijamos, la clase objetos puede variar bastante de unas imágenes a otras dentro y fuera de una misma secuencia, ya que el número de vehículos cambia constantemente, mientras que los píxeles de las clases *L* y *P* tienen poca variabilidad dentro y fuera de la misma secuencia. De la clase Desconocido no se puede extraer demasiada información, tan sólo que al usar MAP, decrecen mucho su proporción en comparación con el criterio ML. Estos porcentajes serán muy útiles a la hora de buscar datos que evidencien problemas en la clasificación de los píxeles, puesto que si las proporciones cambian demasiado con respecto a los porcentajes medios podremos asumir que es necesaria una actualización de los modelos.

Las condiciones para actualizar los modelos se van a basar en que existan variaciones en las proporciones entre clasificaciones; las proporciones no entren dentro de un rango predeterminado.

1. La primera condición para actualizar los modelos será que los píxeles clasificados como pertenecientes a la clase *P* varíen un cierto porcentaje. Puesto que es con diferencia la clase con aparición más frecuente, si la proporción de estos cambia en exceso, será síntoma de que el modelo no está funcionando. Ya que la desviación típica de la proporción media de píxeles de pavimento es de un 1% aproximadamente, no se espera que entre dos fotogramas contiguos cambie mucho esta proporción. Por ello, se usará un 3% de forma prudencial, de forma que se evitan actualizaciones innecesarias y a la vez se verifica que no haya cambios bruscos en la proporción de píxeles de la clase pavimento.

2. La segunda condición de actualización tendrá en cuenta que la proporción de píxeles asignados a la clase U esté siempre dentro de un rango. Su media de proporción medida no alcanza tan siquiera un 0.01%. Sin embargo, este porcentaje en condiciones no ideales (en los que aparezcan otros elementos no asignables a ninguna de las clases en la carretera o haya cambios de iluminación) el porcentaje de píxeles asignados a la clase U puede llegar hasta cerca de un 1%. Por lo tanto, si se supera este 1% se lanzará la actualización de los modelos.
3. El tercer criterio tiene en cuenta que el porcentaje de píxeles pertenecientes a la clase L esté dentro de un rango. El mayor porcentaje de píxeles de la clase L en condiciones ideales se ha situado en un 8% (Tabla I), con una desviación de una unidad. Por lo tanto, no espera que el porcentaje de píxeles de la clase L sea mayor a un 10%, condición que si es sobrepasada lanzará la actualización de los modelos.

La información a priori se debe reiniciar (esto es, que la probabilidad de cada clase sea la misma) cuando se actualicen los modelos. De esta manera evitamos que las probabilidades a priori de una imagen clasificada en la que hay problemas, afecten a la hora de clasificar una imagen nueva.

La Figura 4.25 muestra la misma secuencia solo que ahora ha sido clasificada usando las condiciones de actualización de los modelos explicados. Podemos comprobar, cómo cuando en la imagen #113, al verificarse una de las 3 condiciones, se actualizan los modelos y se obtiene la imagen clasificada #114 de la secuencia (b), donde se obtienen mejores resultados razonables, muchos mejores que usando el modelo de la primera imagen.

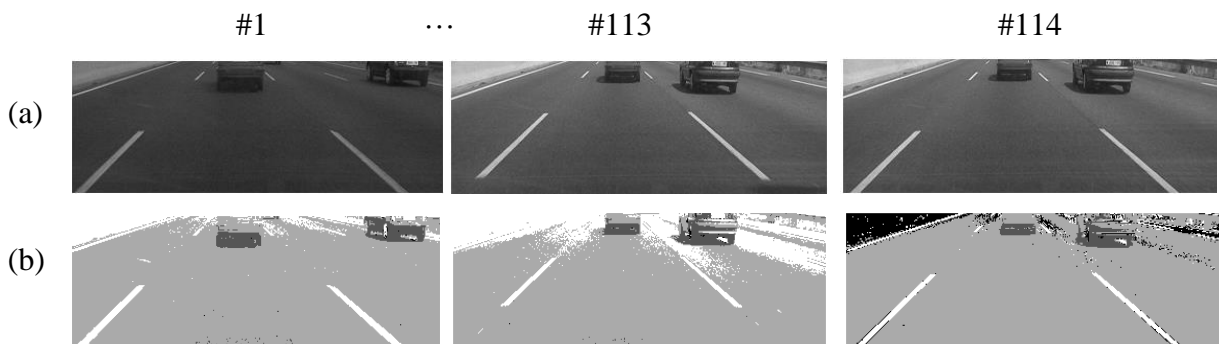


Figura 4.25 Imágenes de una misma secuencia y sus imágenes clasificadas (II)

Hasta este punto, hemos logrado usar dos atributos de los píxeles tales como la luminancia y la respuesta al filtro detector de líneas para crear los modelos de las 4 clases que hemos añadido, clasificar los píxeles por dos criterios diferentes y conseguir que el sistema se adapte automáticamente cuando detecte que ha habido un cambio de iluminación.

Llegados hasta aquí, y viendo que con la última aproximación se consiguen unos resultados razonables, esta se va a escoger como la mejor estrategia para segmentar las imágenes de carretera. En el siguiente apartado se procederá a realizar las pruebas que verifiquen el buen funcionamiento del sistema y evaluar los resultados.

5. Pruebas y resultados

Para verificar el correcto funcionamiento del sistema se van a desarrollar tanto pruebas subjetivas como objetivas. Las pruebas subjetivas se realizarán comparando visualmente la coherencia de las segmentaciones obtenidas con las imágenes originales. Para las pruebas objetivas, se parte de un conjunto de secuencias para las que se ha realizado una anotación manual (segmentación manual) en algunas de sus imágenes. Estas imágenes segmentadas se considerarán segmentaciones de referencia (*ground truth*) y en la evaluación se tendrán en cuenta los aciertos en la clasificación (píxeles cuya clasificación coincide con la de referencia) como aquellos incorrectamente clasificados.

Se han elegido tres secuencias con unas características tales que sus clasificaciones permiten realizar una evaluación relevante del sistema. La secuencia de pruebas 1 presenta cambios de iluminación que se producen gradualmente en varias imágenes. Además, no tiene la misma luminancia en la parte más inferior y la parte más superior dentro de una misma imagen, lo que pone a prueba al sistema debido a que el modelado se ha realizado con la parte más inferior.

La secuencia de pruebas 2, tiene varias características que la hacen apropiada para ser clasificada y evaluada. La primera de ellas es que en ella aparecen muchos vehículos, más que en la primera, por lo que permitirá analizar cómo actúa el sistema con tal cantidad de píxeles de la clase O . Además, durante un momento, un vehículo invade la distancia de seguridad, zona donde no debe haber ningún objeto para poder crear los modelos de las clases.

La secuencia de pruebas 3 tiene cambios de iluminación entre la parte más superior y la más inferior, como ocurre con la secuencia 1, pero a diferencia de esta, no presenta cambios importantes de iluminación gradual entre imágenes y por ello el sistema no actualizará los modelos en tantas ocasiones como se hará con las otras dos secuencias.

Se van a evaluar los resultados de dos de las aproximaciones realizadas junto con el criterio de clasificación MAP, puesto que ya se ha verificado en cada aproximación que ML ofrece peores resultados. Las pruebas evidenciarán pues, si introducir el filtrado de líneas de carretera aporta mejores resultados en la clasificación final.

Primera aproximación evaluada

La estrategia que se ha utilizado para realizar la primera evaluación es la cuarta aproximación detallada en este trabajo, que tiene las siguientes características:

- Atributos de los píxeles para modelado de las clases: Luminancia y respuesta al filtro detector de líneas de carretera.
- Criterio de clasificación MAP.
- Adaptación automática a cambios de iluminación.

Segunda aproximación evaluada

Los segundos resultados que se van a evaluar son los obtenidos con la aproximación 3 (con adaptación a cambios de iluminación) detallada en este trabajo, que tiene las siguientes características:

- Atributos de los píxeles para modelado de las clases: Sólo con luminancia.
- Criterio de clasificación MAP.
- Adaptación automática a cambios de iluminación.

En las Figuras 5.1, 5.2 y 5.3 se muestran algunas de las imágenes originales de las secuencias de pruebas (a), su clasificación con la aproximación 4 (b) y la aproximación 3 (c).

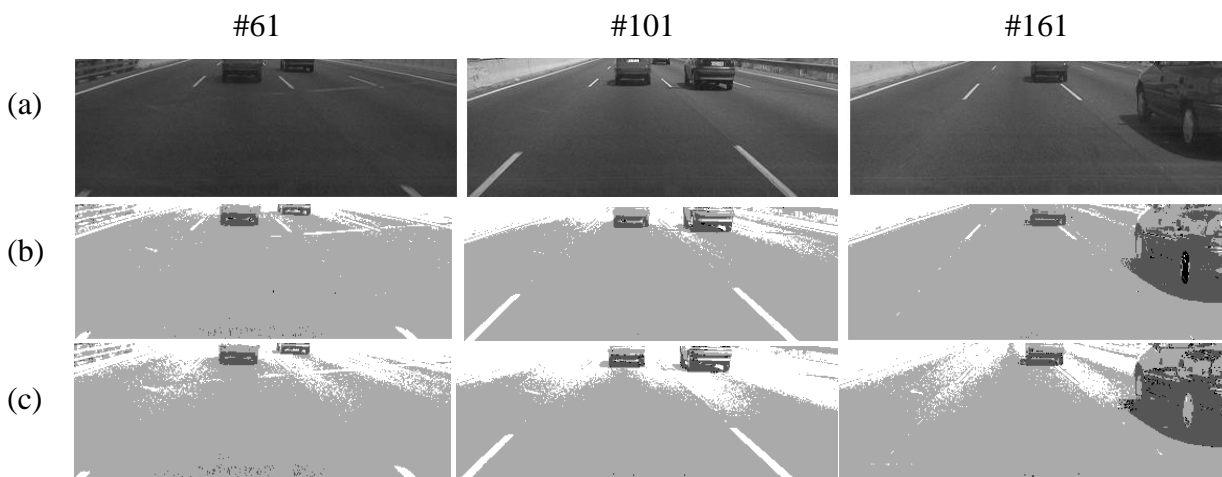


Figura 5.1 Fotogramas de la secuencia de pruebas 1 con imágenes originales y clasificadas con la aprox. 4 y 3



Figura 5.2 Fotogramas de la secuencia de pruebas 2 con imágenes originales y clasificadas con la aprox. 4 y 3

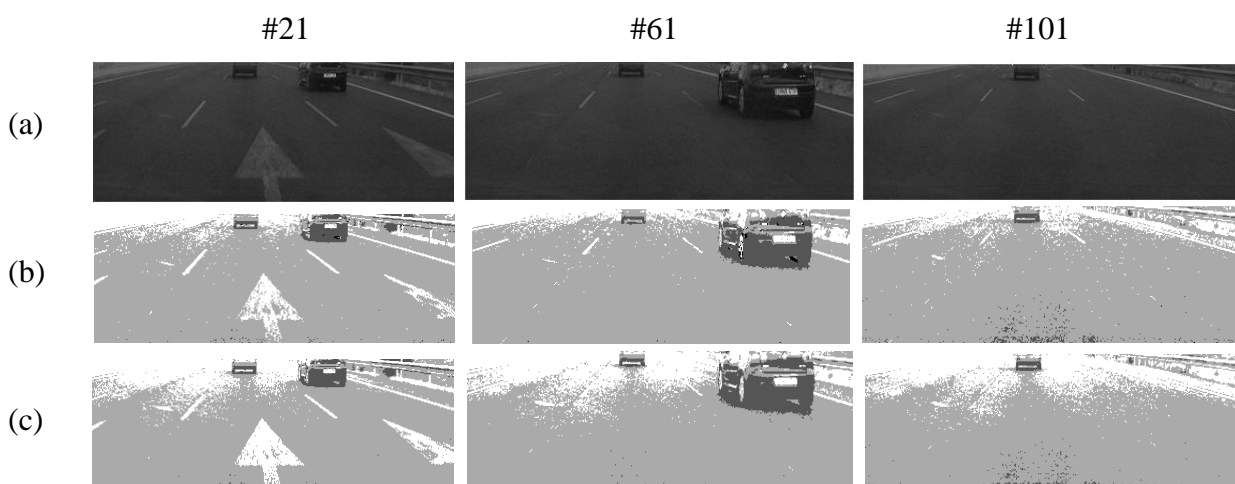


Figura 5.3 Fotogramas de la secuencia de pruebas 3 con imágenes originales y clasificadas con la aprox. 4 y 3

5.1. Pruebas subjetivas

En líneas generales se puede identificar sin problemas los diferentes elementos de interés, sobre todo los clasificados en las clases L y P . En cuanto a la clase O , sólo podemos localizar la sombra de los vehículos y por esta razón las partes superiores de estos se puede clasificar como píxeles pertenecientes a cualquiera de las demás clases. En cualquier caso, no va a ser un error, puesto que nuestro objetivo es localizar las sombras y las partes inferiores de los vehículos, cosa que se consigue con razonables resultados en prácticamente todas las clasificaciones de ambas estrategias. Los errores más comunes que se pueden apreciar, además del comentado, se encuentran siempre en las partes superiores de las imágenes, donde la mayoría de las veces, sus píxeles se clasifican como pertenecientes a la clase L o U . Esto ocurre con más frecuencia en las imágenes clasificadas con la aproximación 3, con la que son demasiados los píxeles que no se clasifican correctamente.

Una de las diferencias principales además de la comentada, se comprueba en la secuencia de pruebas 2. Si observamos detenidamente las sombras de los vehículos, vemos que existen partes de las mismas que con la aproximación 3 no se asignan a ninguna clase y sin embargo sí se hace y además correctamente con la estrategia 4, si bien es cierto que se pierde algo de compacidad como se dijo en capítulos anteriores. El último caso que es necesario valorar, es el de la clasificación de la imagen 181 de la secuencia de pruebas dos. Si nos fijamos, en este caso la aproximación 4 falla prácticamente por completo al realizarla debido a que en estas zonas la respuesta al filtro detector de líneas es baja y la clase más probable es la de pavimento, seguramente propiciado por la información a priori. Con la aproximación 3, los resultados no son mejores, porque, si bien no se clasifican todos esos píxeles como pertenecientes a la clase P , se dejan como pertenecientes a la clase U .

Un último detalle que conviene comentar, tiene que ver con la adaptación del sistema a cambios de iluminación. Cuando la clasificación comienza a ir mal en imágenes sucesivas, lo que se observan son pequeños saltos (sólo observables viendo la secuencia) debidas a la actualización de los modelos y la clasificación, que cambia las asignaciones de clases de los píxeles que en anteriores imágenes no se hacían correctamente, si bien esto no presenta ningún problema, puede resultar molesto, ya que a veces los cambios son algo bruscos.

5.2. Pruebas objetivas

Para las pruebas objetivas se han realizado un total de veinte clasificaciones manuales en total de las tres secuencias de prueba (siete de la primera secuencia, siete de la segunda y seis de la tercera). Las imágenes de las Figuras 5.1, 5.2 y 5.3 se encuentran entre estas clasificaciones manuales. El número de píxeles totales de cada I_{ROI} clasificada es de $128 * 360 = 46080$ píxeles.

Aproximación 4

Tabla IV Aciertos y errores de la secuencia de pruebas uno con la aproximación 4

Secuencia 1 Imagen/Clase	Aciertos en cada clase			Errores en cada clase		
	P	L	O	P	L	O
#61	40116	683	559	1705	31	135
#81	35694	697	662	5065	185	719
#101	37071	1276	544	2933	52	383
#121	38883	649	925	847	143	381
#141	36062	1068	1583	1764	449	254
#161	36101	321	2601	3217	8	327
#181	38149	719	2467	1238	325	700

% Medio de aciertos	93,53	86,78	72,17	% Medio de errores	6,47	13,22	27,83
---------------------	-------	-------	-------	--------------------	------	-------	-------

Tabla V Aciertos y errores de la secuencia de pruebas dos con la aproximación 4

Secuencia 2 Imagen/Clase	Aciertos en cada clase				Errores en cada clase		
	P	L	O		P	L	O
#1	37133	1346	2015		780	60	232
#21	37532	1120	1728		803	190	425
#41	37544	1027	1655		974	185	316
#61	35784	873	2583		2230	95	440
#81	34337	634	3560		1503	110	454
#101	33383	882	3754		2067	128	1043
#181	28137	296	28		617	174	6535
% Medio de aciertos	96,24	88,12	84,38	% Medio de errores	3,76	11,88	15,62

Tabla VI Aciertos y errores de la secuencia de pruebas tres con la aproximación 4

Secuencia 3 Imagen/Clase	Aciertos en cada clase				Errores en cada clase		
	P	L	O		P	L	O
#1	38091	904	849		4154	66	427
#21	35808	2279	933		2725	1877	161
#41	37959	1140	1252		3214	207	170
#61	36544	1054	1304		2958	352	414
#81	35148	771	3658		3085	96	430
#101	38478	513	162		5165	93	132
% Medio de aciertos	91,32	80,20	76,73	% Medio de errores	8,68	19,80	23,27

% Medio total de aciertos	93,70	85,03	77,76	% Medio total de errores	6,30	14,97	22,24
----------------------------------	--------------	--------------	--------------	---------------------------------	-------------	--------------	--------------

Aproximación 3

Tabla VII Aciertos y errores de la secuencia de pruebas uno con la aproximación 3

Secuencia 1 Imagen/Clase	Aciertos en cada clase				Errores en cada clase		
	P	L	O		P	L	O
#61	36051	703	580		5825	12	123
#81	29510	866	704		11252	16	693
#101	31225	1316	538		87777	12	379
#121	31750	7739	972		7739	43	338
#141	32954	1320	1566		4841	366	435
#161	31388	329	2536		7694	0	427
#181	33097	975	3177		6287	71	46
<i>% Medio de aciertos</i>	<i>72,11</i>	<i>95,56</i>	<i>71,60</i>	<i>% Medio de errores</i>	<i>27,89</i>	<i>4,44</i>	<i>28,40</i>

Tabla VIII Aciertos y errores de la secuencia de pruebas dos con la aproximación 3

Secuencia 2 Imagen/Clase	Aciertos en cada clase				Errores en cada clase		
	P	L	O		P	L	O
#1	34634	1388	1973		3321	25	172
#21	34676	1180	1678		3680	136	296
#41	33695	1127	1655		4895	93	216
#61	32134	2547	910		5846	58	361
#81	31156	695	3715		4489	51	274
#101	31806	862	3518		3515	170	653
#181	20075	820	14		16972	189	2092
<i>% Medio de aciertos</i>	<i>88,51</i>	<i>92,46</i>	<i>85,75</i>	<i>% Medio de errores</i>	<i>11,49</i>	<i>7,54</i>	<i>14,25</i>

Tabla IX Aciertos y errores de la secuencia de pruebas tres con la aproximación 3

Secuencia 3 Imagen/Clase	Aciertos en cada clase				Errores en cada clase		
	P	L	O		P	L	O
#1	35134	932	852		7114	38	433
#21	33856	621	938		8228	57	163
#41	34378	1190	1266		6812	157	193
#61	31692	452	2990		7685	87	628
#81	32463	828	3949		5747	39	154
#101	34086	554	162		9560	52	132
<i>% Medio de aciertos</i>	<i>81,77</i>	<i>91,13</i>	<i>78,71</i>	<i>% Medio de errores</i>	<i>18,23</i>	<i>8,87</i>	<i>21,29</i>

% Medio total de aciertos	80,79	93,05	78,69	% Medio total de errores	19,21	6,95	21,31
----------------------------------	--------------	--------------	--------------	---------------------------------	--------------	-------------	--------------

5.3. Análisis de las pruebas

La estrategia número 4 arroja buenos porcentajes en general, que como vemos superan el 90% de aciertos para la clase *P* y ronda el 80% para las clases *P* y *O* en media, aunque es conveniente analizar algunas apreciaciones. La secuencia uno tiene el porcentaje de error más alto para la clase líneas entre las tres secuencias. Esto es debido a que en la parte superior de las imágenes de esta secuencia se concentran más errores que en el resto de zonas debido a que, como dijimos anteriormente, se crea el modelo con píxeles de la zona más inferior. Los porcentajes de las demás clases, reflejan que las clasificaciones se realizan correctamente en la mayoría de las ocasiones, a excepción de los errores que surjan en la parte superior.

En la secuencia 2 hay un mayor porcentaje de errores en la clase *L*. Esto es debido a que al situarse un mayor número de vehículos en las zonas de la imagen, en muchas ocasiones se asignan píxeles de la parte superior del vehículo a clases como pavimento o líneas. Además, en esta secuencia hay un fotograma en particular (#181) que se interpone en la zona con las que se modelan las clases y en esos instantes la clasificación para la clase *O* roza el 100% de errores, aunque puesto que en general los resultados tienen un porcentaje alto de aciertos, prácticamente no afecta al resultado final.

La secuencia 3 tiene en media más errores que el resto. Esto es producido por la gran variabilidad de la luminancia entre la zona superior e inferior de la imagen, como se dijo antes. Es en esta zona es donde en mayor número de ocasiones se encuentran los vehículos de esta secuencia, y puesto que la zona superior es mucho más luminosa que la inferior, los errores en las clasificaciones de la clase *L* se hacen patentes.

En cuanto a los resultados ofrecidos por la aproximación 3, vemos que reflejan lo que se comentaba antes en las pruebas objetivas, en cuanto que hay un mayor número de píxeles que no se clasifica correctamente. Ahora la tasa de aciertos para la clase *P* es de un 80.79%, que puede parecer no tener demasiada diferencia, sin embargo, ya que la clase más frecuente es esta, estamos hablando del orden de miles de píxeles erróneamente clasificados. Las líneas y objetos de clasificación de forma correcta, incluso mejor que anteriormente. Esto es debido a que en la aproximación 4, puede ocurrir que algunas zonas que pertenezcan a la clase *L* se clasifiquen como de otra clase si el valor de la respuesta al filtro detector no es demasiado alto en ese punto. Por lo tanto, la última aproximación explicada en el trabajo, es la mejor opción, es decir, es mejor introducir el filtrado de líneas como atributo de los píxeles.

6. Conclusiones y Trabajo futuro

Después de probar varias estrategias de modelado y segmentación se ha conseguido obtener unos resultados más que razonables. El uso de las distintas clases, unidas al uso de los criterios de clasificación ML y MAP con el uso de dos atributos diferentes de los píxeles de las imágenes han permitido comparar y valorar el mejor método a la hora de conseguir las imágenes con sus píxeles clasificados. Los resultados lo avalan, sin embargo, son varias las características que conviene mejorar con el fin de aumentar con el fin de aumentar los porcentajes de aciertos en las clasificaciones.

Una de ellas afecta a las probabilidades a priori. Al introducirlas, estamos dando por hecho que se distribuyen uniformemente, y que un píxel de la parte superior de la imagen tiene la misma probabilidad de pertenecer a una clase que si el píxel estuviera en la parte inferior.

Otra de las características que se pueden mejorar, es el método de actualización de los modelos para que el sistema se adapte a cambios de iluminación. El método implementado no presenta cambios suaves en los modelos por lo que hay pequeños saltos en las clasificaciones de una imagen a otra.

La última característica que conviene mejorar es la detección de vehículos. Con el método implementado se detectan los vehículos casi la totalidad de los casos, aunque no se detecten todos los píxeles de la sombra y parte baja de los vehículos, pero no se detecta toda la zona del mismo.

Es por ello que se hacen las siguientes propuestas de trabajo futuro:

- Utilizar probabilidades de las clases no uniforme en función de la zona de la imagen. Para ser necesario estudiar qué clases son más probables en cada parte, para conseguir unos mejores resultados.
- Adaptación imagen a imagen de los modelos. De una imagen a otras, los modelos de las distintas clases no cambian en exceso, por ello recalcularlos desde cero no es una solución óptima. Se propone implementar técnicas de adaptación usando para ello EM (*Expectation-Maximization*) como se hace en [1], de forma que se vayan actualizando los modelos poco a poco y no existan cambios demasiado bruscos entre dos imágenes consecutivas.
- Para mejorar la detección de los vehículos se propone usar otros atributos de los mismos además de la sombra y la parte baja. Existen algunos como la simetría, tamaño que pueden mejorar los resultados, así como buscar los objetos de la imagen que estén en movimiento. También se puede implementar algunos de los métodos como [15] y [16] que usan descriptores y filtrado específicos.

Referencias

- [1] Marcos Nieto, Jon Arróspide Laborda, Luis Salgado. “*Road environment modeling using robust perspective analysis and recursive Bayesian segmentation*”. Machine Vision and Applications (2011), Volume 22, Pages 927–945.
- [2] Nieto, M.; Arrospide, J.; Salgado, L.; Jaureguizar, F. “*On-board video based system for robust road modeling*”, International Workshop Publication Year: 2008, Page(s): 109-116
- [3] Sayanan Sivaraman, Mohan Manubhai Trivedi. “*Vehicle Detection by Independent Parts for Urban Driver Assistance*”, IEEE transactions on intelligent transportation systems, vol. 14, no. 4, December 2013, pages 1597-1608.
- [4] Jia He, Hui Rong, Jinfeng Gong, Wei Huang, “*A Lane Detection Method for Lane Departure Warning System*”, 2010 International Conference on Optoelectronics and Image Processing, 2010, pages 28-31.
- [5] P.M.Daigavane, Dr. P.R.Bajaj. “*Road Lane Detection with Improved Canny Edges Using Ant Colony Optimization*”, Third International Conference on Emerging Trends in Engineering and Technology, 2010, pages: 76-80.
- [6] Yong Chen, Mingyi He, Yifan Zhang. “*Robust Lane Detection Based on Gradient Direction*”, 6th IEEE Conference on Industrial Electronics and Applications, pages: 1547-1552.
- [7] Mirko Felisa and Paolo Zani. “*Robust monocular lane detection in urban environments*”, 2010 IEEE Intelligent Vehicles Symposium, pages: 591-576.
- [8] Marcos Nieto, Luis Salgado, Fernando Jaureguizar, Julián Cabrera. “*Stabilization of Inverse Perspective Mapping Images Based on Robust Vanishing Point Estimation*”. Proceedings of the 2007 IEEE Intelligent Vehicles Symposium, Istanbul, Turkey, June 13-15, 2007, pages 315-320.
- [9] Deaho Lee; Youngtae Park “*Robust vehicle detection based on shadow classification Pattern Recognition*”, 2006.18th International Conference, Volume: 3, Publication Year: 2006, Page(s): 1167 – 1170.
- [10] Ross, Sheldon M.”Introduction to probability and statistics for engineers and scientists” 4th Ed.
- [11] A. Kaehler G. Bradski. “*Computer Vision with the OpenCV Library. Learning OpenCV. O'Reilly*”, First edition, September 2008.
- [12] Izumi, T.; Kikawa, M.; Hase, S.; Nishimoto, M. “*Forward vehicle tracking based on shadow detection*”, SICE, 2007 Annual Conference, 2007, Page(s): 1449 – 1453.
- [12] Horra Navarro, Julián de la. ”*Estadística aplicada*”.

- [13] González, Rafael C. “*Digital image processing*”, 3rd ed.
- [14] Akihiro WATANABE, Takashi NAITO, Yoshiki NINOMIYA. “*Lane detection with roadside structure using on-board monocular camera*”. 2009 IEEE
- [15] Y.-M. Chan, S.-S. Huang, L.-C. Fu, P.-Y. Hsiao, M.-F. Lo. “*Vehicle detection and tracking under various lighting conditions using a particle filter*”, IET Intell. Transp. Syst., 2012, Vol. 6, Iss. 1, pp. 1–8.
- [16] S. Yang, J. Xu and M.H. Wang. “*Onboard vehicle detection and tracking using boosted Gabor descriptor and sparse representation*”, ELECTRONICS LETTERS 2nd August 2012 Vol. 48 No. 16.

