

UNIVERSIDAD DE ALCALÁ
ESCUELA POLITÉCNICA SUPERIOR

Departamento de Electrónica



Detección de fatiga en conductores mediante fusión
de sistemas ADAS

Tesis Doctoral

Autor

Iván García Daza

Supervisores

Dr. D. Luis Miguel Bergasa Pascual

Dr. D. Rafael Barea Navarro

2011

Agradecimientos

Sirvan estas líneas para expresar mi más sincero y profundo agradecimiento a todas aquellas personas que con su ayuda han colaborado en la realización de esta tesis doctoral, en especial a los doctores D. Luis Miguel Bergasa Pascual y D. Rafael Barea Navarro, directores de esta investigación por la motivación y el apoyo recibido a lo largo de estos años y sobre todo por los conocimientos que tan generosamente me han transferido.

Especial reconocimiento merece el interés mostrado por mi trabajo y las sugerencias recibidas de los profesores D. Miguel Ángel Sotelo Vázquez y D. Pedro Revenga de Toro, con los que me encuentro en deuda por el ánimo infundido y la confianza en mí depositada.

Quisiera hacer extensiva mi gratitud al equipo investigador de RobeSafe de la Universidad de Alcalá por su amistad, colaboración y apoyo técnico. A David, Jesús, Nacho y Valky por la distensión que me han regalado con los partidos de tenis. A Pablo y Almazán por los agradables momentos vividos en el laboratorio. A Noelia y Pedro que han alegrado los frecuentes viajes de trabajo. A Sergio, Sebastián, Carlos, Miguel, Álvaro, Fernando, Ángel, Oscar, Raúl, Eduardo y José Javier por su compañerismo y amistad.

También agradezco a todos los compañeros de los proyectos CABINTEC y MARTA, su colaboración en el suministro de los datos necesarios para la realización de la parte empírica de esta investigación, en especial a Beatriz Delgado, Matías Sevillano, Ariadna Bartra y Noelia Rodríguez por los buenos momentos pasados en San Sebastián y Barcelona.

No podría olvidarme de los compañeros suecos que hicieron tan fructíferos los meses que pase en Chalmers (Gotemburgo). Para todos ellos mi agradecimiento y cariño esperando volver a verles.

Un agradecimiento muy especial merecen la paciencia y la comprensión de mi familia y amigos.

A todos, muchas gracias.

Resumen

Se ha identificado la somnolencia como una de las causas más importante de accidentes de tráfico, ya que se encuentra implicada en el 20% de los mismos, por lo que existe un interés creciente en encontrar sistemas ADAS (*Advanced Driver Assistance Systems*) capaces de detectar el estado de fatiga del conductor para prevenir posibles accidentes. Esto hace que las investigaciones en este ámbito sean de vital importancia.

Para abordar el análisis de la somnolencia durante la conducción, distintos grupos de investigadores han trabajado en diversas técnicas, entre las que destaca la Visión por Computador dado que permite, mediante el uso de una tecnología relativamente barata, la monitorización del conductor de forma no intrusiva.

En esta tesis se propone una técnica, basada en el procesado de imágenes monoculares consistente en la detección, seguimiento y caracterización de la apertura de los ojos, que trabaja automáticamente con distintos usuarios y en condiciones de conducción real. A partir de esta información y de otras señales relativas a la conducción, se infiere la somnolencia del conductor.

Las señales relativas al conductor se han obtenido de una amplia colección de vídeos de la cara de diferentes conductores, en simulación y en condiciones reales, en estado normal y de privación de sueño. Además, se dispone de numerosas señales relativas a la conducción, como la posición del vehículo en el carril, el ángulo de giro del volante, el error del ángulo de guiñada, etc.

Para la evaluación de la somnolencia en el conductor es necesario generar una señal de referencia que indique el estado del mismo. Esta señal se obtiene, por tres expertos, del estudio del índice de Karolinska (KSS), de la información facial extraída de los vídeos del conductor y de las señales registradas en la conducción. Normalmente, la señal de referencia consta de dos niveles: vigilia y somnolencia, pero en esta tesis se ha ampliado a un tercero, el de fatiga, para intentar mejorar la clasificación del estado del conductor. Tras un análisis de esta opción se ha concluido que la mejor clasificación es la binaria. Esta metodología es nueva ya que fusiona la información KSS con la proporcionada por expertos.

Con los vídeos de la cara del conductor y técnicas de procesado de imágenes se ha evaluado la apertura de los ojos. Para la detección de la cara se ha empleado el algoritmo de detección por apariencia de Viola y Jones, y para la de los ojos se ha mejorado con técnicas de clustering y un filtro de Kalman como predictor. La medida de la apertura de los ojos se ha obtenido aplicando filtros adaptativos, integrales proyectivas y un modelo Gaussiano cuya desviación estándar coincide con la apertura, consiguiendo un sistema en tiempo real y robusto frente a cambios de iluminación. Conocida la apertura se calcula el Porcentaje de Ojo Cerrado (PERCLOS), que es uno de los indicadores más importantes en la detección de somnolencia.

De las señales relativas a la conducción, que son dependientes de cada conductor, se han obtenido indicadores calculados con ventanas temporales como son: desviación estándar (std) y error cuadrático medio (mse) de la posición del vehículo en el carril, fracción de tiempo que el vehículo está fuera del carril (Lanex), tiempo estimado en cruzar los límites del carril (TLC), desviación estándar y movimientos rápidos del ángulo de giro del volante, desviación estándar del error del ángulo de guiñada e indicadores genéricos de variabilidad (GVI). Para eliminar la

dependencia de los indicadores con cada uno de los conductores, se aplican algoritmos genéticos para optimizar una función objetivo que tiene en cuenta a la señal de referencia.

Para determinar el estado del conductor se han fusionado indicadores entre sí mediante una red neuronal Perceptrón Multicapa, en donde el número de neuronas de la capa oculta se ha ajustado con el cálculo de la curva ROC.

Se han expuesto resultados en condiciones de simulación y en condiciones reales. La clasificación se ha realizado con indicadores individuales y fusionándolos, presentando los resultados en diagramas de error, en diagramas de tela de araña y en tablas donde se muestra la tasa de aciertos, especificidad, sensibilidad y valor de la función objetivo.

Los resultados obtenidos sobre la detección de somnolencia demuestran que la utilización del PERCLOS es determinante para la estimación del estado del conductor y que su fusión con otros indicadores de conducción mejora su tasa de aciertos individual. Los indicadores relacionados con la conducción obtienen peores resultados que el PERCLOS porque sus señales no sólo están motivadas por la somnolencia sino también por factores relacionados con la trayectoria de la carretera. El error del ángulo de guiñada se ha evaluado sólo en simulación puesto que en condiciones reales es difícil de estimar y no se disponía de él. En simulación las dos mejores combinaciones son la fusión entre el PERCLOS y la desviación estándar del error del ángulo de guiñada y el PERCLOS y el indicador mse optimizado. En condiciones reales la mejor combinación es el PERCLOS y mse optimizado. Si se hubiera tenido el error del ángulo de guiñada en condiciones reales quizás hubiera dado el mejor porcentaje en la clasificación, superando inclusive la fusión del PERCLOS con el indicador mse optimizado en condiciones reales, por lo tanto, los resultados obtenidos avalan la metodología empleada que traslada los resultados del simulador naturalista a las condiciones reales, extrayendo conclusiones válidas para las condiciones reales con porcentajes de detección menores debido al mayor ruido asociado a las señales.

En simulación, se ha revelado como un buen indicador de somnolencia por su independencia con respecto a la trayectoria. La fusión entre las señales de PERCLOS y ángulo de guiñada ha sido la que ha alcanzado mayor tanto por ciento de detección.

En términos generales, los resultados obtenidos están en concordancia con otros importantes trabajos sobre detección de somnolencia [Sandberg, 2011], a excepción de la discusión sobre la importancia de la variable PERCLOS ya que, en esta tesis, se concluye que es el mejor indicador de somnolencia debido, probablemente, a que el sistema de visión desarrollado para su obtención es propietario y no comercial. Por otro lado, los resultados de detección obtenidos son mejores que en otros importantes trabajos referenciados del estado del arte [Friedrichs & Yang, 2010a, Caterpillar, 2008].

Abstract

Drowsiness has been identified as one of the most important causes of traffic accidents, as it is present in the 20% of them. Therefore, there is a growing interest in looking for ADAS (*Advanced Driver Assistance Systems*), capable of detecting driver's fatigue, to prevent potential accidents. So that, research in this area is vital.

In order to tackle drowsiness analysis while a user is driving, several worldwide research groups have been working on different techniques. Computer Vision techniques are prominent among them, since it allows, by means of relatively cheap technology, to monitor driver's state in a non intrusive way.

In this thesis a technique based on monocular image processing is introduced. This consists of detection, tracking and characterization of eye closure, able to deal with different users and real driving conditions. Using this information and others acquired from the car, the driver behaviour is inferred.

Driver-related signals have been obtained from a wide set of sequences, in which there are different people's faces, either simulation or real driving conditions, awake or sleepy users.

For the purpose of driver drowsiness evaluation is necessary to generate a ground truth, which can provide the real state of the user at a particular time. This signal is obtained by 3 experts, as a result of studying several parameters as: KSS *Karolinska Sleepiness Scale*, facial features obtained from the recorded sequences and registered signals from the driving process. This ground truth usually has 2 levels, awakesness and drowsiness, although in this Thesis it has been extended, including a middle level: fatigue, to precisely classify the user's state. This methodology is new because takes information of the KSS scale and some experts.

Once all the user's face sequences were obtained, the developed computer vision techniques to obtain eye closure have been tested. Face detection is based on Viola & Jones algorithm, which is appearance-based, and eye detection is improved using clustering techniques and Kalman filtering, as predictor. Eye closure is obtained applying, over the located eyes, adaptive filters, projective integration and Gaussian modelling. All these image treatment algorithms makes the system robust against illumination variation and different users, archieving real time operation. Once the closure is known, the parameter PERCLOS (*PERcentage of eye CLOSure*) is computed. This parameter is one of the most relevant in drowsiness detection.

Talking about driving signals, obtained from the car, which depends on the user, some signs filtered using temporal window have been obtained like: standard deviation (STD), root mean square error (MSE) of the vehicle position on the lane, number of lane excedances (Lanex), time to lane crossing (TLC), its standard deviation, fast changes on the steering wheel angle, yaw angle standard deviation, and generic variability indicators (GVI). In order to remove sign-user dependence, parameter optimization using genetic algorithms has been carried out, taking into account the ground truth.

In order to stablish the driver's state, some indicators have been merged using a multilayer perceptron neural network, in which the number of neurons of the hidden layer is set using the ROC curve.

Some results are shown in simulation and real driving conditions. The classification is per-

formed by individual signals as well as fusing them, presenting the results using diagram error for 2 and 3 variables, and tables where the recall rate, the specificity, the sensitivity and the objective function are shown.

The results related to drowsiness detection demonstrates that PERCLOS is a fundamental parameter for the estimation of driver's state, and merging it with other driving signals improves the overall recall rate. Indicators related to driving yield worse results than using PERCLOS because those signals are not only caused by drowsiness but also by real driving conditions that are difficult to estimate. Heading error has been tested only in simulation because in real conditions it has not been provided by the person in charge of driving signals. During the simulation, the best possible combinations are the following: the fusion of PERCLOS and the standard deviation of the heading angle; and PERCLOS and the optimized MSE indicator. During real conditions, the best possible combination is the last one. If heading angle had had taken into account during real operation, the recall rate would have been improved. Therefore, the obtained results guarantee the methodology used, and can be easily extrapolated from the realistic simulator to real driving operation. Conclusions obtained using this methodology are valid for real conditions even though the detection rate is lower since the input signals are noisier.

The results are in line with other important works about this subject [Sandberg, 2011] except in the consideration of the PERCLOS, the best signal for us maybe due to we use our own vision system and not a commercial one. On the other hand, results are better than other important works of the state of the art [Friedrichs & Yang, 2010a, Caterpillar, 2008].

Índice general

1. Introducción	1
1.1. Motivación	1
1.2. Factores que influyen en la fatiga	2
1.3. Síntomas de la fatiga y somnolencia en la conducción	3
1.3.1. Síntomas fisiológicos	3
1.3.2. Síntomas en el comportamiento del conductor	4
1.3.3. Síntomas en la forma de conducir	4
1.4. Validación de estudios de somnolencia	4
1.5. Objetivos generales de la tesis	5
1.6. Estructura y organización de la tesis	7
2. Estado del arte	9
2.1. Estado del arte del método	9
2.1.1. Medidas subjetivas del estado de somnolencia	9
2.1.2. Medidas fisiológicas	10
2.1.3. Medidas relacionadas con el conductor	10
2.1.4. Medidas relacionadas con la conducción	11
2.1.5. Métodos híbridos	12
2.2. Extracción de señales físicas del conductor con técnicas de visión artificial	13
2.2.1. Modelos basados en forma	13
2.2.2. Modelos basados en características	16
2.2.3. Modelos basados en apariencia	17
2.2.4. Modelos híbridos	19
2.2.5. Otros métodos	19
2.3. Sistemas de detección de somnolencia comerciales	21
2.4. Discusión	22
3. Generación de datos para la evaluación del sistema	25
3.1. Ensayos en simulación	26
3.1.1. Diseño de los ensayos	26
3.1.2. Escenario de conducción	27
3.1.3. Señales registradas	28
3.2. Ensayos en condiciones reales	29
3.2.1. Diseño de los ensayos	30
3.2.2. Escenario de conducción	30
3.2.3. Señales registradas	31
3.3. Resumen de ensayos realizados	31
3.4. Generación de la señal de referencia (“ground truth”)	31

4. Cálculo de parámetros visuales de la cara del conductor	35
4.1. Sistema de adquisición de imágenes	36
4.2. Detección de cara	37
4.3. Detección de ojos	40
4.4. Evaluación de la apertura del ojo	43
4.4.1. Preprocesado de imagen	45
4.4.2. Aplicación de filtros adaptativos	45
4.4.3. Evaluación de la apertura y posición del ojo en imágenes filtradas	46
4.5. Seguimiento de los ojos	48
4.5.1. Filtro de Kalman como predictor	49
4.5.1.1. Modelo de movimiento	50
4.5.1.2. Parámetros e inicialización del filtro de Kalman	51
4.6. Definición y cálculo práctico de parámetros visuales de interés	53
5. Método de medida y fusión de datos	57
5.1. Indicadores de fatiga en la conducción	57
5.1.1. Desviación estándar sobre múltiples indicadores	58
5.1.2. Error cuadrático medio de la posición lateral	58
5.1.3. Lanex	58
5.1.4. Tiempo estimado en cruzar la línea del carril (TLC)	60
5.1.5. Movimientos de volante	62
5.1.6. Movimientos rápidos de volante	62
5.1.7. Ángulo de guiñada (Heading error)	62
5.1.8. Indicadores genéricos de variabilidad (GVI)	62
5.2. Optimización de indicadores	63
5.2.1. Técnicas de optimización estocástica	63
5.2.1.1. Algoritmos genéticos	64
5.2.1.2. Función objetivo	66
5.2.1.3. Preparación de los datos	67
5.2.2. Parámetros de los indicadores que han sido optimizados	67
5.3. Fusión de indicadores, métodos de análisis	68
5.3.1. Técnicas de fusión con redes neuronales	69
6. Resultados experimentales	71
6.1. Evaluación de la caracterización de la apertura del ojo	71
6.2. Metodología de evaluación del funcionamiento del sistema	74
6.2.1. Evaluación de un clasificador binario	75
6.2.2. Evaluación de un clasificador terciario	77
6.3. Resultados en condiciones de simulación	78
6.3.1. Análisis individual de los indicadores	81
6.3.2. Análisis de los indicadores optimizados	85
6.3.3. Análisis de la fusión de indicadores	88
6.4. Resultados en condiciones reales	88
6.4.1. Análisis individuales de los indicadores	89
6.4.2. Análisis de los indicadores optimizados	92
6.4.3. Análisis de la fusión de indicadores	93
6.5. Simulación vs Condiciones reales	93

7. Conclusiones y trabajos futuros	95
7.1. Principales contribuciones	96
7.2. Trabajos futuros	98
7.3. Main Contributions	102
7.4. Further work	103
Bibliografía	105

Índice de figuras

1.1. Diagrama general para la detección de somnolencia	6
2.1. Modelo de forma deformable	16
3.1. Diseño de pruebas en simulación	27
3.2. Fotografía del simulador, cedida por CEIT	28
3.3. Características simulador	28
3.4. Trayectoria recorrido real	30
3.5. Sistema PERCLOS vehículo real	31
3.6. Señal ground truth	33
4.1. Arquitectura del desarrollo software	36
4.2. Caras de entrada	37
4.3. Detección de cara ante giros en ensayos reales	39
4.4. Detección de cara ante giros en ensayos de simulación	39
4.5. Características Haar	39
4.6. Arquitecturas del detector de Viola-Jones	40
4.7. Detección de ojos en secuencia de imágenes de exteriores	41
4.8. Detección de ojos en secuencia de imágenes de exteriores	42
4.9. Clustering de candidatos de ojos	42
4.10. Restricciones de posición de los ojos	43
4.11. Candidatos finales de ojos	43
4.12. Técnica de apertura del ojo evaluada por <i>Suzuki et al, 2006</i>	44
4.13. Diagrama general para evaluar la apertura del ojo	44
4.14. Preprocesado de las imágenes de entrada	45
4.15. Filtros adaptativos aplicados a un ojo girado	46
4.16. Filtros adaptativos aplicados a un ojo frontal	46
4.17. Cálculo de proyecciones integrales horizontales y verticales	47
4.18. Convolución entre gaussiana patrón y proyecciones integrales	47
4.19. Evolución del cierre de ojos en ensayo real	48
4.20. Evolución del cierre de ojos en ensayo de simulación	49
4.21. Máquina de estados aplicada al seguimiento de ojos	49
4.22. Ruido del sistema	52
4.23. Filtro de Kalman	52
4.24. Evolución temporal de la posición del ojo	53
4.25. Evolución del cierre de ojos	54
4.26. PERCLOS	54
4.27. Parpadeo	55
4.28. Parpadeos por minuto	55
5.1. Indicadores evaluados en esta tesis	58

5.2. Error cuadrático medio	59
5.3. Indicador Lanex	60
5.4. Configuración del carril	61
5.5. Señal de Time to Lane Crossing	61
5.6. Ángulo de guiñada	62
5.7. Indicadores optimizados	63
5.8. Sobreentrenamiento	64
5.9. Algoritmo genético	65
5.10. Matriz de confusión	66
5.11. Matrices de confusión	67
5.12. División de las muestras de entrada	68
5.13. Red neuronal	70
6.1. Evolución temporal del PERCLOS en ensayo de vigilia	72
6.2. Secuencia de imágenes de ensayo de simulación sin somnolencia	72
6.3. Evolución temporal del PERCLOS en ensayo de somnolencia	72
6.4. Secuencia de imágenes de ensayo de simulación con somnolencia	73
6.5. Secuencia de imágenes para ensayo real	73
6.6. Evolución temporal del PERCLOS en ensayo real	74
6.7. Caracterización en un movimiento de giro de cabeza	74
6.8. Caracterización tras el reflejo del sol	75
6.9. Caracterización tras la salida de un tunel	75
6.10. Curva ROC	76
6.11. Representación gráfica de la matriz de confusión	78
6.12. Porcentajes de clasificación del total de ensayos	78
6.13. Señal de referencia y evolución temporal del PERCLOS	79
6.14. Clasificación en tres estados por medio del PERCLOS	79
6.15. Clasificación en dos estados por medio del PERCLOS	80
6.16. Clasificación en tres estados con desviación estándar de la posición en el carril	80
6.17. Clasificación en dos estados con desviación estándar de la posición en el carril	81
6.18. Indicadores sin y con privación de sueño	81
6.19. PERCLOS y posición del vehículo en el carril, estados de vigilia y somnolencia	82
6.20. Indicadores en simulación	83
6.21. Barras de error para tres indicadores y tres estados	83
6.22. Representación en tela de araña para el indicador TLC inferior a 2 segundos	84
6.23. Barras de error para dos indicadores y dos estados	85
6.24. Representación temporal de indicadores optimizados	86
6.25. Barras de error para dos indicadores con dos estados	87
6.26. Porcentajes de clasificación del total de ensayos	89
6.27. Señal de referencia y evolución temporal de PERCLOS	89
6.28. PERCLOS con clasificación binaria	90
6.29. Porcentaje STD LP dos estados	90
6.30. Indicadores sin optimizar	91

Índice de tablas

2.1. Métodos y estudios de detección de somnolencia	14
2.2. Modelos de detección de ojos	23
3.1. Resumen ensayos conducción	26
3.2. Características de los ensayos realizados	32
3.3. Escala del índice de somnolencia KSS	33
4.1. Características del sensor	38
5.1. Parámetros optimizados de indicadores genéricos	68
6.1. Matriz de confusión binaria	75
6.2. Matriz de confusión de tres estados	77
6.3. Ejemplo de matriz de confusión de tres estados	77
6.4. Clasificación tres estados	84
6.5. Indicadores sin optimizar	85
6.6. Parámetros de indicadores optimizados	86
6.7. Parámetros de indicadores genéricos	86
6.8. Indicadores optimizados	87
6.9. Combinación de indicadores simulación	88
6.10. Indicadores sin optimizar	91
6.11. Parámetros de indicadores optimizados	92
6.12. Parámetros de indicadores optimizados genéricos	92
6.13. Indicadores optimizados	93
6.14. Combinación de indicadores	93

Capítulo 1

Introducción

1.1. Motivación

La inatención debida a la fatiga o somnolencia al volante es bien conocida como un factor de riesgo determinante en accidentes de tráfico. Muchos conductores combinan la privación de sueño con excesivas horas de conducción existiendo un grave conflicto entre las necesidades fisiológicas y las actividades sociales o profesionales del conductor. Por lo tanto, la clave para conseguir la prevención de accidentes es el entendimiento de los límites humanos entre la fatiga, la somnolencia y la privación de sueño. La fatiga y somnolencia en conductores es un factor muy importante a tener en cuenta por ocasionar muchos accidentes de tráfico. Por esta razón durante la última década el problema de la somnolencia en conductores ha recibido un incremento de atención por parte de la comunidad científica, las compañías de vehículos y en general por la sociedad.

En el año 2007, en la Unión Europea se produjeron 42.854 muertes como consecuencia de accidentes de tráfico ocasionados por la fatiga o somnolencia [Mahieu, 2009]. En el año 2006 la cifra fue de 44.400 muertos [UN-ECE, 2007]. En ese mismo año, tuvieron lugar 1.25 millones de accidentes de tráfico y más de 1.5 millones de personas se vieron involucradas con heridas de menor o mayor gravedad. Los efectos de la fatiga o somnolencia están presentes entre el 15 % y el 20 % de los accidentes y por este motivo existe un interés creciente en encontrar sistemas automáticos capaces de detectar el estado de fatiga del conductor. En un esfuerzo para reducir estas tragedias, la Comisión Europea estableció en 2003 el programa de acción de seguridad en carretera [Commission, 2003], cuyo objetivo era reducir a la mitad el número de víctimas en accidentes de tráfico para 2010. La Comisión Europea presentó los primeros resultados en junio de 2009 [Commission, 2009]. Este documento no sólo resume los logros alcanzados con respecto al plan de 2003, sino que también identifica las políticas futuras de transporte. En sus resultados se muestra una reducción del 12 % en los accidentes de tráfico con heridos desde 1991 hasta 2007, mientras que el número de muertes en accidentes de tráfico disminuyó en un 44 % en el mismo periodo. Por otro lado, la National Highway Traffic Safety Administration (NHTSA) estima, que aproximadamente el 25 % de los informes redactados por la policía después de un accidente, se relacionan con factores de inatención del conductor, siendo los factores más comunes los debidos a distracciones, fatiga o somnolencia [Ranney et al., 2001]. Un estudio realizado por la American Automobile Association Foundation for Traffic Safety (AAA FTS) [Foundation for Traffic Safety, 2010] determina que la inatención de un conductor puede clasificarse en cinco categorías: atento, distraído, mirando pero sin ver, dormido o fatigado y desconocido.

Existen trabajos muy importantes, como el realizado en [Klauer et al., 2006] que demuestra que la somnolencia durante el acto de conducción aumenta el riesgo de sufrir un accidente de cuatro a seis veces, comparándolo con un estado de vigilia. Otros como [Akerstedt et al., 2001,

Connor et al., 2002] confirman que el riesgo de sufrir un accidente durante la conducción nocturna es mayor que en la diurna debido al ritmo circadiano al que nos vemos sujetos las personas.

El fenómeno de la fatiga es diferente al de las distracciones. El término fatiga se refiere a una combinación de síntomas tales que deterioran las características de la conducción y producen un subjetivo sentimiento de somnolencia [Sub-Group, 2006]. Incluso con la investigación intensiva que ha sido realizada hasta el momento, el término de fatiga no tiene todavía una definición aceptada universalmente. Por esta razón, es complicado determinar un nivel de fatiga a cada uno de los accidentes producidos. Sin embargo, como se comentó anteriormente, la mayor parte de los estudios realizados concluyen que en el 25-30 % de los accidentes están involucrados aspectos de la fatiga o somnolencia [Fatigue & Accidents, 2001]. La European Transport Safety Council (ETSC) define la fatiga o somnolencia como *estado concerniente a la imposibilidad o estar poco dispuesto a continuar una actividad, generalmente porque la actividad ha sido realizada continuadamente un largo periodo de tiempo*. Desde el punto de vista del funcionamiento del organismo existen diferentes clases de fatiga: fatiga física local, fatiga física general, fatiga del sistema central nervioso y fatiga mental (no tener energía para hacer nada). La fatiga del sistema nervioso central y la fatiga mental son las más peligrosas, porque eventualmente pueden desembocar en un estado de somnolencia, incrementando la probabilidad de accidente. Por lo tanto, los estudios realizados por la ETSC han llevado a definir cuatro niveles de somnolencia basados en términos de comportamiento [Croo et al., 2001]: completamente despierto, moderadamente despierto, somnolencia severa y dormido.

Cuando un conductor alcanza el estado de fatiga y comienza a caer dormido, pueden observarse los siguientes síntomas: bostezos repetitivos, confusión, sensación deprimida e irritable, reacciones y respuestas lentas, ensoñación, dificultad en mantener los ojos abiertos y sensación de quemazón en los mismos, movimientos del volante perezosos, dificultad en mantener la concentración, bandeo del vehículo en la carretera o cambios de carril involuntarios, cabeceos en el volante, etc, tal y como concluyen muchos de los trabajos referenciados en la tesis. Estos síntomas se manifiestan de forma diferente en cada conductor y están asociados a grados de fatiga o somnolencia también diferentes para cada uno, por esta razón, hay que utilizar más de una evidencia para medir el nivel de fatiga. Estudios de la ETSC [Croo et al., 2001] muestran que el nivel de somnolencia (la última expresión de la fatiga) es función de la cantidad de actividad realizada en relación a la capacidad fisiológica del cerebro de estar despierto. Muchos factores pueden influir en la capacidad fisiológica del cerebro, tales como el ritmo circadiano, el exceso de trabajo antes de conducir, etc, lo que va a dar lugar a niveles de fatiga diferentes para cada uno de los conductores [Sub-Group, 2006] [Croo et al., 2001]. Estos factores son independientes de la actividad de conducir que está siendo realizada en un momento determinado, pero condicionan la aparición de fatiga, es decir, la fatiga no puede verse solamente como una función del tiempo que se lleva conduciendo.

El tema central de esta tesis es la detección de somnolencia en conductores y se aportan procedimientos innovadores de procesamiento de imagen, fusión de información y optimización de variables, intentando decrementar el número de falsos positivos que normalmente producen estos sistemas. Se profundizará en la relación que existe entre la somnolencia y la conducción.

1.2. Factores que influyen en la fatiga

Se han identificado los factores principales que influyen en la fatiga y somnolencia del conductor: dormir poco, horarios de trabajo, edad del conductor, experiencia en la conducción, sueño acumulado, la presencia de desordenes del sueño y la hora del día [Gander & James, 1999].

Los efectos de la **privación de sueño** son acumulativos y la pérdida de una o dos horas de sueño al día puede ocasionar serios problemas. El dormir por intervalos también es causa de pérdida de sueño [Dinges, 1995]. Dormir menos de 4 horas por la noche afecta acusadamente a

las características de la conducción. Los conductores que de media duermen aproximadamente 5 horas, aumentan el riesgo de tener un accidente de tráfico en 5 veces [Stutts et al., 2003]. Los factores más importantes para distinguir entre si un accidente ha sido consecuencia de la somnolencia o no son, la duración del último periodo de sueño y la cantidad de horas dormidas en las últimas 24 horas [Sweeney et al., 1995].

Los **horarios de los conductores** tienen impacto en la fatiga de los mismos, particularmente en vehículos comerciales. El riesgo de sufrir un accidente debido a la fatiga está muy relacionado con las horas totales de conducción y con factores propios del conductor: tener más de un trabajo, tener turno de noche y trabajar más de 60 horas semanales [Stutts et al., 2003].

Los **desórdenes del sueño** como la apnea o si la calidad del sueño es mala, causan un cansancio excesivo por el día, de manera que, derivan en un aumento del riesgo de sufrir accidentes [Young et al., 1997] [Stutts et al., 2003].

Los accidentes de tráfico, debidos a somnolencia, dependen de la **hora del día**, siendo más probables durante las horas tempranas de la mañana, desde las 2 hasta las 6 de la madrugada y también después de la comida, desde las 2 hasta las 4 de la tarde. Mackie y Miller encontraron una gran correlación entre la hora del día y el nivel de fatiga del conductor. Debido al ritmo circadiano, las personas suelen tener más sueño durante la tarde y noche, aunque no hayan sufrido ninguna privación de sueño [Dinges, 1995]. Los horarios de los conductores que interfieren con el ritmo natural circadiano, pueden generar privación de sueño y por lo tanto, tienen mayor riesgo de sufrir un accidente de tráfico [Jovanis et al., 1991].

Se dice que una situación se vuelve monótona cuando los estímulos recibidos de la misma, no cambian o los cambios son predecibles. La **monotonía de la carretera y de las condiciones ambientales** han sido identificadas como causa de somnolencia en una gran cantidad de estudios. [Åkerstedt, 1995] concluye que los accidentes debidos a somnolencia son más comunes durante la noche, en autovías anchas y monótonas. Por otro lado, [McCartt et al., 1996] establece que el 40% de los accidentes relacionados con somnolencia ocurren en autovías y el 30% de los accidentes en carreteras rurales son debidos a una conducción monótona. También [Thiffault & Bergeron, 2003] concluye que la fatiga es más temprana cuando la conducción es monótona y la demanda ambiental es baja.

La **personalidad del conductor y la edad** tienen influencia en la susceptibilidad a la somnolencia. [Verwey & Zaidel, 2000] comenta que existen grandes diferencias individuales en la susceptibilidad de caer dormido cuando se está conduciendo, mientras que [Campagne et al., 2004] ha comparado las características de tres grupos de edad en un simulador y encuentra que la vigilancia en la conducción se va deteriorando a medida que la edad avanza.

1.3. Síntomas de la fatiga y somnolencia en la conducción

La exploración y estudio de estos síntomas proporciona información muy valiosa para el desarrollo de sistemas de detección de fatiga o somnolencia. Cuando el conductor se encuentra fatigado se detectan en él diversos síntomas fisiológicos. Por ejemplo, la fatiga produce cambios en las ondas cerebrales, la actividad de los ojos varía, las expresiones faciales cambian, los movimientos de la cabeza disminuyen, el pulso cardíaco se decelera, la presión que se realiza sobre el volante disminuye, etc. En la última década se han aplicado diferentes técnicas de monitorización para la detección de fatiga, pudiéndose dividir en tres grandes grupos en función de los síntomas a monitorizar [Bergasa et al., 2006].

1.3.1. Síntomas fisiológicos

Las variables fisiológicas como la medida de la respuesta galvánica de la piel (GSR) y la conductividad están relacionadas con el estado psicológico de la persona [Berka et al., 2007].

Así mismo, la temperatura corporal del mismo, es un parámetro fisiológico importante que está relacionado con el estado del conductor: la temperatura corporal aumenta con infecciones, fiebre y además permite conocer de forma indirecta el estado del sistema nervioso autónomo [Oron-Gilad et al., 2008]. El electroencefalograma (EEG) y el electrocardiograma (ECG) proporciona información psicofisiológica acerca del estado de estrés, somnolencia, o reacciones emocionales [Papadelis et al., 2007].

1.3.2. Síntomas en el comportamiento del conductor

Existen patrones muy relacionados con la fatiga del conductor [Eskandarian et al., 2007]: los movimientos de la cabeza son significativamente menos frecuentes; el número de veces que el conductor se toca la barbilla, la cara, la cabeza, la oreja y los ojos se incrementa; se inclina ligeramente la cabeza hacia un lado a consecuencia de la relajación muscular del cuello; la actividad de los ojos se incrementa radicalmente; se producen episodios de cabeceo con más frecuencia y se tiende a adoptar posturas de relajación durante la conducción. En concreto, los patrones de parpadeo y el PERCLOS [F. & Richard, 1998], definido como el porcentaje de tiempo que los ojos se encuentran cerrados por debajo del 80 % de su nivel base, son las medidas más aceptadas en la literatura para la detección de fatiga o somnolencia en conductores.

La mayoría de estos patrones son analizados visualmente con técnicas de procesamiento de imagen, que son no intrusivas y monitorizan la somnolencia del conductor partiendo de la imagen capturada por cámaras instaladas en el vehículo. Estas técnicas son bastante efectivas porque la somnolencia se puede reflejar a través de la apariencia de la cara, de la actividad de la cabeza y los ojos.

1.3.3. Síntomas en la forma de conducir

Las medidas de las variables o señales del vehículo reflejan el comportamiento del conductor, por lo que se relacionan con el estado del conductor de una forma indirecta. Las variables usadas normalmente son: la fuerza realizada sobre los pedales, los cambios que sufre la velocidad del vehículo, los movimientos efectuados sobre el volante, la posición lateral del vehículo, los cambios involuntarios de carril y la distancia al vehículo precedente. Algunas de estas variables son sencillas de medir con el bus CAN del vehículo mientras que otras requieren de sistemas ADAS más complicados.

1.4. Validación de estudios de somnolencia

Independientemente del tipo de medida que se utilice para medir la inatención en conductores, uno de los mayores problemas encontrados en estos estudios es la dificultad de obtener una señal de referencia o control (Ground Truth) para su validación. Los estudios realizados, por razones de seguridad, son a menudo ejecutados en simuladores de conducción, ya que los ejercicios en condiciones reales tienen grandes limitaciones y dificultades.

Prácticamente todos los estudios mencionados han empleado simuladores de conducción, desde simuladores simples de mesa hasta sofisticados simuladores con movimiento en su base y de gran inmersión. Sin embargo, es extraño encontrar estudios donde la conducción se realice en carreteras reales, dos excepciones son los trabajos realizados por [Philip et al., 2005a, Philip et al., 2005b] y [Sagaspe et al., 2008] que realizan los ensayos bajo condiciones reales en un escenario de autovía.

Las pruebas en laboratorio permiten controlar aspectos ambientales, y por lo tanto la somnolencia puede ser inducida, y es posible usar equipos de medida sofisticados que son difíciles de integrar en vehículos reales. Sin embargo, las principales desventajas de los ejercicios en

simulación son el bajo realismo y el riesgo de que se interfiera en la somnolencia del conductor [Bos et al., 2008]. Otro problema importante, tanto en condiciones reales como en simulación, es la alteración del comportamiento espontáneo del conductor: la somnolencia en una conducción real es causada por una acumulación de fatiga y por la pesadez de una tarea monótona, especialmente en carreteras familiares y vehículos conocidos. La inusual experiencia en la participación en tales experimentos, especialmente si hay instrumentación invasiva, o existe un “efecto de guante blanco” debido a la presencia de investigadores, podría dificultar la somnolencia; o por el contrario, el alto nivel de estímulos en carreteras reales reduce la somnolencia [Åkerstedt et al., 2005]. Estos hechos limitan la eficacia de los experimentos. Para la validación de los modelos de detección de la somnolencia, es importante tener una gran cantidad de registros reales de usuarios tanto en estado de vigilia como de somnolencia.

1.5. Objetivos generales de la tesis

En el estado del arte existen diferentes propuestas para la detección de fatiga y somnolencia en conductores, pero se encuentran enfocadas, principalmente, a ambientes controlados en un simulador, y no resuelven el problema general. Muchos trabajos realizan la detección de somnolencia por medio de medidas fisiológicas del conductor, pero es muy complicado y nada obvio encontrar un patrón de somnolencia en estas señales, además, suelen ser métodos invasivos. Por ello, los métodos más prometedores, en el contexto de conducción real, se basan en medidas del comportamiento del conductor y de parámetros de la conducción realizada, debido a que son no intrusivas.

La detección de somnolencia puede dividirse en dos etapas. La primera, mide parámetros del conductor, habitualmente mediante técnicas de procesamiento de imágenes, y parámetros de la conducción a través de los propios sistemas ADAS del vehículo o añadiendo sensores específicos. La segunda, traduce las medidas anteriores en un nivel de somnolencia. Así como en la primera etapa ya se han obtenido resultados aceptables en el estado del arte y la investigación actual se centra en robustecer los algoritmos, hacerlos más rápidos, reducir los procesos de calibración, etc, en la segunda etapa hay todavía mucho trabajo por realizar.

Desde el año 2004, el grupo de investigación RobeSafe ha estado trabajando en el desarrollo de sistemas de asistencia a la conducción, tales como estimación de somnolencia [Bergasa et al., 2006], condiciones peligrosas en la conducción [Alcantarilla et al., 2008] y seguimiento robusto de la cara de conductores [Nuevo, 2009]. En este último trabajo, Nuevo expone un método novedoso para robustecer el seguimiento de la cara con una sola cámara. Sin embargo, no aborda el cálculo de parámetros de somnolencia. Otros trabajos previos en el grupo [Jiménez, 2011] usan la iluminación activa para realizar el seguimiento 3D de la mirada del conductor y a partir de ella inferir distracciones en la conducción.

Durante la última década, el número de trabajos relacionados con el seguimiento de la cara y evaluación visual de síntomas de somnolencia se ha incrementado notablemente, existiendo una investigación muy activa en los últimos años [Hansen & Ji, 2009]. Existen empresas que han comercializado algún producto, aunque no han publicado información técnica de los algoritmos empleados ni proporcionan un análisis de los resultados obtenidos con un cierto rigor científico. Además, las grandes compañías de automoción (Volvo, Toyota, Nissan, Mercedes-Benz, Saab) también han centrado su investigación en sistemas de detección de inatención en conductores. Algunas, como Saab y Toyota, emplean visión artificial para determinar características de la cara, frecuencia de parpadeo, etc. que permitan deducir el estado de somnolencia. Otras, como Mercedes Benz y Nissan, se inclinan por basarse en el comportamiento de la conducción cuando se le compara con un patrón previamente establecido de conducción normal. Volvo prefiere analizar el progreso del vehículo en la carretera para deducir el estado de fatiga del conductor.

En esta tesis se propone una técnica no intrusiva, basada en el procesado de imágenes mo-

noculares, consistente en la detección, seguimiento y caracterización de la apertura de los ojos, que trabaja óptimamente con distintos usuarios y en condiciones de conducción real soportado sobre un hardware propietario desarrollado para esta tesis. Esta medida, junto con otras señales derivadas de la conducción, entre las que hay que destacar el error de guiñada o heading error, se fusionarán para obtener un nivel de somnolencia del conductor, tanto en simulación como en condiciones reales. Por otro lado, se proporciona el desarrollo de todas las técnicas empleadas y un exhaustivo análisis de los resultados obtenidos. Por tanto, el objetivo principal de la tesis será desarrollar un sistema robusto para detectar somnolencia en conductores y contribuir a la mejora del conocimiento de cómo la somnolencia degrada el comportamiento en la conducción.

Específicamente, la tesis desarrollará un método robusto de cálculo del PERCLOS; analizará indicadores sensibles a la somnolencia; evaluará cómo los indicadores pueden mejorar si se les aplica técnicas de optimización estocástica e investigará la fusión de indicadores para mejorar la detección de somnolencia. Se considerará la fusión de señales de conducción y del conductor.

La validación de la propuesta se realizará en un simulador de tercera generación y en un vehículo real, llevándose a cabo una comparación entre detección de somnolencia en condiciones reales y en simulación. Las señales del simulador o vehículo real se agruparán en dos categorías: señales concernientes al conductor y señales correspondientes a la conducción. Del primer grupo se obtendrán imágenes de la cara del conductor y del segundo se registrarán los movimientos del volante, el ángulo de orientación del eje longitudinal del vehículo respecto a la dirección de la carretera, la posición del vehículo en el carril (separación a los márgenes), el error de guiñada, etc. En la figura 1.1 se representa el diagrama de bloques de las distintas etapas de la tesis.

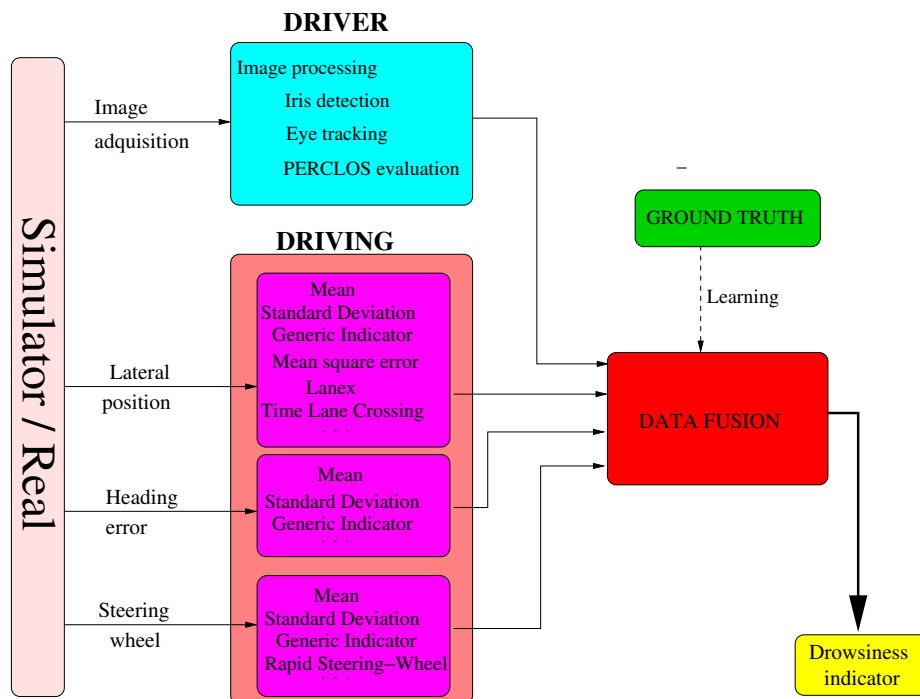


Figura 1.1: Diagrama general para la detección de somnolencia

Se implementará un sistema de captura de imágenes propio, formado por una cámara de infrarrojos, un sistema de iluminación sincronizada con la captura de imágenes y un hardware de control. Las imágenes obtenidas serán procesadas para detectar los ojos del conductor. Una vez identificados, se efectuará un seguimiento de los mismos ayudándose de un filtro de Kalman. En una etapa posterior, se realizarán medidas del iris del ojo y se calculará el PERCLOS, todo ello de forma automática mediante el uso de técnicas de procesamiento de imágenes.

Las señales de conducción proporcionadas por el simulador o los sistemas ADAS del vehículo real, (ángulo de giro del volante, posición en el carril, etc) también serán procesadas para obtener una serie de estadísticas de las mismas. Además, todas las señales a fusionar han de ser normalizadas, optimizadas y coincidentes en el tiempo o en el espacio.

Se estudiarán las combinaciones óptimas de señales a ser fusionadas mediante una red neuronal, que convenientemente entrenada por medio de una señal de referencia, determinará el nivel de somnolencia del conductor.

1.6. Estructura y organización de la tesis

Tras la introducción inicial, en el capítulo 2 se describe el estado actual de las diferentes técnicas de detección de somnolencia basadas en medidas subjetivas, medidas fisiológicas, medidas físicas ligadas a la detección de gestos o aspectos de la cara del conductor, medidas realizadas al vehículo y por último métodos híbridos que aprovechan las características más importantes de las diferentes medidas para obtener mejores resultados. Como las personas con fatiga muestran cambios visualmente observables, sobre todo en los ojos, la visión artificial puede ser una técnica natural y no intrusiva para detectar características visuales típicas de la somnolencia y se ha incidido, particularmente, en el estado actual del conocimiento en este campo. La referencia a algunos sistemas comerciales de detección de somnolencia y la comparativa de las características de los diferentes métodos finalizan este capítulo.

En el capítulo 3 se explica la metodología seguida para obtener la señal de referencia (“ground truth”) que refleja el estado real de cómo se encuentra el conductor, y se describen los diseños de los ensayos realizados, las peculiaridades de los escenarios de conducción y las señales registradas, tanto en ensayos de simulación como en condiciones reales.

El capítulo 4 expone las técnicas utilizadas y los métodos implementados para detectar la somnolencia del conductor mediante visión computacional, tanto en ensayos de simulación como en ensayos reales de conducción, en tiempo real, de forma robusta ante posibles cambios de iluminación, para distintos usuarios y sin requerir un proceso previo de calibración para cada usuario. Se parte de la adquisición de imágenes con un sistema monocular propio de alta resolución y frecuencia, que generará los vídeos de la cara del conductor. Mediante técnicas de procesamiento de imagen basadas en apariencia, se detectan la cara y los ojos del conductor para medir la apertura de los ojos por medio de técnicas de distribución de los niveles de gris. Por último, se evaluará el PERCLOS que según el estado del arte es un buen indicador del estado de fatiga del conductor.

En el capítulo 5 se describen diferentes indicadores útiles para detectar somnolencia, muchos de los cuales tienen parámetros que han de ser fijados de antemano y para no hacerlo de forma aleatoria se han empleado algoritmos genéticos que han permitido encontrar los valores óptimos. La clasificación de los estados se ha llevado a cabo, con cada uno de los indicadores individualmente, mediante una red neuronal y posteriormente se han fusionado los indicadores más relevantes.

Los resultados se exponen en el capítulo 6 analizando la calidad de las clasificaciones mediante la matriz de confusión y los valores de la sensibilidad y especificidad.

Por último las conclusiones y las líneas de trabajo que han quedado abiertas para el futuro se enumeran en el capítulo 7.

Capítulo 2

Estado del arte

En este capítulo se describe el estado actual de las diferentes técnicas de detección de somnolencia, que están basadas en medidas de las siguientes categorías: medidas subjetivas que utilizan la experiencia adquirida durante los estudios realizados; medidas fisiológicas tales como ondas cerebrales, velocidad de respiración, velocidad del pulso, etc, que son intrusivas al necesitar colocar electrodos en el sujeto a estudiar; medidas físicas ligadas a la detección de gestos o aspectos de la cara del conductor; medidas realizadas al vehículo que puedan inferir cuál es el estado en el que se encuentra el conductor, por ejemplo, los movimientos del volante o de la posición del vehículo en el carril; y por último, métodos híbridos que aprovechan las características más importantes de las diferentes medidas para obtener mejores resultados.

Dado que las personas con fatiga muestran cambios, visualmente observables, en ciertas características de la cara: larga duración de los parpadeos, movimiento lento de los ojos, pequeña apertura de los mismos, frecuentes cabeceos, indolencia facial, expresión y posturas de cansancio, etc, la visión artificial puede ser una técnica natural y no intrusiva para detectar características visuales típicas de la somnolencia y el estado de tales técnicas será descrito en el apartado 2.2.

La referencia a algunos sistemas comerciales de detección de somnolencia y la comparativa de las características de los diferentes métodos cierran este capítulo.

2.1. Estado del arte del método

Los métodos basados en medidas subjetivas de estimación del cansancio del conductor y en medidas fisiológicas no se emplean en condiciones reales de conducción, y sólo están desarrollados para la validación de sistemas. Los métodos híbridos proporcionan mayor fiabilidad si son comparados con los métodos de medidas físicas del conductor o con los métodos de medidas de la conducción, minimizando el número de falsas alarmas y haciendo de ellos sistemas que prometen aceptación en el mercado automovilístico.

2.1.1. Medidas subjetivas del estado de somnolencia

La escala de somnolencia de Karolinska (KSS) es una de las más empleadas para determinar el estado del conductor. En [Kaida et al., 2006] se investiga la validez y la fiabilidad del índice KSS usando la información del EEG, el comportamiento y otros indicadores subjetivos de somnolencia.

En [Ingre et al., 2006a] se comprueba la relación entre las medidas de somnolencia realizadas con el índice KSS y la duración del parpadeo y la desviación en el carril, obtenida en un simulador de altas prestaciones de base móvil. Según aumenta el índice KSS, se produce un aumento en la duración del parpadeo y en la desviación estandar del vehículo sobre el carril.

2.1.2. Medidas fisiológicas

El análisis espectral del EEG muestra que la transición de despierto a dormido puede ser detectada con cambios en las bajas frecuencias del EEG, hasta los 20 Hz, porque cuando los conductores comienzan a dormirse la actividad de la onda α es reemplazada por la actividad de la onda θ y este hecho es un indicador de somnolencia en el conductor.

También hay señales indicativas de somnolencia en el electrocardiograma (ECG) y la variabilidad del ritmo cardíaco (HRV). De estas señales, se ha calculado el espectro de potencia mediante la transformada discreta de Fourier y estudiado la correspondencia que existe entre el estado de la persona y su espectro para determinar las condiciones psicofisiológicas del conductor. Los parámetros de interés son la potencia total (de 0.03 a 0.4 Hz), la potencia de frecuencias bajas (de 0.05 a 0.15 Hz) y la potencia de frecuencias altas (de 0.15 a 0.4 Hz) [Faber, 2004] [Wright et al., 2007].

2.1.3. Medidas relacionadas con el conductor

La mayoría de los patrones de somnolencia son analizados visualmente con técnicas de procesamiento de imagen, que son no intrusivas y monitorizan la somnolencia del conductor partiendo de la imagen capturada por cámaras instaladas en el vehículo. Estas técnicas son bastante efectivas porque la somnolencia se puede reflejar a través de la apariencia de la cara, de la actividad de la cabeza y los ojos. Se han empleado para ello diferentes métodos: basándose en el espectro visible [D’Orazio et al., 2007a] y [Suzuki et al., 2006b] o en el espectro infrarrojo [Bergasa et al., 2006], [Ji & Yang, 2002b], [Bretzner & Krantz, 2005] y [Heinzmann et al., 2008], tanto con sistemas monoculares como estéreos [SmartEye, 2011] [SeeingMachines, 2011]. Basadas en estas técnicas se comercializan productos como Smart Eye [Bretzner & Krantz, 2005], Seeing Machines Face Lab [SeeingMachines, 2011], Seeing Machines DSS [Heinzmann et al., 2008] y Smart Eye Pro [SmartEye, 2011]. Sin embargo, están todavía limitados a condiciones ambientales controladas y requieren de un proceso complicado de calibración. Por esta razón, queda todavía un largo camino para la obtención de un producto de este tipo que se implante comercialmente en los vehículos.

En [Bergasa et al., 2006], el PERCLOS, la duración del cierre de los ojos, la frecuencia de parpadeo, y la mirada fija son normalizados y usados como entradas en un sistema de lógica borrosa para detección de fatiga. Diferentes términos lingüísticos y sus correspondientes reglas borrosas fueron distribuidas para cada una de las entradas empleando un conocimiento inducido basado en métodos jerárquicos borrosos, eligiendo el algoritmo más rápido para generar automáticamente las reglas que son consistentes, no tienen redundancia y son interpretables. Después, se aplica un proceso de simplificación para lograr una base de conocimiento más compacta.

En [Fan et al., 2010] se utilizan características Gabor aplicadas sobre la cara para detectar la fatiga. Primero se localiza la cara y se obtienen las características Gabor a diferentes escalas y orientaciones. Las características de la misma escala se fusionan para reducir la dimensionalidad y finalmente, se utiliza el algoritmo de AdaBoost para extraer las características más críticas y construir un clasificador robusto que detecte la somnolencia. Este método es robusto ante un amplio rango de sujetos expuestos a diferentes posturas y condiciones de iluminación.

En [Friedrichs & Yang, 2010b] se exploran 18 características del movimiento de los ojos para la detección de somnolencia. Empleando Principal Component Analysis (PCA) o Linear Discriminate Analysis (LDA) se reduce la dimensión de las características y con el algoritmo (SFFS) se seleccionan las características más representativas para construir el clasificador.

En [D’Orazio et al., 2007a] se usan los modelos de mezclas de Gaussianas para modelar estadísticas del comportamiento normal de la duración del cierre de los ojos (ECD), y de la frecuencia del cierre de los ojos (FEC), para cada una de las personas e identificar comportamientos anómalos en el cierre de los ojos.

2.1.4. Medidas relacionadas con la conducción

En [Farid et al., 2006a] se emplea la distancia al vehículo precedente, junto con el ángulo de giro del volante, mientras que [Wakita et al., 2005] identifican el estado del conductor basándose en señales relacionadas con la fuerza ejercida sobre los pedales y la velocidad del vehículo. La ventaja de estas técnicas es que las señales son sencillas de adquirir y tienen un significado muy claro de lo que está ocurriendo. Esta es una de las razones por las que estas señales se han introducido en los sistemas comerciales [Volvo, 2011], [Mercedes, 2011] y [Lexus, 2011] pero hay muy poca información técnica acerca de ellos. Además, se encuentran sujetos a limitaciones tales como el tipo de vehículo, la experiencia del conductor, características geométricas de la vía, condiciones de la carretera, y requieren entrenamiento previo del comportamiento del conductor [Torkkola et al., 2004].

Otras variables como los movimientos reversos o el paso por cero del volante también son utilizados para la detección de somnolencia. Los movimientos que hay que realizar en el volante para tener una correcta conducción, también se modifican con la fatiga. El número de movimientos de reverso, o pasos por cero del volante, es mayor cuando el conductor no está fatigado, mientras que disminuye cuando existen síntomas evidentes de fatiga. Con somnolencia se detectan movimientos lentos en el volante, movimientos de gran amplitud y grandes desviaciones estándares. Cuando el conductor se encuentra en un estado de fatiga, el ángulo de giro del volante y la trayectoria del vehículo pueden ser irregulares y el grado de desviación generalmente aumenta [Zhong et al., 2007]. Variables como las salidas que se producen del carril, el indicador SDLP (Standard Deviation Lateral Position), y la desviación máxima de carril se encuentran muy relacionadas con el cierre de los ojos del conductor. El error cuadrático medio de la desviación en el carril y el indicador de la desviación estandar de la posición del vehículo en el carril muestran un gran potencial como indicadores de fatiga o somnolencia [Stein, 1995], también la varianza y la media de la desviación de la orientación del vehículo (calculadas sobre una ventana temporal de tres minutos) son prometedores indicadores de fatiga [Sandberg & Wahde, 2008]. Sin embargo, no se encuentra correlación entre la fatiga y la aceleración o frenado que se realiza sobre el vehículo.

Un cambio en el estado mental induce un cambio en las características de la conducción. En [Furugori et al., 2005] la presión que se realiza sobre el asiento se mide durante largas sesiones de simulación, y los resultados muestran que esta presión está relacionada con la posición del vehículo sobre el carril y el índice subjetivo de fatiga. El algoritmo estima un índice de fatiga en intervalos de 10 minutos.

En [Farid et al., 2006b] se trata de distinguir entre conducción atenta y no atenta analizando la distancia al vehículo precedente y el ángulo de giro del volante. Se desarrolla un modelo, que funciona en tiempo real, usando Modelos de Markov con mezcla de Gaussianas para inferir la atención del conductor. En [Zhong et al., 2007] se hace un análisis espectral de la dinámica del ángulo del giro del volante y la posición del vehículo en el carril para detectar la fatiga. En [Takei & Furukawa, 2005] se emplea la teoría del caos sobre la variable del ángulo de giro del volante para estimar la fatiga de los conductores. En [Wakita et al., 2005] un modelo de mezcla de Gaussianas identifica el estado del conductor, basándose en la fuerza que se ejerce sobre los pedales y la velocidad del vehículo.

En [Torkkola et al., 2004] se utiliza la posición del giro del volante, la posición del pedal del acelerador, los extremos del carril y la curvatura de la carretera para inferir el estado del conductor. Inicialmente, las señales originales son preprocesadas (media, entropía, etc.), obteniendo una gran cantidad de características. Posteriormente, el algoritmo Random Forests (RF) [Breiman, 2001], es empleado para seleccionar los parámetros óptimos y extraer características. El clasificador también emplea técnicas RF.

En [Ersal et al., 2010] se utiliza un modelo de redes neuronales para caracterizar un patrón de

conducción normal. En unión con un sistema SVM, es capaz de discriminar entre una conducción normal y otra inatenta. La dinámica del vehículo y datos de la conducción tales como la posición del vehículo sobre el carril, la velocidad y la aceleración son utilizadas para obtener el patrón de conducción normal. La media y la desviación estandar de los residuos (la diferencia entre el comportamiento actual y el modelo) son elegidas como entradas al clasificador SVM. Los resultados muestran que la exactitud de la medida varía en función del individuo.

2.1.5. Métodos híbridos

Combinando medidas del conductor y de la conducción se puede incrementar la exactitud en la detección de inatención en conductores, como se muestra en la descripción de las siguientes referencias bibliográficas.

En [Eskandarian et al., 2007] se utilizan Redes Neuronales Artificiales (ANN) para analizar parámetros del vehículo y datos de cierre de ojos para inferir la fatiga o somnolencia del conductor. Dentro de los parámetros del vehículo se incluyen la velocidad, aceleración, posición del vehículo en el carril, ángulo de giro del volante, el freno y el ángulo de orientación del vehículo, los cuales son almacenados a una frecuencia de 20 Hz. La apertura de los ojos se toma a una frecuencia de 60Hz usando un PC equipado con Applied Scency Laboratory, el cual almacena el diámetro de la pupila gracias a la reflexión que genera la propia pupila. Se analizan los datos para identificar las variables que se encuentran relacionadas con la somnolencia. En este análisis se encuentran tres variables: el PERCLOS, el desplazamiento lateral del vehículo y el ángulo del giro del volante. Se implementan dos ANN detectoras de fatiga: una tomando como entrada el ángulo de giro del volante y la otra el ángulo de giro del volante y el PERCLOS. El ángulo de giro del volante es preprocesado antes de introducirlo a la entrada de la red basándose en la normalización del grado de curvatura de la carretera, en la discretización en diferentes rangos, en la codificación y en la acumulación cada 15 segundos. Para las señales de ojos, se aplica el mismo preprocesado excepto en la etapa de normalización. La conclusión que se obtiene es que el ángulo de giro del volante junto con el PERCLOS tiene más precisión y menores falsas alarmas que cuando sólo se emplea el ángulo de giro del volante.

En [Miyaji et al., 2009] para mejorar la exactitud de la medida de somnolencia se combina la desviación estandar de la dirección de la mirada, la orientación de la cara, los diámetros de las pupilas y el pulso cardiaco. Los parámetros de los ojos y de la cabeza son obtenidos con el sistema desarrollado por faceLAB, mientras que el pulso cardiaco se obtiene del electrocardiograma. Se utilizan dos técnicas de clasificación SVM y AdaBoost bajo las mismas condiciones para concluir que la clasificación realizada por AdaBoost obtiene mejores resultados que los obtenidos por SVM, con tiempos de proceso muy similares.

En [Sandberg, 2011] se utilizan técnicas de optimización estocástica para implementar nuevos indicadores con los que evaluar la somnolencia. Estos nuevos indicadores son más independientes de las características de cada conductor, mejorando la tasa de aciertos del estado del conductor, tanto en escenarios de simulación como reales. Se combinan, mediante redes neuronales, parámetros relacionados con la conducción y el conductor, en concreto, la información del conductor utilizada es un modelo del ciclo circadiano, denominado Sleep Wake Predictor (SWP). En la obtención de resultados de fusión no se utiliza el indicador de PERCLOS ya que los resultados obtenidos individualmente con él no son satisfactorios.

En la tabla 2.1 se resumen algunas características de los trabajos relacionados con métodos o estudios de detección de somnolencia. Se representan primero aquellos trabajos que están relacionados con parámetros del conductor, luego los relacionados con las variables de conducción y por último métodos híbridos que emplean las dos variables anteriores. En la columna *Parámetros analizados + Método* se presentan las variables de entrada utilizadas para la detección de somnolencia y los métodos de preprocesado y acondicionamiento para que éstas sean más repre-

sentativas de la somnolencia del conductor. La columna *Método de estimación somnolencia + (Tasa de acierto)* indica el método empleado para inferir el grado de somnolencia del conductor a partir de las señales de entrada de la columna anterior y la tasa de aciertos correspondiente a cada uno de los métodos. En aquellos casos que no se muestra es porque no se proporciona por los autores. La columna de tiempo real indica el tiempo de ejecución, cuando se pone un “ es porque no se proporciona.

2.2. Extracción de señales físicas del conductor con técnicas de visión artificial

El parpadeo, la expresión de la cara, la mirada perdida, son señales de las que un observador humano podría intuir, fácilmente, el estado de somnolencia de una persona. Los ojos se convierten en el objeto principal de estudio y han de ser detectados dentro de una escena. A continuación se presentan las principales técnicas existentes de detección de ojos.

2.2.1. Modelos basados en forma

En estos métodos se emplea un modelo genérico de ojo, basado en la forma, como diseño inicial. Esta plantilla se utiliza para encontrar los ojos dentro de la imagen. En [Nixon, 1985] se propone una aproximación haciendo uso de la transformada de Hough para medir el espacio entre ojos. El iris se modela por un círculo, mientras que el límite de la esclerótica se modela con una elipse. Este método consume mucho tiempo de cómputo, necesita imágenes muy contrastadas y solamente funciona con imágenes frontales de cara. También se puede utilizar plantillas deformables para la detección de los ojos cuando éstos no se encuentran en una posición frontal a la cámara. Para que se detecten los ojos fácilmente, se requiere que el modelo sea correctamente inicializado en las proximidades de los ojos. El método es computacionalmente costoso y requiere imágenes con muy buen contraste para converger rápidamente.

Los modelos de forma basados en aproximación pueden dividirse en dos grandes grupos: de forma fija o de forma variable. A los métodos de forma variable también se les denomina modelos deformables. Los modelos se construyen utilizando cualquier punto característico del ojo o su contorno. Puntos característicos pueden ser bordes, esquinas o alguna propiedad de la imagen después de pasar por algún filtro específico. El borde del iris junto con la córnea y la pupila, son características muy habituales a utilizar en estos modelos.

Los ojos abiertos están muy bien definidos por su forma, incluyendo el contorno del iris, de la pupila y la forma exterior de los párpados. La clasificación de los modelos de forma dependen de si el modelo es una simple elipse o por el contrario es un modelo más complejo. En este caso se puede trabajar con un modelo elíptico para la detección de los párpados o un modelo de naturaleza más compleja. Estos modelos normalmente están compuestos por dos elementos: uno correspondiente al modelo geométrico y otro de medida. Los parámetros del modelo geométrico definen las deformaciones posibles de la plantilla, conteniendo parámetros para las transformaciones rígidas y parámetros para las deformaciones de la plantilla no rígidas. Los modelos deformables dependen de un patrón deformable genérico y el ojo se localiza deformando el modelo por medio de una función de energía que se va minimizando. Lo importante de estos modelos son sus posibilidades ante cambios de escala y rotación.

■ Modelos elípticos simples

Muchas de las aplicaciones de seguimiento sólo necesitan la detección del iris o pupila para realizarlo. Dependiendo del punto de vista, tanto el iris como la pupila tienen forma elíptica y consecuentemente pueden ser modelados por los cinco parámetros que determinan la elipse en el plano: las dos coordenadas del centro, el semieje mayor, el semieje

Tipo	Referencia	Parámetros analizados + Método	Método de estimación somnolencia + (Tasa de acierto)	Tiempo real
Conductor	[D'Orazio et al., 2007b]	ECD (Duración cierre de ojos) y FEC (Frecuencia de ojos cerrados) + ANN	Mezcla de Gausianas para comparar modelos	No
	[Senaratne et al., 2007]	PERCLOS, Orientación de la cabeza y posición + SVM	Reglas Fuzzy (88%)	No - 6 fps
	[Suzuki et al., 2006a]	Duración y número de parpadeos + Características	Redes neuronales	-
	[Fan et al., 2010]	Expresiones faciales (bostezos) + Características Gabor	AdaBoost (70%)	No
	[Bergasa et al., 2006]	PERCLOS, dirección de la cara, frecuencia parpadeo + Características	Reglas Fuzzy (80%)	Si
	[Krajewski et al., 2009]	Giro de volante + STD	SVM + K-NN + Árboles de decisión (86.1%)	Si
Conducción	[Eskandarian & Mortazavi, 2007]	Giro de volante	ANN (85%)	-
	[Friedrichs & Yang, 2010c]	Posición vehículo carril (Lanex, ZigZags, TLC, etc), giro del volante (Velocidad, Pequeños movimientos de ajuste)	ANN (55%) + Clasificador Bayesiano (50%)	-
	[Zhong et al., 2007]	Giro de volante, posición vehículo carril + Análisis de energía(FFT)	-	-
	[Torkkola et al., 2004]	Giro volante, posición vehículo carril + Media, STD, entropía	Random Forest Classifier (80%)	-
	[Ersal et al., 2010]	Dinámica del vehículo, posición del vehículo, velocidad, aceleración + ANN modelo de conducción	SVM (80%)	Si
Híbrido	[Liang et al., 2007]	Parámetros relacionados con ojos, giro de volante y posición vehículo carril	SVM (81.1%)	No
	[Boyras et al., 2008]	Cierre de ojos, dirección de la cabeza, ángulo de giro del volante, velocidad y fuerza aplicada sobre el volante	ANN (90%) + Fuzzy (89%)	-
	[Sandberg & Wahde, 2008]	Parpadeos, PERCLOS, posición lateral del vehículo, giro de volante + TLC, Lanex, Indicadores genéricos de variabilidad, Modelo SWP + Algoritmos genéticos	ANN (90%)	-

Tabla 2.1: Métodos y estudios de detección de somnolencia

menor y la dirección del eje mayor. En [Kim & Ramakrishna, 1999] y [Pérez et al., 2003] se usan técnicas de umbralización de la imagen para estimar el centro de la pupila como centro de la elipse. Se usan técnicas de detección de bordes para extraer los límites de la pupila y de la esclerótica. Muchas regiones en la imagen tendrán una intensidad similar tanto en el iris como en la pupila y los umbrales sólo serán aplicables bajo determinadas condiciones. La transformada de Hough puede ser utilizada eficientemente para extraer el iris o la pupila [Nixon, 1985] [Zhu et al., 2002b], pero requiere una forma explícita, como puede ser una restricción simple de circularidad y, consecuentemente, el modelo sólo funcionaría con caras frontales y primeros planos. La demanda computacional de estos modelos puede ser reducida observando que sólo son necesarios dos parámetros para contemplar las deformaciones que puede sufrir el iris o la pupila ante movimientos de la cabeza. En [Hansen & Pece, 2005] se modela también el iris como una elipse, pero se ajusta a la imagen a través de los métodos de optimización EM y RANSAC. Son métodos probabilísticos que incorporan la información de los vecinos dentro del modelo, evitando así una detección de características implícitas. El objetivo del método es funcionar correctamente con aquellas imágenes donde es complicado localizar los niveles de umbralización y hacer los algoritmos más robustos frente a cambios de iluminación. En [Li et al., 2005] se propone el algoritmo de Starburst que detecta el iris por medio de un modelo de forma elíptico. Este método localiza los puntos donde existe mayor diferencia en niveles de gris, por medio de una técnica de rayos recursiva que posteriormente maximiza al modelo elíptico. El método que mejor maximiza a la elipse es encontrado con RANSAC. El algoritmo de Starburst es esencialmente un modelo de forma activo. Normalmente, los modelos de forma simples son eficientes y pueden modelar características como las del iris o la pupila desde cualquier punto de vista. Sin embargo, estos modelos no son capaces de adaptarse a variaciones del ojo tales como pestañas, las esquinas y las cejas.

■ Modelos de forma complejos

Con estos modelos, el ojo puede ser modelado con más detalle. Un ejemplo es el modelo deformable propuesto por [Yuille et al., 1989b]. El modelo de ojo deformable consiste en dos parábolas que representan los párpados y un círculo el iris, como se muestra en la figura 2.1. El modelo se ajusta a la imagen a través de la actualización de reglas, las cuales incorporan funciones de energía para los valles, bordes y picos en la imagen. Las investigaciones realizadas con estos modelos corroboran que la inicialización es crucial. Otros de los problemas que tienen estos modelos son su complejidad, y la dificultad de su funcionamiento cuando existen oclusiones debidas al cierre de ojos o movimientos de la cabeza muy acusados. El método propuesto por [Yuille et al., 1989a] requiere la presencia de cuatro esquinas en cada ojo: las esquinas izquierda y derecha de los párpados y los puntos de corte del iris con el párpado superior. Habrá cuatro esquinas cuando el iris esté parcialmente ocluido por el párpado superior.

Estos métodos son lógicos, generalmente precisos y bastante genéricos, pero tienen bastantes limitaciones, como son:

1. Computacionalmente pesados.
2. Requieren de imágenes muy bien contrastadas.
3. Necesitan ser inicializados con exactitud.
4. Los modelos deformables tienen problemas cuando la escena es iluminada con infrarrojos, debido a que desaparecen gran cantidad de bordes.
5. No funcionan correctamente cuando existen movimientos acusados de cabeza o existen oclusiones.

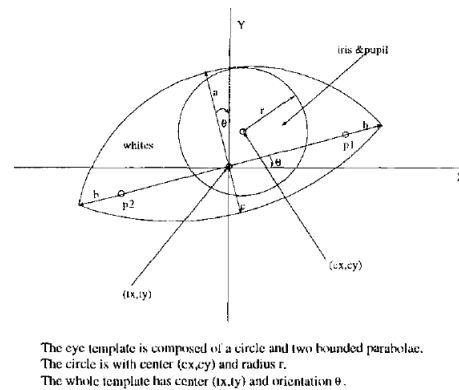


Figura 2.1: Modelo de forma deformable

2.2.2. Modelos basados en características

Estos modelos exploran las características locales del ojo para identificar un conjunto de ellas que lo definan, tales como la intensidad del iris, la pupila o las reflexiones de la córnea. Los modelos son menos sensibles a las variaciones de iluminación y puntos de vista. La detección no es fácil cuando los ojos están cerrados o parcialmente ocluidos por el pelo y también depende de la postura que adopte el sujeto.

■ Características locales de intensidad

La región de los ojos contiene muchas características que pueden ser detectadas por las diferencias encontradas en sus niveles de gris. En [Herpers et al., 1996] se propone un método que detecta características locales tales como bordes y líneas, sus orientaciones, longitudes y escalas, además de emplear un modelo de forma del ojo a priori. Inicialmente el método localiza un borde particular y posteriormente usa filtros de Gabor para hacer el seguimiento del borde del iris y las esquinas de los ojos. Basándose en el conocimiento del modelo del ojo y de las características, se inicia una estrategia de búsqueda secuencial para localizar la posición.

En [Reinders et al., 1996] se sugiere realizar un método basado en dos partes, una de ellas, con el nombre de microestructura, compuesta por el contorno de los párpados y las esquinas de los ojos. Por otra parte, se entrena una red neuronal solamente con imágenes de caras frontales y se obtienen muy buenos resultados de detección de ojos rotados y escalados, además de funcionar correctamente bajo diferentes condiciones de iluminación. Posteriormente se usa un modelo detallado de ojo para definir su localización.

■ Respuesta de filtros a características locales

Al aplicar un filtro sobre una imagen se realzan características particulares, mientras que se atenúan otras, por lo tanto, una batería de filtros puede realzar características deseadas de la imagen y minimizar características irrelevantes. De esta forma las regiones con características particulares pueden ser extraídas a través de la similitud de sus valores. En [Sirohey et al., 2002] se presenta un método para la detección del ojo usando filtros lineales y no lineales. Los bordes de la esclerótica son detectados con cuatro filtros lineales de Gabor, mientras que el filtro no lineal, se aplica para detectar las esquinas izquierda y derecha del ojo. Las esquinas sirven para determinar la región de búsqueda. También se aplica un método de detección de bordes para localizar el límite del iris. Aunque la parte superior del iris no sea visible, se van evaluando los píxeles del borde para formar una región angular en el borde del iris y poder marcar cual es su centro. Para detectar

el borde del párpado superior, todos los segmentos de borde se examinan y se ajustan a un polinomio de tercer grado. Los filtros no lineales obtienen mejores ratios de detección que los tradicionales métodos de filtrado lineal. Las imágenes de alta calidad son muy apropiadas para este método. En [D'Orazio et al., 2004] se convoluciona una imagen con un filtro circular y el valor máximo de la convolución proporciona el candidato del centro del iris. Realizando un filtrado de los candidatos por simetría y distancias heurísticas se localizan ambos ojos.

■ **Detección de pupila**

Cuando el ojo se encuentra parcialmente cerrado, la pupila es una característica determinante para la detección del ojo. La pupila y el iris deben ser oscuros comparándolos con su entorno y se puede aplicar una umbralización si el contraste es lo suficientemente grande. En [Yang et al., 1998] se introduce un método de umbralización adaptativa para la localización de las pupilas como dos regiones oscuras que satisfacen ciertas restricciones de un modelo de color de piel. Este método se encuentra limitado por los resultados obtenidos por el modelo de color de piel y falla en presencia de otras regiones oscuras tales como pestañas y cejas. Incluso es un error aplicar el mismo umbral para ambos ojos, especialmente si se consideran cambios en la orientación de la cara o condiciones variables de luminosidad. Para efectuar el seguimiento de los ojos se aplica una búsqueda en un área centrada alrededor de la posición del último ojo encontrado. Esta filosofía falla cuando hay otras regiones con intensidades similares o durante el cierre de los ojos. La detección de regiones oscuras puede ser apropiada cuando se usa luz infrarroja y no lo es tanto cuando se emplea luz visible.

La mayoría de los métodos descritos, anteriormente, se encuentran limitados por el hecho de no tener la posibilidad de detectar cuándo los ojos se encuentran cerrados. En [Tian et al., 2000] se propone un método para el seguimiento de los ojos y reconstruir los parámetros a través de un modelo de estado dual (abierto/ cerrado) para superar la limitación anterior. Este método requiere una inicialización manual del modelo del ojo. Las esquinas interiores y los párpados son seguidos usando un algoritmo modificado de Lucas-Kanade. Los bordes y la intensidad del iris se usan para extraer la información de la forma de los ojos, empleando los modelos deformables de [Yuille et al., 1989b]. Este método requiere imágenes muy contrastadas para un buen funcionamiento.

Los modelos basados en características generalmente son robustos ante cambios de iluminación. En las imágenes con un amplio campo de visión, los candidatos de ojos deben ser filtrados, debido a que muchas regiones pueden ser similares a los ojos. La detección de la pupila puede ser más eficiente a través de técnicas de generación de ojos rojos. Estas técnicas se comportan mejor en espacios interiores e incluso en entornos nocturnos, y es más complicado aplicarlas en espacios abiertos porque con sol o iluminación ambiente, el efecto de ojo rojo producido por la iluminación infrarroja desaparece.

2.2.3. Modelos basados en apariencia

Aunque la forma de los ojos sea un factor importante cuando se quiere diseñar un algoritmo de detección, también lo es su apariencia. Estos métodos detectan los ojos basándose en su apariencia fotométrica, como puede ser su distribución de color, y son independientes del objeto a detectar, por lo que a veces detectan objetos que no son ojos. Pueden ser aplicados en el dominio espacial, o en el dominio transformado. Uno de los principales beneficios de la detección de los ojos en el dominio transformado es que alivia los efectos de las variaciones de luminosidad por preservar bandas que son poco sensibles a los cambios de luz y eliminar bandas que son muy sensibles a dichos cambios. En la práctica, tales técnicas son solamente tolerantes a pequeños

cambios de iluminación.

Los métodos basados en apariencia pueden ser aplicados a imágenes patrón, donde la información espacial y de intensidad de cada píxel se tiene en cuenta, o imágenes integrales donde solo se tiene en cuenta la distribución de intensidad y se ignora la información espacial. Ambos métodos tienen problemas inherentes con los cambios de escala y rotación. Incluso cambios en la posición de la cabeza y movimientos del ojo pueden influenciar negativamente al funcionamiento del algoritmo.

Los modelos integrales utilizan técnicas estadísticas para analizar la distribución de intensidad de la apariencia de los objetos de entrada. Dada una imagen de test, la similitud entre los prototipos almacenados y la imagen de test se lleva a cabo en un espacio transformado. Normalmente se necesita una gran cantidad de imágenes de entrenamiento, bajo diferentes condiciones de iluminación, de diferentes sujetos y bajo diferentes orientaciones, pudiendo de esta forma alcanzar el diseño de un clasificador generalista.

Se pueden dividir en métodos de intensidad y de respuesta a filtros.

■ Dominio de la intensidad

Los métodos de intensidad usan la intensidad de la imagen como plantilla y marcan un umbral de detección en función de la similitud de las imágenes para realizar el posterior emparejamiento.

El seguimiento y la detección de ojos a través de la correlación de imágenes patrón es un método simple y efectivo. En [Grauman et al., 2001] se aplica la técnica de diferenciación de fondo y restricciones antropomórficas para iniciar los métodos de correlación. En [Hallinan, 1991] se propone un modelo de dos regiones de intensidad uniforme. Una región corresponde a la zona oscura del iris y la otra a la zona clara de la esclerótica. Esta aproximación obtiene el modelo de un ojo idealizado y por lo tanto se usan medidas estadísticas para tener en cuenta las variaciones de intensidad encontradas en las imágenes de ojos. En [Zhu et al., 2002a] se detectan los ojos usando máquinas de soporte vectorial. Todos los métodos aquí descritos tienen problemas con la invarianza a la escala, necesitando realizar un ajuste en el tamaño de la cara o realizar un aprendizaje con diferentes escalas de la imagen.

■ Respuesta de filtros

Los métodos que utilizan la respuesta a filtros usan la imagen filtrada directamente, sin hacer una selección de características a utilizar, como ocurría en los métodos de características.

En [Huang & Wechsler, 1999] se representan imágenes transformadas de ojos aplicando transformadas Wavelets. Se estudia la detección del ojo como un clasificador binomial. En los experimentos, se muestra la mejoría de la detección cuando se aplica un clasificador Wavelet RBF si se compara con la detección por imágenes de intensidad de niveles de gris. Después de la detección de los ojos, se obtiene una precisa información de la localización, como es el centro y el radio, haciendo uso de métodos de detección de contornos y regiones.

El algoritmo de Viola and Jones [Viola & Jones, 2004] aprende características Haar para la detección de la cara y los ojos por medio del clasificador en cascada Adaboost. En [Hansen & Hansen, 2006] se mejora la detección de los ojos por la información del reflejo que un foco luminoso produce en la córnea. En todos estos algoritmos, inicialmente se localiza la cara a múltiples escalas, y posteriormente se localizan los ojos dentro del área de la cara. La principal ventaja de las características Haar es su eficiencia computacional aunque su eficacia discriminatoria es limitada, especialmente en las etapas finales de los clasificadores en cascada. Para patrones complejos, el número de clasificadores en cascada debe ser elevado.

Las características y los procedimientos de selección usados en el detector de Viola and Jones son simples e intuitivos, sin embargo, el procedimiento de selección de características se basa en una búsqueda por fuerza bruta, entre una multitud de características predefinidas, lo que hace que el tiempo de cómputo y recursos empleados sean elevados. Las características Haar se aplican, frecuentemente, en la detección de ojos en caras frontales.

2.2.4. Modelos híbridos

Los modelos híbridos combinan métodos de forma y apariencia para beneficiarse de las virtudes de cada una de ellos.

Tratando de construir un modelo general de la forma del objeto, para la localización de ese objeto en imágenes particulares. La forma se modela implícitamente, mientras que la apariencia se modela explícitamente. En [Xie et al., 1998] se emplea un modelo basado en partes que emplea un modelo de forma a priori compuesto por subcomponentes. La región de los ojos se detecta, inicialmente umbralizando, y por búsquedas binarias, después se divide en varias regiones: dos regiones representando la esclerótica, la región del iris, la porción de iris que se encuentra ocluido y en la que no lo está. El iris y las pestañas se modelan por círculos y parábolas que tienen parámetros y una distribución de intensidad determinados.

En [Matsumoto & Zelinsky, 2000] se usan patrones de imagen de dos dimensiones para representar características faciales localizadas en un modelo tridimensional. El iris se localiza por la transformada de Hough. La imagen patrón en 2D asociada con el modelo 3D se utiliza para el emparejamiento. Las limitaciones de estos modelos basados en partes es que no modelan la imagen de intensidad directamente en las áreas correspondientes y que los modelos han de ser específicos para cada uno de los usuarios.

Otros métodos combinan modelos de forma y apariencia más explícitamente. Por ejemplo, en [Hansen et al., 2002a] se propone un método que combina la forma y la apariencia a través de un Modelo de Apariencia Activo (AAM). La forma y la apariencia son combinados en un modelo generativo, que se va ajustando a la imagen por cambios en sus parámetros de acuerdo a la deformación aprendida por el modelo.

Para la detección de caras, en [Cristinacce & Cootes, 2006] se usa un modelo AAM modificado, aplicando un modelo de apariencia para cada marca que se quiere detectar y un modelo de forma para las restricciones espaciales que tiene que cumplir el modelo. Los modelos activos de apariencia y sus variantes son capaces de modelar tanto las variaciones de forma como de textura. Los modelos de apariencia activos tienen una demanda computacional alta, aunque [Ishikawa et al., 2004] realiza una modificación para aumentar su eficiencia. Al igual que los modelos deformables, los modelos activos de apariencia necesitan ser inicializados cerca de la posición del ojo para obtener buenos resultados. Cometan errores cuando existe un giro acusado de la cabeza y tienen dificultades modelando la variabilidad de la apariencia del ojo.

2.2.5. Otros métodos

- **Detección de ojos bajo condiciones de iluminación infrarroja.**

En vídeos de interior se utiliza iluminación infrarroja para la detección y seguimiento de los ojos. Los métodos sin fuentes adicionales de iluminación se denominan métodos de iluminación pasiva, mientras que si existe iluminación adicional, se les llama métodos de iluminación activa. En los métodos de iluminación activa, las fuentes de iluminación infrarroja con una longitud de onda entre los 780 - 880 nm, son las más utilizadas. Hay cámaras comerciales sensibles a estas frecuencias, que son invisibles para el ojo humano y, por tanto, no distraen al usuario. La cantidad de luz emitida por estos sistemas, se encuentra sujeta a los estándares de seguridad internacional actualmente en desarrollo.

Si la fuente de luz se localiza en el eje óptico de la cámara, la imagen capturada muestra una pupila brillante. Este efecto es parecido al ocurrido en fotografía con los ojos rojos a consecuencia del flash, y es por ello que se llama imagen de ojos rojos. Sin embargo, si la fuente luminosa se encuentra fuera del eje óptico de la cámara se produce una iluminación de la escena en donde la pupila aparece negra.

Muchos investigadores han encontrado la relación entre la intensidad del brillo de la pupila y parámetros como la posición de la cabeza y la dirección de la mirada [Agustin et al., 2006] [Greivenkamp & Guyton, 1994] [Nguyen et al., 2002]. Sus estudios demuestran que la respuesta del brillo de la pupila varía mucho en función del sujeto y de los grupos étnicos.

Algunos objetos de fondo en una imagen pueden ser confundidos por pupilas oscuras, sin embargo, los brillos generados por la iluminación en las pupilas rara vez ocurren en otros objetos que no sean los ojos. Los modelos de ojos basados en iluminación activa normalmente usan la diferencia entre imágenes con y sin brillo de pupilas, necesiéndose una sincronización entre la cámara y la iluminación para que las imágenes se tomen en el momento en el que las pupilas están iluminadas [Bergasa et al., 2006] [Yoo & Chung, 2005] [Ji & Yang, 2002b]. Las ventajas de este método de diferenciación de imágenes es la robustez ante cambios de iluminación, simplicidad y eficiencia.

Los umbrales definidos manualmente son poco efectivos cuando hay variaciones en la iluminación, por eso deberían ser adaptativos a las variaciones de la respuesta de la pupila. En [Ji & Yang, 2002a] se usa la distancia de Kullback-Leibler para marcar un umbral. Son utilizados criterios temporales y de geometría para filtrar candidatos. Los movimientos de cabeza rápidos y acusados producen grandes diferencias a la hora de restar las imágenes, lo que constituye un problema a la hora de realizar la detección. En [Beymer & Flickner, 2003] y [D’Orazio et al., 2007a], para cerciorarse que los ojos se encuentran dentro de la imagen, se emplean varias cámaras en el sistema.

Existen muchísimos seguidores de ojos basados en la iluminación activa. Estos sistemas son particularmente eficientes en ambientes controlados donde no existen cambios bruscos de iluminación. Muchos de estos métodos requieren efectos de ojos rojos bien acusados para funcionar correctamente y también dependen del brillo y tamaño de la pupila. El brillo se ve afectado por muchos factores, tales como el cierre de los ojos, la oclusión debida a un movimiento de la cabeza, interferencias debidas a la iluminación externa, la distancia del sujeto a la cámara y las propiedades intrínsecas de cada uno de los ojos. Para solucionar alguno de estos problemas en [Haro et al., 2000] se propone combinar el seguimiento de la pupila por medio de apariencia de imágenes iluminadas con infrarrojo y características de movimiento, de tal forma que las pupilas puedan ser distinguidas de otros objetos brillantes de la escena. Para llevarlo a cabo, se verifican los contornos y el interior de la pupila usando un matching convencional y características del movimiento de los ojos. El seguimiento de los ojos no puede hacerse cuando se encuentran cerrados u ocluidos pero sí cuando haya cambios de iluminación en la escena. En [Zhu et al., 2002b] se propone un sistema en tiempo real y un método robusto para el seguimiento de ojos ante cambios de iluminación y orientaciones de la cara. El efecto de ojo rojo y su apariencia se utilizan simultáneamente para la detección y el seguimiento.

■ Operadores de simetría.

La simetría es una característica importante de la percepción humana y se ha investigado con el propósito de automatizar la detección del ojo y la cara. El operador de simetría de Reisfeld, tiene en cuenta regiones de alto contraste y simetría radial. Se basa más en intuiciones que en parámetros formales. Se analiza el gradiente de un punto con sus vecinos. Dentro de los vecinos, el gradiente de pares de puntos simétricamente situados sobre el

píxel central se usa como evidencia de una simetría radial, y es una contribución a una medida de simetría del punto central. El uso de esta técnica para la detección y seguimiento de ojos se encuentra limitada por la necesidad de umbrales para hacer la selección de las características. En [Loy & Zelinsky, 2003] se proponen la Transformada Simétrica Radial Rápida que tiene en cuenta la contribución de los puntos vecinos a la simetría del punto central.

- **Movimiento y parpadeo.**

Los parpadeos son involuntarios y periódicos y habitualmente simultáneos en cada uno de los ojos. Parpadear es necesario para mantener los ojos frescos y limpios. Esta dinámica de parpadeo puede ser aprovechada para su detección. En [Bala, 1997] se extrae la región de la cara gracias a la diferencia del fondo y de la información del color de la piel, después analiza las diferencias de luminancia entre sucesivas imágenes para la detección del parpadeo. Una vez determinada la región en donde se encuentra el ojo, se busca dentro del área un círculo negro como pupila. El centro de la pupila se toma como centro del ojo y se almacena para el siguiente proceso de matching. Un trabajo similar fue propuesto por [Crowley & Berard, 1997] donde la detección del parpadeo se basa en la diferencia de luminancia en sucesivas imágenes en pequeñas áreas en las proximidades del ojo. Sin embargo, la detección basada en parpadeos se encuentra actualmente limitada a detectar los ojos sólo en imágenes frontales.

2.3. Sistemas de detección de somnolencia comerciales

Aunque existen algunos productos comerciales para detectar la somnolencia, la mayoría de ellos se encuentran en fase experimental o son prototipos. Prácticamente todos hacen uso de cámaras y técnicas de procesamiento de imagen para no ser intrusivos. Las compañías más famosas de automoción (Volvo, Toyota, Nissan, Mercedes-Benz, Saab) se encuentran actualmente investigando en sistemas de detección de inatención en conductores.

Saab hace uso del sistema *SmartEye* para evaluar el número y frecuencia de parpadeos, la dirección y la orientación de la cara para la estimación de la somnolencia en el conductor. Sin embargo, no hay documentos explicativos acerca de la robustez del sistema en condiciones diurnas, nocturnas y no se proporciona una señal de referencia de la somnolencia del conductor.

Mercedes-Benz en 2009 introduce *Attention Assistem* que observa el comportamiento del conductor, y utiliza esa información para generar un perfil de conducción. Durante la conducción normal se va calculando la desviación respecto al perfil y el sistema determina cuando existe una desviación lo suficientemente grande para generar una alarma. Sólomente usa parámetros del vehículo para determinar la somnolencia, no requiriendo un hardware adicional. Sin embargo, este sistema necesita determinar un perfil de conducción para cada uno de los conductores, lo que puede afectar a su aceptación en el mercado.

En los últimos años han emergido nuevas empresas de componentes que fabrican sistemas para detectar y monitorizar niveles peligrosos de inatención del conductor. Muchos de estos proyectos se encuentran en desarrollo, validación o fase experimental. A continuación se describen las principales empresas que comercializan dispositivos o prototipos de este tipo.

Smart Eye ha desarrollado un sensor no intrusivo que mide movimientos de la cara y los ojos para una gran variedad de aplicaciones, incluyendo la seguridad en el transporte en simulación. Es un sistema de visión que estima la posición de la cabeza empleando un método robusto de seguimiento de características particulares de la cara y con un modelo tridimensional de la cabeza. El modelo inicial de la cabeza es genérico y adaptable a cada uno de los usuarios. Mientras la cara está siendo seguida, se evalúa la dirección de la mirada y la posición de los párpados gracias a la aplicación de un modelo 3D que utiliza entre 3 y 5 cámaras [SmartEye, 2011].

Seeing Machines faceLAB da la posibilidad de analizar una amplia gama de parámetros relacionados con la posición de la cabeza, la dirección de la mirada, el cierre de párpados bajo condiciones controladas de iluminación. Es una gran herramienta para analizar el estado de conductores en ejercicios de simulación, perdiendo efectividad en vehículos reales de test. Ha sido instalado en multitud de simuladores, incluyendo a National Advanced Driving Simulator en la Universidad de Iowa y en la universidad de Minnesota. También se ha instalado un sistema monocular en una cabina de demostración de Volvo. La somnolencia se estima en tiempo real por medio de la evaluación del PERCLOS realizando una calibración para cada uno de los usuarios, los datos pueden ser almacenados para su uso en sucesivos experimentos [SeeingMachines, 2011].

Attention Technology, Inc. ha desarrollado el sistema de monitorización de fatiga conocido con el nombre de *Driver Fatigue Monitor (DFM)*, que determina el nivel de somnolencia basándose en la velocidad de parpadeo del conductor [DriverFatigueMonitor, 2011]. No se tiene conocimiento que haya sido adoptado por ningún fabricante de automóviles.

Delphi Electronics and Safety ha comercializado el *Driver State Monitor (DSM)* que infiere el estado de somnolencia midiendo el AVERCLOS, que representa el porcentaje de tiempo que los ojos se encuentran completamente cerrados en una ventana temporal de un minuto [Edenborough et al., 2005]. Este sistema ha sido probado y testeado, como estudio previo a la comercialización, por Volvo Car Corporation.

2.4. Discusión

La cantidad de trabajos descritos en las secciones anteriores indica el interés de los investigadores en este área, aunque muchos de ellos no tengan por finalidad la detección de somnolencia. En la tabla 2.2 se resume y compara los diferentes métodos y se muestran las técnicas empleadas y las condiciones de funcionamiento de los algoritmos usando técnicas de visión.

La gran mayoría de los trabajos referenciados están desarrollados en condiciones de simulación y muy pocos los están en condiciones reales, por ello, el objetivo fundamental de detección de somnolencia de esta tesis se centrará en experimentos reales. Para el desarrollo de estos experimentos se utilizará, previamente, el conocimiento adquirido en los ensayos realizados en simulación, utilizando un simulador lo más cercano a la realidad, un simulador naturalista de tercera generación.

Los modelos descritos en el estado del arte para la detección y seguimiento de ojos son: modelos de forma, modelos de características, modelos de apariencia y modelos híbridos. El hecho de haber desarrollado tantas técnicas y tantos trabajos sólo indica que el tema no está cerrado y mucho menos resuelto de una forma satisfactoria para su uso en aplicaciones reales, fundamentalmente en exteriores. Los modelos híbridos son los que más se han aproximado a la resolución del problema, empleando conjuntamente métodos de forma y apariencia. En esta tesis, se empleará también un modelo híbrido, con la novedad de unir métodos de apariencia y características. La rapidez del algoritmo se mejorará aplicando técnicas de seguimiento en una secuencia de imágenes, mediante un filtro de Kalman empleado como predictor.

Del análisis del estado del arte se deduce que el PERCLOS es un índice robusto para detectar somnolencia por lo que es el método propuesto en esta tesis. Se obtendrá midiendo la altura del iris por medio de un ajuste de la proyección integral horizontal de la imagen a una distribución Gaussiana, cuya desviación típica proporcionará la apertura del ojo. De esta forma se logrará mejorar su robustez fundamentalmente en entornos exteriores.

La mayoría de los trabajos expuestos en el estado del arte están relacionados con técnicas de procesamiento de imagen pero no tienen como objetivo final medir la somnolencia, sino proporcionar métodos que evalúan parámetros faciales del conductor, como pueden ser la frecuencia de parpadeo, la posición del ojo en la imagen o simplemente su apertura. Los sistemas de seguimiento son complejos además de no funcionar correctamente ante cambios de iluminación.

Método	Información	Iluminación	Invariante	Requerimientos	Referencias
	P, I, E, C	D, F, IR	E, C, O	R, C, T, I	
Forma circular	P	IR	E, C	R, C	[Kim & Ramakrishna, 1999]
Forma elipsoidal	I, P	D, F, IR	E, C	C	[Ji & Yang, 2002a]
Forma elipsoidal	P	IR	E, C	C, T	[Haro et al., 2000]
Forma compleja	P, I, E	D, F	E, C	R, G	[Yuille et al., 1989b] [Yuille et al., 1989b]
Características	I, E	D	-	-	[Herpers et al., 1996]
Características	I	D	-	C	[D'Orazio et al., 2004] [Bala, 1997]
Características	I, E	D	-	C	[Sirohey et al., 2002]
Características	C	D	-	C	[Yang et al., 1998] [Tian et al., 2000]
Características	-	D	E, C, O	-	[Feris et al., 2000]
Características	P	D	-	C	[Yang et al., 1998] [Grauman et al., 2001] [Feris et al., 2000]
Apariencia	C	D, F	E, C	-	[Hansen & Hansen, 2006]
Simetría	P, I	-	E	-	[Loy & Zelinsky, 2003]
Movimiento	C	D, F	-	T	[Tian et al., 2000] [Grauman et al., 2001] [Bala, 1997]
Híbridos	P, I, E, C	D, F, IR	E, C	I	[Xie et al., 1998] [Matsumoto & Zelinsky, 2000] [Hansen et al., 2002b]

Tabla 2.2: Modelos de detección de ojos. La columna 'Método' corresponde a la categoría del método. La columna de 'Información' representa la información que puede ser obtenida directamente del modelo: Pupila (P), Iris (I), Esquinas (E), Ojo Completo (C). La columna 'Iluminación' indica bajo qué condiciones de luminosidad funciona el método: Dentro (D), Fuera (F), Iluminación IR (IR). La columna 'Invariante' considera la robustez a la escala (E), posición de la cabeza (C) o a la oclusión (O) debido al parpadeo de los ojos u ojos cerrados. La columna 'Requerimientos' incluye alta resolución (R) de la imagen, alto contraste (C), dependencias temporales (T), buena inicialización (I)

Por ejemplo, [Ryan et al., 2008] para determinar la apertura de los ojos utiliza un sistema de rayos muy lento computacionalmente y un modelo de forma, a pesar de que estos últimos son excesivamente complejos y no permiten trabajar en tiempo real. En [Zhu & Ji, 2005] y otros trabajos, se emplean técnicas de iluminación activa en funcionamiento nocturno pero se olvidan de las condiciones diurnas. Además, muchos de estos trabajos no tienen fiabilidad suficiente para determinar la somnolencia. Con respecto a sistemas comerciales, *SmartEye* [SmartEye, 2011] evalúa el número y frecuencia de parpadeos, la dirección y la orientación de la cara para la estimación de la somnolencia en el conductor. Sin embargo, no hay documentos explicativos acerca de la robustez del sistema en condiciones diurnas, nocturnas y no se proporciona una señal de referencia de la somnolencia del conductor. El sistema *faceLAB* de [SeeingMachines, 2011] da la posibilidad de analizar una amplia gama de parámetros relacionados con la posición de la cabeza, la dirección de la mirada, el cierre de párpados bajo condiciones controladas de iluminación. Es utilizado habitualmente en ejercicios de simulación, perdiendo efectividad en vehículos reales de test. Al igual que en el caso anterior al tratarse de un sistema comercial no hay referencias técnicas ni análisis de resultados que permitan una comparación rigurosa con nuestro sistema.

En la tesis se desarrollará un método que permitirá trabajar on-line y se ajustará a los cambios de iluminación y a las diferentes condiciones nocturnas y diurnas sin excesiva complejidad, utilizando iluminación activa pulsada cuando se requiera. El sistema de iluminación activo tiene el control de todos los parámetros de la cámara utilizada.

Además del PERCLOS, se trabajará con señales relacionadas con la conducción, tales como la posición del vehículo en el carril o el ángulo de guiñada, que serán obtenidas de forma diferente cuando los ensayos sean reales o en condiciones de simulación. En el caso real se emplearán sistemas ADAS embarcados para la obtención de las señales y en el caso de simulación, técnicas de adquisición de datos y registro. No será el objetivo de esta tesis profundizar en el funcionamiento de estos sistemas, pero sí se hará una descripción de lo que representa cada uno de ellos para fundamentar la fusión de información. Se estudiará cuáles son las variables más relacionadas con el estado del conductor, se les aplicará técnicas de optimización, y se seleccionará las redes neuronales por consideraciones de robustez y sencillez, además de ser muy utilizados en otros trabajos del estado del arte [Eskandarian & Mortazavi, 2007], [Friedrichs & Yang, 2010c], [Boyraz et al., 2008], [Sandberg & Wahde, 2008] presentando resultados en la tasa de acierto comprendidos en el rango del 55 % al 90 %, dependiendo ésta de los parámetros de entrada utilizados a la red.

En muy pocos desarrollos se trabaja con índices genéricos de variabilidad obtenidos de señales de entrada del vehículo [Sandberg, 2011], en lugar de trabajar con éstas directamente. En este trabajo se emplearán ambas técnicas. Además, los resultados no se limitarán a mostrar la tasa de aciertos, sino que se utilizarán diagramas gráficos en tela de araña cuando la clasificación a realizar sea de tres estados, mostrando las tasas de falsos positivos y falsos negativos que se generan.

Respecto a la fusión de datos, existen trabajos en el estado del arte [Bergasa et al., 2006], [Sandberg & Wahde, 2008], que sólo fusionan información relacionada con la cámara o con la conducción. En otros trabajos se hace fusión de datos del vehículo, pero no existen trabajos centrados en la fusión de parámetros del conductor y de la conducción. En esta tesis se demostrará que el porcentaje de aciertos es mucho mayor cuando se introduce el PERCLOS en el sistema. Además, en nuestro estudio se considerará otra variable, el heading error, que es independiente de la carretera y está muy relacionada con la somnolencia.

Capítulo 3

Generación de datos para la evaluación del sistema

Los datos obtenidos en esta tesis están ligados a dos grandes proyectos ubicados dentro del desarrollo de sistemas avanzados de asistencia a la conducción (ADAS) para vehículos de última generación. El primero de estos proyectos se encuentra englobado dentro de un proyecto estratégico singular del Ministerio de Ciencia e Innovación, llamado Cabina Inteligente (CABINTEC <http://www.cabintec.net>). Este proyecto basa su desarrollo en un simulador naturalista de conducción de tercera generación equivalente a la cabina de un camión. El segundo proyecto se encuentra englobado dentro del proyecto CENIT de Movilidad y Automoción con Redes de Transporte Avanzado (MARTA). Uno de los objetivos de este proyecto se centra en el desarrollo de un sistema que detecte la fatiga del conductor mediante técnicas de visión computacional en condiciones reales para la validación de otros sistemas fisiológicos también de detección de somnolencia, como por ejemplo, el estudio de la respiración del conductor.

El propósito de este capítulo es explicar la metodología seguida para obtener la señal de referencia (*ground truth*) que indica el estado del conductor, y describir el diseño de los ensayos realizados, las peculiaridades de los escenarios de conducción y las señales registradas, tanto en ensayos de simulación como en condiciones reales. El empleo de simulación para el entrenamiento e investigación en el campo de transporte por carretera es eficiente en relación al coste-resultado, ya que la simulación proporciona al conductor la posibilidad de sentirse inmerso en su espacio de trabajo habitual, dotando a los ensayos de la realidad necesaria. Sin embargo, no es suficiente para el desarrollo de un sistema los estudios de simulación pues siempre sufren desviaciones con respecto a los ensayos reales. De esta manera se ha contado con un total de 20 conductores, de los que se ha obtenido información muy valiosa y diversa para el desarrollo de esta tesis. Diez de estos conductores realizaron los ensayos de conducción en un simulador naturalista y, aprovechando la seguridad que proporciona el simulador, hicieron los ensayos de conducción con y sin privación de sueño, mientras que con los diez usuarios que realizaron las pruebas de conducción real no se ha trabajado con conductores previamente cansados por los riesgos de seguridad que conlleva. En los ensayos de simulación se definen dos sesiones para cada usuario, una de ellas de duración 1 hora que se realiza sin privación de sueño y la otra de duración 2 horas con privación de sueño. Los ensayos reales de conducción se distribuyen en 3 sesiones de duración 1 hora con un descanso de precaución entre sesiones de media hora. En la tabla 3.1 se hace referencia a la distribución y características de los ejercicios de conducción llevados a cabo.

Ensayos	Condiciones controladas	Condiciones reales
N de usuarios	10	10
Condiciones	Sin y con privación de sueño	Sin privación de sueño
Sesiones por condición	1(SPS) 2(PS)	3(SPS)
Duración de las sesiones	1 hora 1 hora	1 hora
Características	Simulador (A-1)	Diurno (AP-2 + AP-7)
Vehículo	IVECO Stralis	Audi A4

Tabla 3.1: Distribución y características de los ensayos de conducción. SPS: Sin Privación de Sueño. CPS: Con Privación de Sueño

3.1. Ensayos en simulación

Como se ha comentado, estos ensayos se realizaron dentro del proyecto singular estratégico CABINTEC usando un simulador naturalista de conducción avanzado de tercera generación con base móvil instalado en el Centro de Estudios e Investigaciones Técnicas de Gipuzkoa [CEIT, 2011], en San Sebastián. El propósito principal de los ensayos es llegar a inferir somnolencia en alguno de ellos debido a los siguientes aspectos: privación de sueño, hora del día o tiempo que se lleva conduciendo. Para ello se han realizado ensayos donde el conductor se encontraba privado del sueño o en una hora propicia para mostrar somnolencia. A continuación se va a describir el diseño del ensayo, el ejercicio de conducción y las señales registradas en los ensayos de simulación.

3.1.1. Diseño de los ensayos

Han participado 10 usuarios, 4 de ellos conductores profesionales y con larga experiencia en transporte por carretera. Todos los conductores son hombres, en un rango de edad de 32 a 60 años, siendo su valor medio de 45,25 años y la desviación típica de $\pm 5,15$ años. Todos los usuarios participantes gozan de buena salud y no sufren problemas de sueño como podría ser la apnea, tampoco han tomado cafeína ni ningún tipo de medicación al menos durante 24 horas antes, y no necesitan gafas para conducir.

Los usuarios se ven sometidos a dos pruebas, divididas en sesiones de una y dos horas, bajo diferentes condiciones de cansancio: sueño normal y cansancio acumulado, resultando un total de 30 sesiones de una hora de conducción, divididas en: 10 sesiones sin privación de sueño y 20 sesiones con cansancio acumulado para un estudio completo. Los usuarios han realizado un hábito normal a la hora de dormir con una media de 7,56 horas de sueño y una desviación estándar de $\pm 0,89$ horas, comprendidas en el rango de 12:00 de la noche a 9:00 de la mañana, permitiendo una hora de desviación del horario marcado durante las dos noches anteriores al ensayo. Se muestra la distribución de alguna de las pruebas realizadas a cada uno de los usuarios en la figura 3.1.

El proceso general del protocolo experimental busca recrear ambientalmente la estancia habitual del conductor para permitir a los investigadores tener una garantía de fiabilidad en los ensayos realizados. La tarea de conducción ha sido estudiada por el Instituto de Investigación en Seguridad y Factores Humanos (ESM) [ESM, 2011], centro privado de investigación de ámbito internacional, con una dilatada experiencia en el conocimiento de los factores humanos para el control de los riesgos en el transporte y la industria, avalándoles una gran cantidad de estudios acerca de somnolencia en conductores profesionales realizados en multitud de proyectos.

Con el diseño de estos ejercicios se persigue tener tramos bien definidos de somnolencia y de vigilia, por lo que se necesita realizar ensayos en los que el conductor muestre condiciones iniciales de fatiga y donde no las muestre, para poder comparar comportamientos y sacar conclusiones.

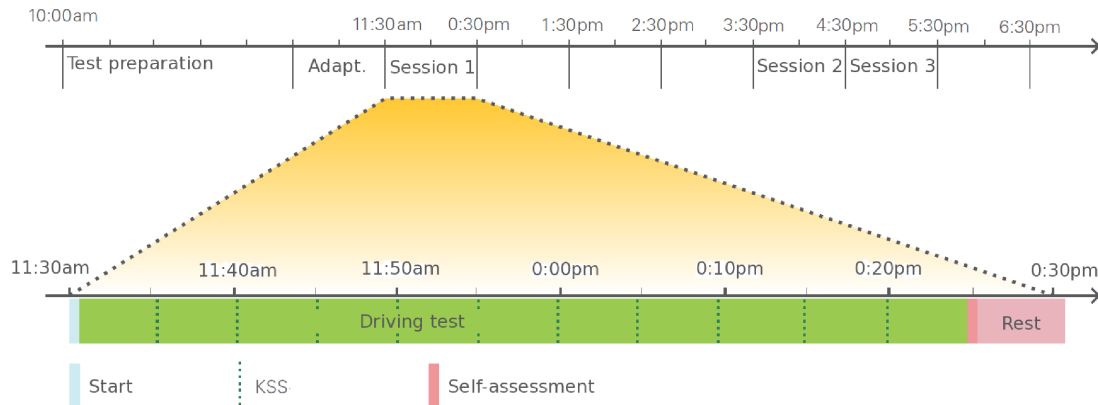


Figura 3.1: Diseño de pruebas en simulación

Los escenarios en los que se tienen condiciones iniciales de fatiga pueden ser programados teniendo en cuenta dos aspectos: privación de sueño (no dejar dormir lo suficiente al usuario antes de realizar la prueba de conducción) o cansancio acumulado (aprovechar que el usuario está cansado del trabajo efectuado durante todo el día). Este último caso será el empleado en las pruebas de madrugada realizadas sobre el conductor. Destacar que bajo estas condiciones se consiguió generar altos síntomas de somnolencia en la conducción en 5 de los usuarios participantes, obteniendo una información muy valiosa para la interpretación de la somnolencia en periodos de la conducción.

3.1.2. Escenario de conducción

La cabina del simulador es de un camión real de la marca Iveco Stralis que se encuentra anclada a una plataforma móvil con 8 grados de libertad de la empresa MOOG . La base móvil permite la simulación de aceleraciones, desaceleraciones y fuerzas centrífugas. Un sistema de pantalla cubre un campo de visión aproximado de 180 grados que se usa para recrear óptimamente el ambiente de conducción. El nivel de ruido y vibraciones dentro de la cabina se asemeja a las condiciones reales, haciendo que el conductor se vea inmerso en su tarea cotidiana de conducción. El simulador también tiene un puesto de control que sirve para comunicarse con el conductor y realizar una inspección visual de su comportamiento. Se tiene la posibilidad de trabajar con tres escenarios diseñados por ESM (urbano, interurbano y de montaña) con más de 250 km de carretera, además de ser ajustables parámetros como la hora del día en la que se realiza el ensayo, las condiciones meteorológicas, etc. Una imagen del simulador se muestra en la figura 3.2.

Los escenarios de conducción han sido elaborados tras un minucioso análisis con conductores profesionales de camión y autobús. En ellos se ha intentado recrear los diferentes entornos de trabajo integrando todos los tipos de vías existentes: carretera convencional, autopista y entorno urbano. El vehículo seleccionado para los ejercicios ha sido un camión rígido por ser el más conocido por los profesionales, además de ser el más genérico y con menor necesidad de experiencia para su manejo. De esta forma, se aminora el tiempo de adaptación al simulador, facilitando la dinámica de trabajo y evitando variables interdependientes. El conductor debe enfrentarse a dos bloques prácticos: el primero consta de un ejercicio específico de adaptación con el que se persigue que se habitúe a la conducción en el simulador y así interferir lo menos posible en su patrón de somnolencia y el segundo bloque está formado por ejercicios de conducción, útiles para la investigación de detección de somnolencia.

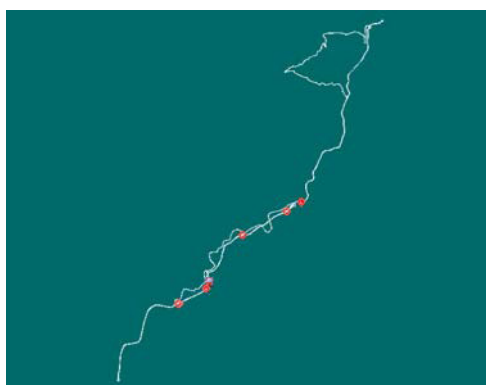
El ejercicio en concreto discurre, exclusivamente, por una autovía con las características



Figura 3.2: Fotografía del simulador, cedida por CEIT

propias de este tipo de vías: dos carriles para cada sentido, tarea relativamente monótona para el conductor de vehículos pesados y baja percepción de riesgo en general, lo cual favorece la fatiga y somnolencia del conductor. El trazado largo, sin percepción de riesgo y con un recorrido muy monótono por la poca o nula actividad cognitiva de la tarea que se desempeña en este tipo de vía, es propicio para que el conductor muestre síntomas de fatiga y somnolencia.

Concretamente, la autovía del escenario interurbano que recrea es la A-1 (fig. 3.3b). El conductor comienza en una carretera convencional que inmediatamente se une a la autovía por un carril de aceleración. El trayecto es cíclico, porque termina en una rotonda que permite realizar un cambio de sentido, en el extremo de inicio también existe una rotonda que permite, nuevamente, realizar otro cambio de sentido y así proseguir indefinidamente. Cada tramo de conducción es de aproximadamente 1 hora.



(a) Trayectoria



(b) Imagen de A-1 simulador

Figura 3.3: Características simulador

3.1.3. Señales registradas

Los psicólogos necesitan que el simulador tenga instalado sistemas específicos para registrar los comportamientos del profesional y así poder determinar el estado en el que se encuentra. El simulador naturalista registra variables del conductor y de la conducción. Las variables de la

conducción registradas son: el ángulo de giro del volante, la posición lateral del vehículo sobre el carril, el heading error o error de guiñada, la velocidad, la aceleración y muchas más variables que son almacenadas a una frecuencia de 60 Hz. También se registra la cara del conductor con un sistema de visión artificial basado en iluminación infrarroja como se detallará en el capítulo 4.1, proporcionando gran flexibilidad para estudiar la reacción y el comportamiento de los conductores [Kaida et al., 2006]. Atendiendo al tipo de variables almacenadas, se han dividido estas en tres grupos: relativas a la conducción, relativas al uso de mandos en el vehículo y otras medidas relativas a la carretera.

Medidas relativas a la conducción:

Velocidad. Velocidad del vehículo principal (Km/h).

Posición lateral. Distancia desde el punto central del vehículo al extremo derecho del carril en el que circula (m).

Ancho carril. Anchura del carril por el que se circula (m).

TTLC. Tiempo para cruzar la línea. Positivo o negativo dependiendo de si se cruza la línea derecha o la izquierda del carril. Siempre se da el valor del menor tiempo. Para la evaluación de este parámetro se tiene en cuenta la velocidad lateral del vehículo como se indica en la expresión $TTLC = \frac{velocidad\ lateral}{posicion\ lateral}$.

Error de ángulo de guiñada. Ángulo entre el eje longitudinal del vehículo y la tangente de la carretera.

Medidas relativas al uso de los mandos del vehículo:

Posición del volante. Ángulo de giro del volante (grados).

Porcentaje uso freno. Porcentaje de presión sobre el pedal del freno (0-100 %).

Porcentaje uso acelerador. Porcentaje de presión sobre el pedal del acelerador (0-100 %).

Otras medidas:

Trayectoria (X,Y) del vehículo. Coordenadas (X,Y) del centro del eje de dirección del vehículo dentro del escenario 3D por cada ciclo de simulación.

Trayectoria de la carretera (X,Y). Coordenadas cartesianas dentro del escenario 3D de la línea derecha del carril por donde circula el vehículo, para cada ciclo de simulación.

3.2. Ensayos en condiciones reales

La mayoría de las investigaciones de somnolencia en conductores han sido realizadas en simuladores de conducción porque en condiciones reales son más complicadas al ser necesario un conductor acompañante que resuelva las situaciones de peligro cuando el conductor objeto de estudio está somnoliento. A pesar de ello, existen algunos estudios en condiciones reales realizados a camioneros [Kecklund & Akerstedt, 1993] y [Philip et al., 2005a] con algunos casos registrados de somnolencia. A continuación se describen los principales aspectos de los ensayos en condiciones reales.

3.2.1. Diseño de los ensayos

Cada usuario lleva a cabo tres sesiones dentro del mismo ensayo. Cada sesión de conducción tiene una duración aproximada de 1 hora. Por razones de seguridad se ha exigido a los participantes dormir al menos 7 horas en el intervalo de 12:00 de la noche a 9:00 de la mañana, durante las tres noches anteriores al ensayo. La prueba es realizada por 10 usuarios con criterios de selección análogos a los realizados en el ensayo de simulación. La media de edad es de 41 años (comprendidas en el rango de 25 a 50) y fueron reclutados con los mismos criterios que los especificados en los ensayos de simulación.

3.2.2. Escenario de conducción

Todas las sesiones de conducción se inician en el centro de investigación de IDIADA y posteriormente se toma la carretera AP-2 dirección Lleida. Llegando a Lleida se conecta con la autovía A-2 hasta llegar a Jorba, cambiando de sentido y se finaliza de nuevo en IDIADA. El recorrido circular tiene una longitud total de 300 km, tardando aproximadamente 3 horas en realizarlo. Estas tres horas se han dividido en tres sesiones de aproximadamente 1 hora de duración. Las características de la vía en la mayoría del recorrido es de 9 metros de anchura, con un carril de conducción de 3.75 metros y la velocidad límite es de 120 Km/h, aunque hay segmentos de carretera donde la velocidad es inferior. El trazado de la trayectoria se representa en la figura 3.4. El coche empleado ha sido un Audi A4, al que se dotó de un sistema de aviso de salida de carril por visión artificial, y doble pedal para evitar problemas de seguridad. Una imagen del sistema de visión instalado en el vehículo se muestra en la figura 3.5.

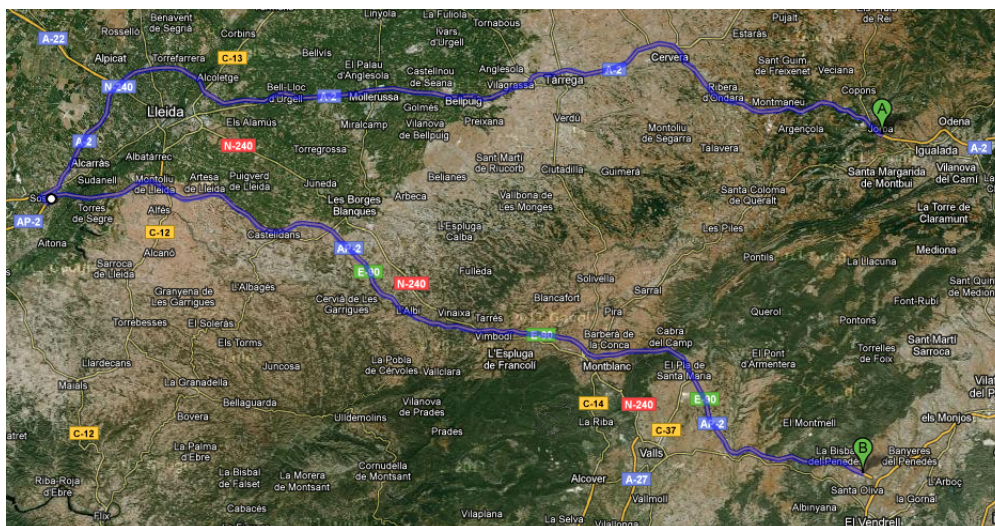


Figura 3.4: Trayectoria recorrido real

El acompañante del conductor es el responsable de la seguridad del vehículo y de sus ocupantes en el caso de que el conductor se durmiera u ocurriera cualquier otro percance. Además, tiene un asistente en la parte trasera del vehículo, responsable de hacer observaciones acerca de la somnolencia de los conductores. La conversación entre los conductores y cualquier ocupante del vehículo se permite para generar el nivel KSS de somnolencia del conductor y si suceden eventos de emergencia durante la sesión de conducción.



Figura 3.5: Sistema PERCLOS vehículo real

3.2.3. Señales registradas

El conjunto de señales registradas en este ensayo es similar al realizado en simulación. La estimación subjetiva del conductor se realiza cada cinco minutos.

Las señales de velocidad y ángulo de giro del volante fueron fácilmente muestreadas del bus CAN del vehículo, que tiene una frecuencia de refresco de 50 Hz. La posición lateral y las velocidades laterales fueron medidas por un sistema de Lane Departure Warning (LDW), consistente en una cámara y algoritmos de visión para detectar los bordes del carril y calcular la distancia del vehículo al mismo, además proporciona una señal de fiabilidad que indica la calidad de las medidas de la posición sobre el carril. La frecuencia de muestreo de este sistema es de 30 Hz. Los sistemas utilizados para el registro de estas señales quedan fuera de los objetivos de esta tesis.

3.3. Resumen de ensayos realizados

En la tabla 3.2 se muestra un resumen de los ensayos realizados en el transcurso de esta tesis, y se contempla la información relacionada con los ensayos de simulación en condiciones nocturnas, y los ensayos reales llevados a cabo en condiciones diurnas. Además, se divide la información en el diseño del ejercicio realizado, en las medidas realizadas sobre el conductor y en las medidas de conducción registradas. En las condiciones del usuario lo que se indica es si ha sido privado de sueño (PS) o no (NPS).

3.4. Generación de la señal de referencia (“ground truth”)

El problema de estimar la somnolencia entraña gran dificultad y se estudia, normalmente, como un problema de clasificación binaria, es decir, determinar si el conductor se encuentra en estado de vigilia o de somnolencia. El indicador de la escala de somnolencia de Karolinska (KSS) fue validado, originalmente, por la actividad eléctrica del cerebro, mediante las ondas alfa y teta del EEG, y movimientos lentos de ojos medidos por el EOG [Akerstedt & Gillberg, 1990]. Además,

Ensayos	Simulador	Real
Diseño		
N de usuarios	10	10
Condiciones usuario	NPS, PS	NPS
Sesiones por condición	1 2	3
Duración de las sesiones	1 hora	1 hora
Características	Condiciones controladas	Condiciones reales
Vehículo	IVECO Stralis	Audi A4
Medidas de somnolencia		
KSS	Si	Si
EEG, EOG	Si	Si
PERCLOS	Si	Si
Medidas del conductor		
Posición lateral	Si	Si
Ángulo de giro del volante	Si	Si
Error de guiñada	Si	No
Velocidad del vehículo	Si	Si

Tabla 3.2: Características de los ensayos realizados

[Kaida et al., 2006] han validado, recientemente, el indicador KSS por medio de la atenuación de la onda alfa del EEG y variables del comportamiento del conductor. La privación de sueño tiene efectos en la estimación del KSS como se indica en los trabajos de [Fairclough & Graham, 1999] y [Otmani et al., 2005]. Otros estudios [Hoddes et al., 1973] emplean ratios subjetivos basados en otras escalas, por ejemplo la SSS, que tiene resultados similares. La ventaja del índice KSS es que refleja fielmente el nivel de somnolencia del conductor.

En esta tesis se persigue clasificar en tres categorías: vigilia, fatiga y somnolencia aunque, en condiciones reales, como se explicará en la sección 6.4, sólo se ha clasificado en dos estados: vigilia y somnolencia. Se propone el uso de la escala KSS por ser una medida subjetiva que está muy relacionada con el estado del conductor, que se compone de nueve niveles, comprendidos entre extremadamente alerta hasta muy dormido como se muestra en la tabla 3.3. Para complementar esta medida subjetiva se propone la incorporación de la experiencia adquirida por expertos en el análisis de la conducción que tendrán en cuenta información de la cara del conductor y de señales de la conducción realizada, para generar la señal de referencia de tres niveles, como se muestra en la figura 3.6.

Como el índice de Karolinska se ha obtenido preguntando cada 5 minutos al conductor, se puede considerar que es un método intrusivo, pudiendo modificar el estado del conductor. Hay autores que dudan de la capacidad del conductor para autoevaluarse después de tres o cuatro horas de trabajo o por el cansancio acumulado [Schmidt et al., 2009] [Friedrichs & Yang, 2010b]. Sin embargo, hay otros que no lo consideran así [Akerstedt & Gillberg, 1990] [Sandberg et al., 2011]. En nuestro caso, como la frecuencia de las preguntas al conductor es muy baja y sus respuestas influyen poco en el cambio de su estado, no se ha detectado en las pruebas experimentales que sea un método intrusivo. El problema de detección de somnolencia puede ser definido como un problema de clasificación de datos, donde cada periodo de conducción es clasificado en dos o más clases. En principio, los nueve niveles definidos en el KSS pueden ser empleados para definir nueve clases diferentes, sin embargo, en el caso de clasificar en dos estados, *vigilia* y *somnolencia*, es posible agrupar los niveles de KSS del 1 al 6 para el estado de *vigilia* y del 7 al 9 para el estado de *somnolencia*.

Los resultados obtenidos por [Akerstedt & Gillberg, 1990] muestran que las modificaciones

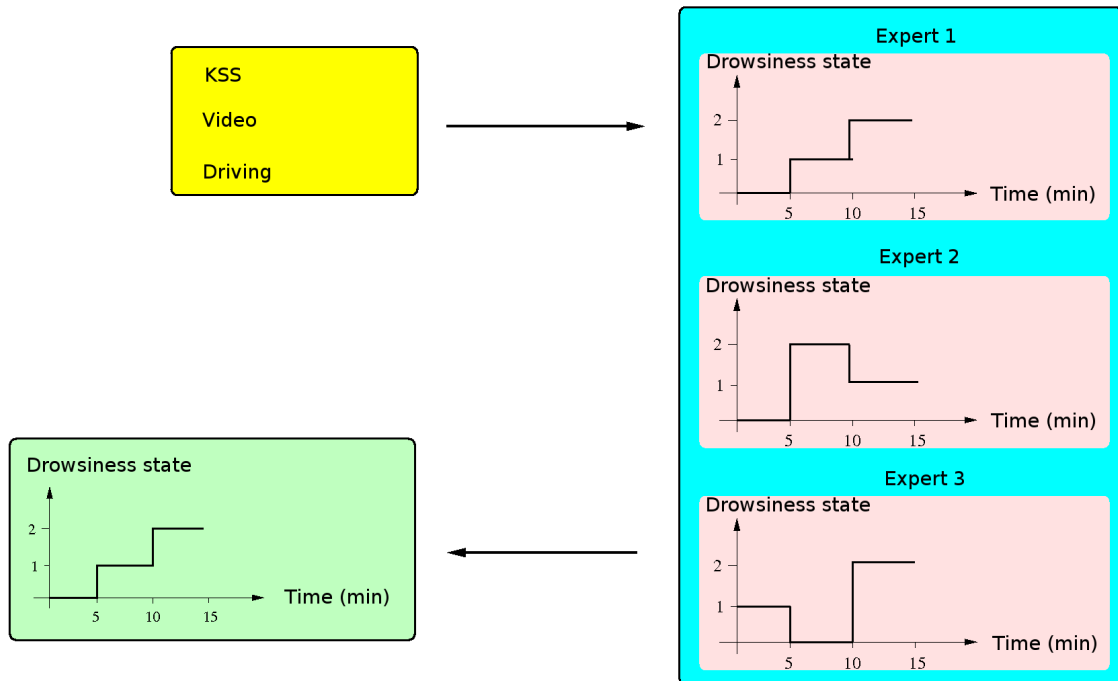


Figura 3.6: Señal ground truth

de actividad en las ondas cerebrales y los movimientos lentos de los ojos sólo se perciben cuando el nivel de Karolinska es superior a 7. En [Ingre et al., 2006b] se concluye que si el riesgo de tener un accidente con $KSS=6$ fuese uno, con $KSS = 7$ sería tres, con $KSS = 8$ sería trece, y con $KSS = 9$ sería ochenta y siete veces superior. Dado que en una aplicación práctica lo importante es avisar al conductor cuando se tenga cierta certeza de su fatiga, con disponer de dos estados de clasificación es suficiente en la práctica.

KSS	Descripción
1	Extremadamente alerta
2	Muy alerta
3	Alerta
4	Bastante alerta
5	Ni alerta ni dormido
6	Algunos síntomas de somnolencia
7	Dormido, no se hace un esfuerzo para estar despierto
8	Dormido, hay que hacer esfuerzos por estar despierto
9	Muy dormido, hay que realizar grandes esfuerzos para estar despierto

Tabla 3.3: Escala del índice de somnolencia KSS

Para generar la señal de referencia de una forma más objetiva, adicionalmente han participado tres expertos en el estudio del análisis de tarea y estado del conductor, pertenecientes al centro de investigación ESM [ESM, 2011], que, de forma independiente, analizaron off-line los ensayos efectuados por los conductores. Cada experto determinó, subjetivamente, el estado de somnolencia del conductor basándose en información visual de la cara (secuencia de vídeo), en información de la conducción (posición del vehículo dentro del carril, anticipación a la hora de tomar una curva, etc.), y sobre todo con el registro del índice KSS realizado por el propio

conductor. Para eliminar, en parte, la subjetividad de los expertos, la señal de referencia final se obtiene fusionando las tres señales anteriores, de modo que, se toma como estado el más repetido y en caso de no existir repetición el valor medio (fig. 3.6), con el objetivo de clasificar correctamente tantos periodos de vigilia, fatiga y somnolencia como sea posible.

Capítulo 4

Cálculo de parámetros visuales de la cara del conductor

Cada vez es más frecuente incluir técnicas de visión computacional en los sistemas avanzados de asistencia a conductores (ADAS). El objetivo de uno de ellos es detectar fatiga en conductores, ya que sin interferir en el conductor es posible identificar cabeceos, gestos faciales y características del ojo, parámetros que están muy relacionados con la somnolencia. De entre estos parámetros, los más utilizados son los que se derivan de la caracterización del ojo: la dirección de la mirada, el PERCLOS, el AVERCLOS (porcentaje de tiempo que los ojos están completamente cerrados evaluado en un minuto), la frecuencia de parpadeo, el intervalo donde los ojos se encuentran cerrados, cambios en el tamaño de la pupila, etc.

Existen numerosos trabajos basados en la caracterización del ojo para obtener parámetros característicos que determinen la detección de somnolencia, entre los que destacan los realizados por [D’Orazio et al., 2007a], [Senaratne et al., 2007], [Suzuki et al., 2006b] que engloban las diferentes técnicas expuestas en el estado del arte. En [D’Orazio et al., 2007a] se emplean dos cámaras instaladas en el salpicadero del vehículo para detectar los ojos del conductor por medio de simetría en la imagen. No se emplea una etapa previa de detección de cara por lo que la probabilidad de producirse falsos positivos aumenta. Detectada la región que comprende los ojos se emplea un clasificador neuronal para clasificar si el ojo se encuentra abierto o cerrado y a partir de ahí un modelo de mezcla de Gaussianas, que necesita ser inicializado, determina la somnolencia del conductor. Su elevado tiempo de cómputo, (15 fps), es una de las desventajas junto con un proceso inicial de calibración del modelo. En [Senaratne et al., 2007] se emplea un modelo complejo de marcas *Landmark Model Matching* (LMM) para determinar donde se localiza la cara y los ojos en la imagen, con objeto de aplicar técnicas de clasificación SVM sobre la región de los ojos y determinar cuando se encuentran abiertos o cerrados. De esta información se evalúa el PERCLOS para determinar el estado del conductor. El gran inconveniente radica en el tiempo de cómputo que es de 6 fps. En [Suzuki et al., 2006b] se evalúa el parpadeo del conductor para inferir la somnolencia. Se realiza aplicando una etapa inicial de detección de la cara en la imagen, una posterior localización de los ojos con técnicas de redes neuronales y características direccionales en cuatro direcciones de los bordes de la imagen, y aplicando técnicas de análisis de los niveles de gris de la región de los ojos se determina la separación entre el párpado superior e inferior.

A diferencia de la mayoría de los trabajos referenciados en el estado del arte sobre sistemas comerciales, en este capítulo, se exponen las técnicas utilizadas y los métodos implementados para detectar la somnolencia del conductor, tanto en ensayos de simulación como en ensayos reales de conducción, en tiempo real, con una velocidad de proceso nunca inferior a 30 fps, de forma robusta ante posibles cambios de iluminación, para distintos usuarios y sin requerir un proceso previo de calibración para cada usuario. En el diagrama de la figura 4.1 se representan los

distintos bloques en los que se divide el sistema de evaluación de parámetros de somnolencia. Está compuesto por cinco módulos, el primero de ellos consiste en la adquisición de imágenes con un sistema monocular propio de alta resolución y frecuencia, que generará las imágenes de entrada al sistema. En los dos siguientes módulos, mediante técnicas de procesado de imagen basadas en apariencia, se detectan la cara y los ojos del conductor para medir, de forma robusta en un cuarto módulo, la apertura del ojo por medio de técnicas de distribución de los niveles de gris. Por último, se evaluarán algunos parámetros visuales: el PERCLOS, la frecuencia de parpadeo y la mirada fija como indicadores del estado de fatiga del conductor, por ser considerados en el estado del arte como los que dan mayor información. No es objetivo de esta tesis el análisis de parámetros visuales en usuarios con gafas.

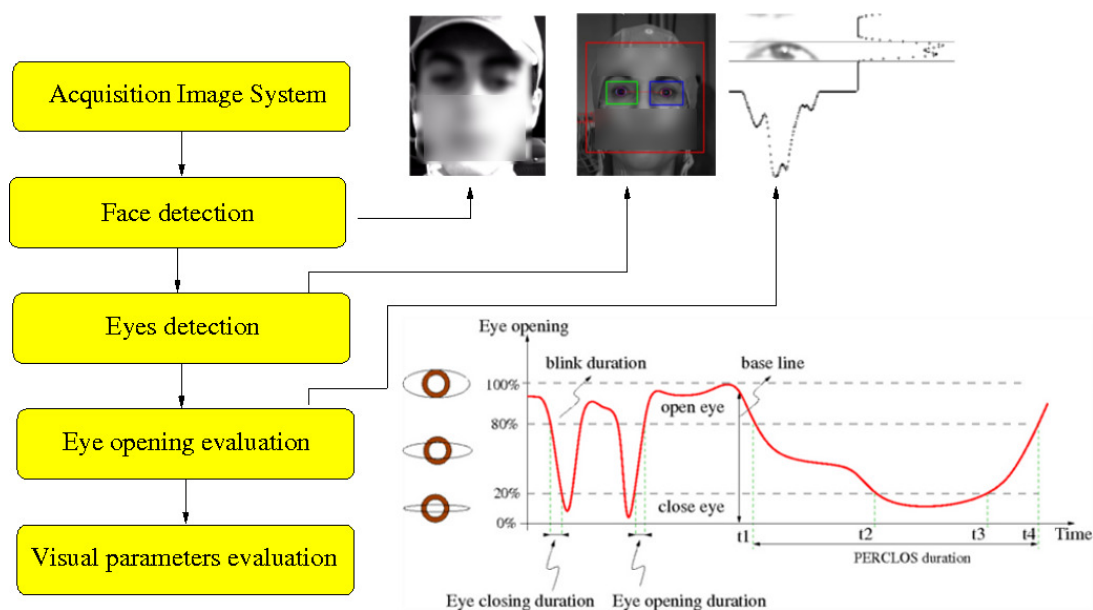


Figura 4.1: Arquitectura del desarrollo software

4.1. Sistema de adquisición de imágenes

El sistema de adquisición de imágenes está compuesto por una cámara de alta resolución y alta frecuencia de adquisición, control automático de ganancia e iluminación infrarroja que proporciona las imágenes de entrada al sistema. La captura puede realizarse en condiciones de iluminación muy diversas, en ambientes diurnos o nocturnos. En condiciones diurnas es necesario un control automático de ganancia para mitigar los cambios de iluminación que se producen debidos a sombras o deslumbramientos. Ante deslumbramientos muy acentuados, en una zona concreta de la imagen, el control automático de ganancia no solventará el problema, proporcionando imágenes como la mostrada en la figura 4.2.a. Los algoritmos desarrollados, tomarán la imagen de entrada para proporcionar medidas correctas de los parámetros de somnolencia.

En condiciones nocturnas, no se encuentran problemas debidos a los cambios de iluminación pero si será imprescindible iluminar la escena. Se ha optado por una iluminación infrarroja pulsada, sincronizada con la captura de imágenes, para provocar el menor cansancio posible al conductor. Es frecuente valerse del efecto de ojos rojos para la evaluación de parámetros relacionados con la cara del conductor, por ejemplo en [Zhu & Ji, 2005] se emplea un sistema de iluminación infrarroja compuesta por dos anillos concéntricos, el anillo interior genera efecto de ojos rojos y el exterior la iluminación general de la escena. Se va alternando la iluminación

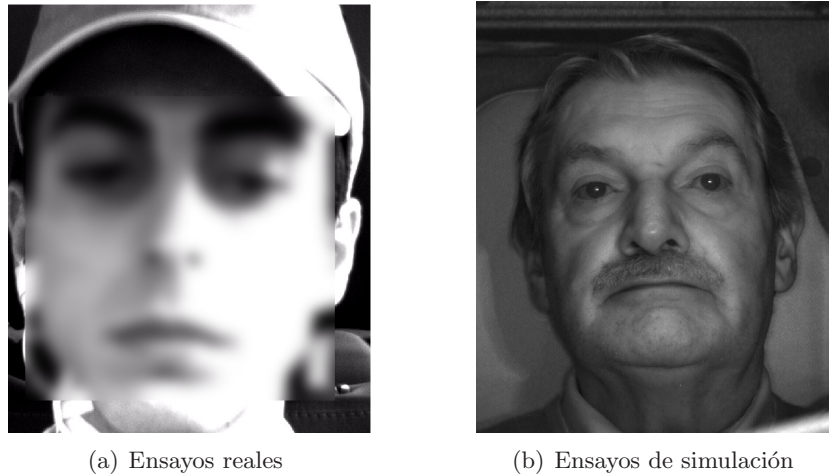


Figura 4.2: Imágenes de entrada: (a) Condiciones reales, con deslumbramiento en la derecha de la imagen; (b) Condiciones de simulación, sin deslumbramiento.

de los anillos para que, por diferencia de imágenes, sea relativamente sencilla la localización de los ojos. El inconveniente es la elevada intensidad de iluminación que incide sobre el conductor, necesaria para generar el efecto de ojos rojos, provocando cansancio y molestias en conductores profesionales que hagan largos recorridos en ambientes nocturnos. El sistema desarrollado en esta tesis parte de un nivel mínimo de iluminación de la escena en general, que lo proporciona un anillo circular centrado en el eje óptico de la cámara, con 10 diodos emisores de luz infrarroja, y una placa de control que sincroniza la captura de la imagen (sincronización externa de la cámara) con la iluminación, de esta forma no se busca el efecto de ojos rojos y la iluminación pulsada no fatiga al conductor. En la figura 4.2.b se muestra una imagen capturada en ambiente nocturno.

En las imágenes es importante disponer de la máxima información por unidad de superficie por lo que se requieren sensores CCD de alta resolución y ópticas de elevada calidad. Distancias focales comprendidas entre 10 y 15 mm, a una distancia del plano imagen de 1 metro, obtienen gran cantidad de información de la cara, pero el riesgo de salirse del plano de imagen cuando el conductor se mueve es alto. Con distancias focales comprendidas entre 5 y 9 mm el riesgo de salirse del plano imagen es menor, pero disminuye la información por unidad de superficie de la cara, realizándose medidas de parámetros menos exactas. Se ha optado por una óptica de distancia focal 12,5 mm que alcanza un buen compromiso entre la movilidad y la exactitud.

Otro factor a tener en cuenta es la velocidad de adquisición de imágenes. La cara del conductor puede realizar movimientos muy rápidos, que producirían imágenes borrosas si la velocidad de adquisición no fuese alta. Por ello, se ha elegido una cámara Basler Scout, dotada de uno de los mejores sensores CCD de la compañía Sony, que obtiene imágenes de excelente calidad, con alta sensibilidad al infrarrojo, con ganancias de 50dB en longitudes de onda cercanas a 800 nm, y una velocidad de adquisición de 30 frames por segundo a la resolución máxima de 1392x1040 píxeles, como se indica en la tabla 4.1.

4.2. Detección de cara

En el capítulo del estado del arte se han expuesto tres métodos para la detección de objetos en general, basándose en forma, apariencia y características. En esta tesis se ha optado por el algoritmo desarrollado por Paul Viola y Michael Jones [Viola & Jones, 2001] y las extensiones

Características	Valor
Tamaño del sensor en pixels	1392 x 1040
Tipo de sensor	Sony ICX285
Tecnología del sensor	Escaneo del sensor CCD progresivo
Tamaño óptico	2/3"
Tamaño del píxel	6.45 μm x 6.45 μm
Máxima velocidad de captura	30 frames por segundo
Sensibilidad infrarroja a 800 nm	50dB

Tabla 4.1: Características del sensor

realizada por Rainer Lienhart y Jochen Maydt [Lienhart & Maydt, 2002]. Este algoritmo está basado en técnicas de apariencia y proporciona muy buenos resultados para la detección de caras y ojos frontales. La eficiencia y versatilidad de su esquema hacen que el denominado detector de Viola & Jones, sea muy interesante para cualquier aplicación donde se necesite detectar objetos con pocos cambios de su apariencia. Este método estadístico, no necesita información previa de la tipología del objeto, ya que a partir de un conjunto de muestras de entrenamiento extrae la información relevante que lo caracteriza.

La librería de visión computacional OpenCV, escrita en C y C++, ha implementado el detector de objetos de Viola & Jones junto con la extensión de Rainer Lienhart y Jochen Maydt, denominándolo clasificador Haar. Además, la librería proporciona un conjunto de modelos preentrenados para la detección de algunos objetos. La función *cvHaarDetectObjects()* es la encargada de la detección de objetos con poca variabilidad estructural, haciendo uso de un modelo del objeto que previamente ha sido entrenado. Para la detección de caras frontales, el modelo preentrenado que mejores resultados proporciona es *Haarcascade_frontalface_alt2.xml*, aunque también tiene la posibilidad de entrenar y crear nuevos modelos mediante las funciones *cvCreateSamples()* y *cvHaarTraining()*. A pesar de su elevada tasa de aciertos en la detección, el modelo comete errores cuando la cara se presenta de perfil, debido a que la apariencia de la cara cambia considerablemente. En las figuras 4.3 y 4.4 se muestran secuencias de imágenes donde se aprecia el límite del detector de Viola & Jones cuando se producen giros de la cara del conductor. En exteriores, la no detección de la cara cuando hay giros de cabeza, no influye en el cálculo de los parámetros característicos del ojo puesto que éste desaparece y no puede ser caracterizado, como se muestra en la secuencia de figuras 4.3.c - 4.3.f. En interiores, el detector de Viola & Jones es capaz de soportar amplios giros de cabeza pero hay un límite a partir del cual los parámetros característicos del ojo no pueden ser evaluados por oclusión, como se muestra en la secuencia de figuras 4.4.c - 4.4.f. Sólo se ha aplicado un detector de caras frontales porque si está girada es porque no hay síntomas de somnolencia.

La arquitectura del detector de Viola & Jones está compuesta de nodos en cascada, cada uno de los cuales emplea AdaBoost como algoritmo de aprendizaje para la clasificación realizada por el nodo, e incorpora conceptos innovadores como los siguientes:

1. Se emplean **características Haar-like wavelets** aplicando un umbral a la suma y diferencia de las regiones rectangulares seleccionadas en la imagen. Las características Haar-like empleadas en el detector se muestran en la figura 4.5.
2. La **imagen integral** se emplea para acelerar el cálculo de las características Haar-like wavelets, consiguiendo una evaluación rápida del sumatorio de los valores de las regiones rectangulares.
3. Se emplean **algoritmos de aprendizaje AdaBoost** en los clasificadores débiles, que



Figura 4.3: Detección de cara ante giros en ensayos reales

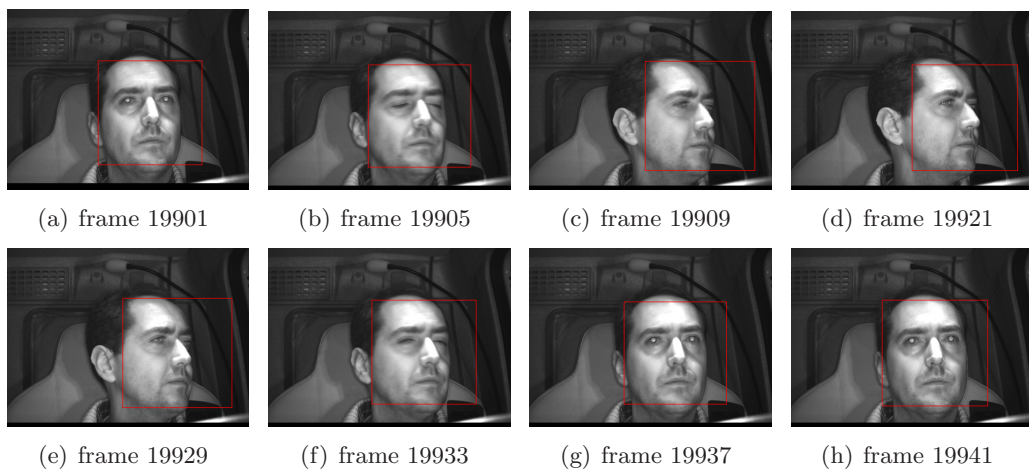


Figura 4.4: Detección de cara ante giros en ensayos de simulación

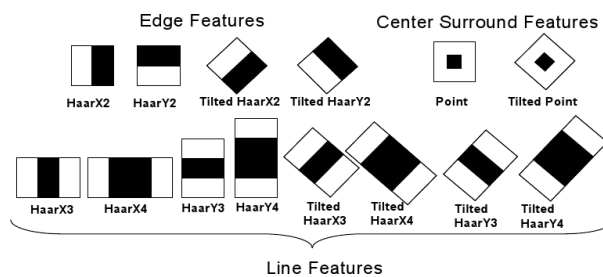


Figura 4.5: Características Haar-like utilizadas en el detector de Viola-Jones

evalúan las características Haar-like citadas, para determinar si la región de la imagen corresponde al objeto que se quiere detectar. La baja tasa de acierto de estos clasificadores débiles aislados, hace que se agrupen formando una arquitectura de clasificador fuerte donde la tasa de acierto es mucho mayor.

4. Los **clasificadores fuertes AdaBoost** se agrupan en cascada, para conseguir una clasi-

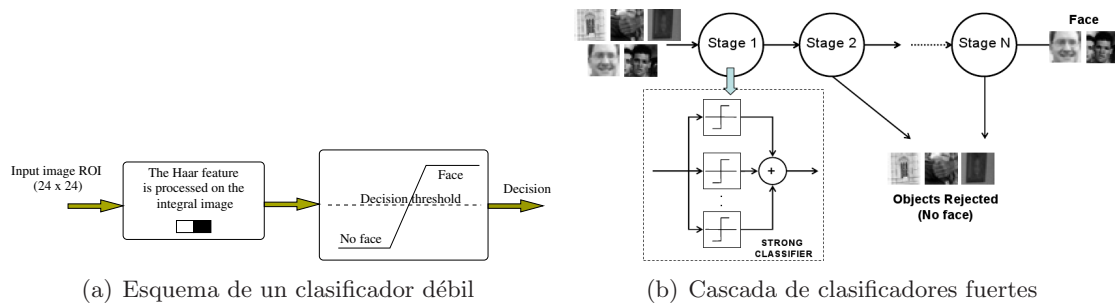


Figura 4.6: Arquitecturas del detector de Viola-Jones

ficación más restrictiva, de tal forma que las primeras etapas se encargan de descartar regiones de imágenes sencillas de clasificar, mientras que las últimas etapas discriminan entre imágenes más complicadas.

4.3. Detección de ojos

La exigencia de funcionamiento en tiempo real de la medida del PERCLOS, en una secuencia de 30 imágenes por segundo, junto a unas condiciones de iluminación cambiante, limita muchas de las técnicas descritas en el estado del arte que, o bien, tienen desarrollos matemáticos que exceden el tiempo de cómputo, o bien, con métodos computacionales eficientes, no soportan cambios de iluminación.

En [Wang et al., 2005] se ha desarrollado un sistema automático de detección de ojos empleando una base de datos FRGC 1.0. Aunque las características Haar dan buenos resultados ante cambios de patrones, tamaños y posiciones, solamente se pueden representar formas rectangulares mientras que, en el caso del ojo, la característica más importante es la forma redondeada del iris. Por ello, el método propone un aprendizaje estadístico de características discriminantes que separen el ojo de otros objetos. Igual que en el algoritmo de Viola & Jones, se utilizan múltiples clasificadores AdaBoost para mejorar la exactitud del detector y se aplican restricciones geométricas para localizar los ojos, por ejemplo, sólo se busca en la mitad superior de la cara. Normalmente, aparecen muchos candidatos próximos a la pupila, seleccionándose el valor medio de las detecciones realizadas. La frecuencia de funcionamiento de este proceso es de 10 fps, en una plataforma Pentium 4 de 2,6 G. Además de no proporcionar suficiente velocidad, la base de datos empleada tiene imágenes con iluminación controlada.

En [D’Orazio et al., 2007a] la detección del iris se realiza en tres etapas. En la primera se emplea una modificación de la transformada de Hough para la detección del iris sobre una región de la imagen, en la segunda etapa se busca una región simétrica y finalmente se aplica un clasificador neuronal que hace uso de transformaciones en la imagen para dar rapidez a la detección. Si el ojo es detectado correctamente, su posición se usa para delimitar la búsqueda en la imagen siguiente. El método no puede asegurar procesados a 30 fps, sobre todo cuando hay pérdida del ojo y su búsqueda no se encuentre encaminada, además no ha sido probado en condiciones cambiantes de iluminación.

En esta tesis se parte de que el rectángulo que encierra la cara, acota la zona de búsqueda de los ojos, aunque su detección es un problema complejo ya que su apariencia varía dependiendo de si están abiertos o cerrados, además de sufrir variaciones por colores, expresiones, tamaños relativos y condiciones de iluminación de la escena. Como en la detección de la cara, se ha empleado el detector de Viola & Jones para la localización de los ojos en la imagen, empleando la función *cvHaarDetectObjects* y los clasificadores preentrenados *haarcascade_righteye_2splits.xml*

y *haarcascade_lefteye_2splits.xml*. Los resultados se muestran en la figura 4.7, correspondiendo la primera fila a las detecciones realizadas por el clasificador del ojo izquierdo y la segunda fila a las realizadas por el clasificador del ojo derecho. Es importante notar que la clasificación de ojos es mucho menos precisa que la detección de cara y en la mayoría de las imágenes se proponen varios candidatos, como se muestra en 4.7.e, por lo que se hace necesario realizar un filtrado. En todas las imágenes de la secuencia se ha detectado el objeto correctamente, excepto en la imagen 4.7.d, que se ha clasificado el ojo izquierdo como derecho.

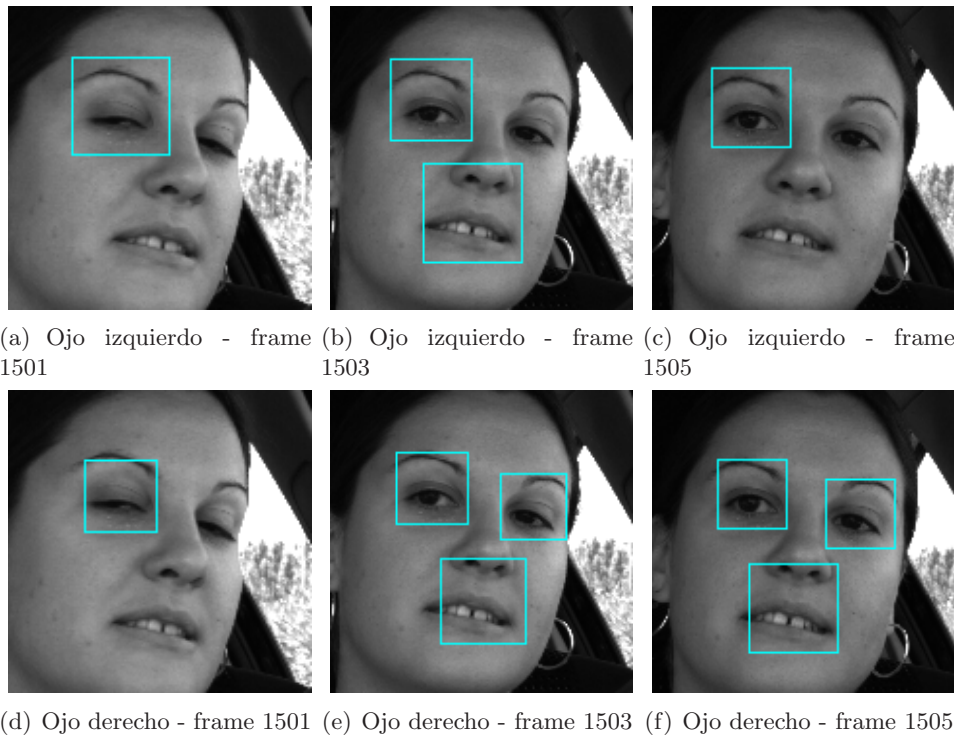


Figura 4.7: Detección de ojos en secuencia de imágenes de exteriores

Para solucionar este fallo se ha trabajado con el parámetro de entrada *min_neighbors* de la función *cvHaarDetectObjects*, que tiene como objetivo controlar las falsas detecciones ya que se tienen que encontrar más candidatos superpuestos del objeto, que el número indicado en el parámetro para que la detección se considere correcta. A pesar de todo se siguen produciendo falsos positivos, como se muestra en la secuencia de la figura 4.7. En esta tesis se ha desarrollado una nueva técnica de clustering para mejorar la detección de los ojos que parte de todos los candidatos encontrados (fijando el parámetro de la función *min_neighbors* a cero), como se muestra en la figura 4.8. En la figura 4.8.d puede apreciarse que ya proporciona un candidato para el ojo derecho.

El conjunto total de posibles candidatos son todas las detecciones realizadas por el clasificador de ojo izquierdo y derecho, como se aprecia en la figura 4.9.a, donde los cuadrados contenidos en ella son los cuadrados encontrados en las figuras 4.8.a y 4.8.d correspondientes al frame 1501 de una secuencia de exteriores. A continuación, un clasificador K-Means, agrupa los candidatos de ojos en cuatro clases (que se representan en diferentes colores), ya que experimentalmente se ha comprobado que los candidatos están mayoritariamente en ambos ojos, la nariz y la boca. La clasificación de los candidatos se indica en la secuencia de imágenes de la figura 4.9 donde aparecen tres clases por no haberse detectado ningún candidato de la cuarta.

Teniendo los candidatos clasificados, se evalúa el centro medio de los rectángulos que componen cada uno de los conjuntos y se aplica, sobre los centros calculados, restricciones de posición

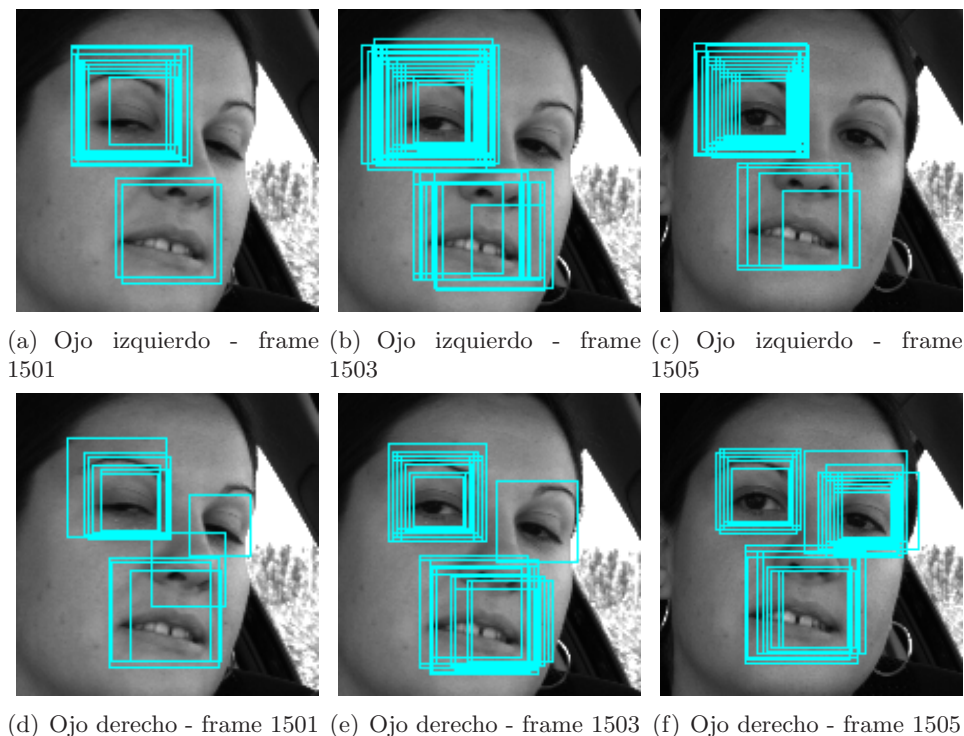


Figura 4.8: Detección de ojos en secuencia de imágenes de exteriores

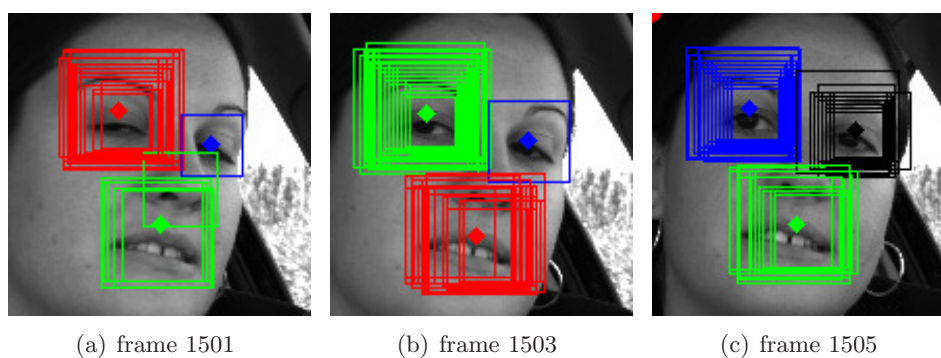


Figura 4.9: Clustering de candidatos de ojos

que tienen que cumplir, dos a dos, para que sean clasificados como ojos. Las restricciones se enumeran a continuación:

- Modelo genérico de la posición de los ojos.

Los candidatos son filtrados por características, de modo que la distancia entre los ojos según el eje x , tiene que estar comprendida entre un valor mínimo y máximo, y la distancia según el eje y , tiene que ser menor que una distancia máxima, como se indica en la figura 4.10.

- Distancia de Mahalanobis

Se calcula la distancia de Mahalanobis, dada por la ecuación 4.1, entre los centros proporcionados por el clasificador K-Means x_1 y la posición, ya calculada, de los ojos en el

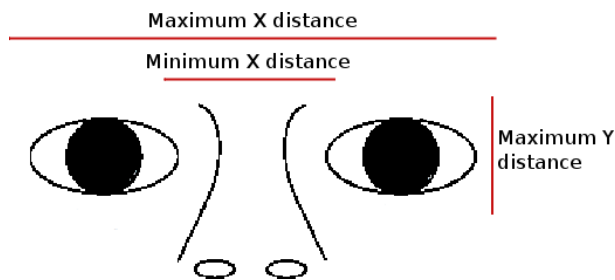


Figura 4.10: Restricciones de posición de los ojos

frame anterior \mathbf{x}_2 , ponderada por la matriz de covarianza \mathbf{S} que define los pesos entre las componentes de cada variable. Se elegirán aquellos centros cuya distancia sea mínima.

$$d(\mathbf{x}_1, \mathbf{x}_2) = \sqrt{(\mathbf{x}_1 - \mathbf{x}_2)^T \mathbf{S}^{-1} (\mathbf{x}_1 - \mathbf{x}_2)} \quad (4.1)$$

Con estas restricciones se determina, finalmente, la posición de los ojos. En la secuencia de imágenes de la figura 4.11 se muestran los resultados obtenidos cuando la cara está girando.

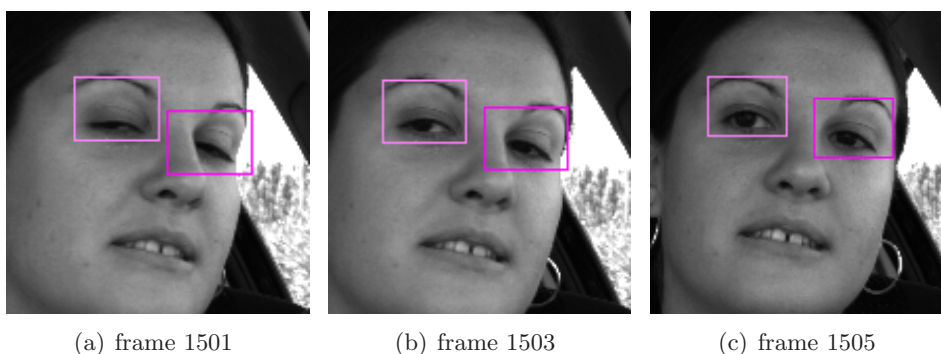


Figura 4.11: Candidatos finales de ojos

4.4. Evaluación de la apertura del ojo

Una vez que el ojo ha sido detectado, se trata de medir su apertura vertical y para ello se han empleado características en la distribución de niveles de gris junto con un modelo Gaussiano.

Estas técnicas se apoyan en la simetría del ojo, existiendo algunas propuestas como la de [Suzuki et al., 2006b] que detecta los párpados superior e inferior cortando verticalmente la imagen del ojo en varias secciones. En cada sección vertical se halla la curva de niveles de gris y se busca el valor máximo para determinar el centro de la pupila, que es la zona más oscura de la imagen. Después, en la derivada de la curva de niveles de gris, se buscan los valores máximo y mínimo a cada uno de los lados del centro anteriormente calculado, que corresponden a los párpados. Se repite el procedimiento para varias secciones de la subimagen que contiene el ojo y la distancia media entre los pares de puntos encontrados proporciona la medida entre párpados y por lo tanto la apertura del ojo (figura 4.12). El método no es lo suficientemente robusto cuando se producen cambios de iluminación, pues no siempre el valor máximo y mínimo corresponde con los párpados, depende mucho de la iluminación y de artefactos como las pestañas o sombras.

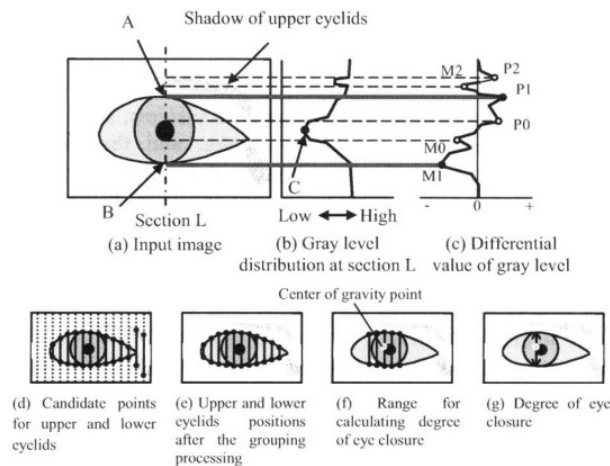


Figura 4.12: Técnica de apertura del ojo evaluada por *Suzuki et al, 2006*

En esta tesis, se ha desarrollado un método que permite evaluar la apertura del ojo en cualquier condición de iluminación y cualquier usuario aplicando inicialmente técnicas de pre-procesado para el acondicionamiento de la imagen de entrada y posteriormente se ha diseñado un sistema adaptativo para evaluar la apertura el ojo. El sistema se basa en aplicar sobre la imagen preprocesada una secuencia de filtros en función de la estimación de la apertura. Para evaluar dicha apertura se emplea un modelo Gaussiano a la proyección vertical obtenida de la imagen filtrada. Con este sistema se consigue mitigar los efectos ocasionados por los cambios de iluminación y posibles brillos de la imagen. Un esquema gráfico del proceso desarrollado se muestra en la figura 4.13.

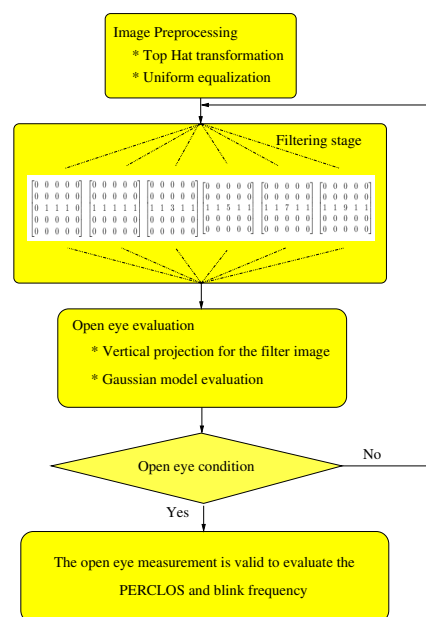


Figura 4.13: Diagrama general para evaluar la apertura del ojo

4.4.1. Preprocesado de imagen

Determinadas las subimágenes de cada uno de los ojos, se preprocesan modificando sus características (niveles de gris, contraste, eliminación de ruido, realce, etc) para conseguir una mejora de la imagen de entrada. Para ello, se han aplicado operaciones morfológicas y una ecualización uniforme del histograma para compensar efectos de sombras, reflejos, etc.

La operación morfológica aplicada ha sido la transformada del sombrero [Dougherty, 1992], mediante la función $cvMorphologyEx()$ con el parámetro de entrada CV_MOP_TOPHAT , que consiste en restar a la imagen original, la resultante de aplicar la operación morfológica de apertura (erosión y dilatación del proceso anterior). Con esta operación se eliminan objetos pequeños, debidos al ruido de la imagen, y se rellenan los huecos existentes en el interior de un objeto. Son operaciones rápidas de implementar que, en parte, eliminan brillos producidos por la iluminación infrarroja o deslumbramientos que se producen en condiciones diurnas. La imagen original se muestra en la figura 4.14.a y la resultante de aplicar la transformada del sombrero se muestra en la figura 4.14.b. Para contrarrestar las sombras y el realce que produce la transformada del sombrero, se realiza una ecualización uniforme del histograma con la función $cvEqualizeHist$, mostrándose el resultado en la figura 4.14.c.

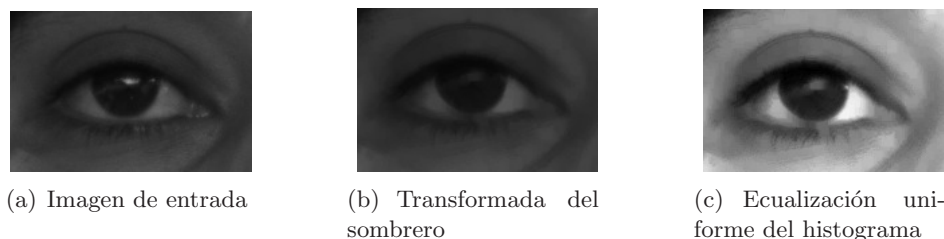


Figura 4.14: Preprocesado de las imágenes de entrada

4.4.2. Aplicación de filtros adaptativos

Para facilitar el proceso de localización del iris y evaluación de la apertura del ojo se aplica la secuencia de filtros que se muestran en la ecuación (4.2), con todos sus elementos a *cero* excepto la fila central que está compuesta por *unos* menos el término ancla que va aumentando desde 1 hasta 9 en función del filtro que se utilice.

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 3 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (4.2)$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 5 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 7 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 9 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Con esta operación de filtrado se pretende mitigar toda la información que proporciona la piel y posibles sombras que se puedan generar próximas al ojo para aislarlo en la subimagen que lo contiene. Además, realza las características horizontales que son, prácticamente, todas las que detallan a los ojos y los párpados. En las figuras 4.15 y 4.16 se muestran las imágenes resultantes al aplicar estos filtros a los ojos del conductor. La imagen de entrada 4.15.a es bastante oscura

y se tiene dificultad para apreciar detalles en ella. Aplicando sucesivamente los filtros de la ecuación 4.2 se puede discernir la información que caracteriza al ojo, párpados e iris. Aplicar el conjunto completo de filtros, hace que haya imágenes donde se pierde toda la información, como se muestra en la figura 4.15.g, por lo que se aplica un método adaptativo que determina cuántos filtros hay que aplicar a cada imagen.

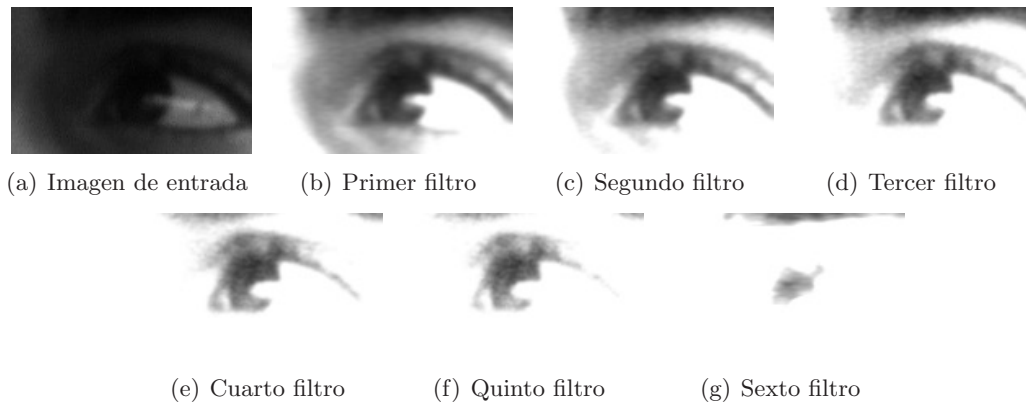


Figura 4.15: Filtros adaptativos aplicados a un ojo girado

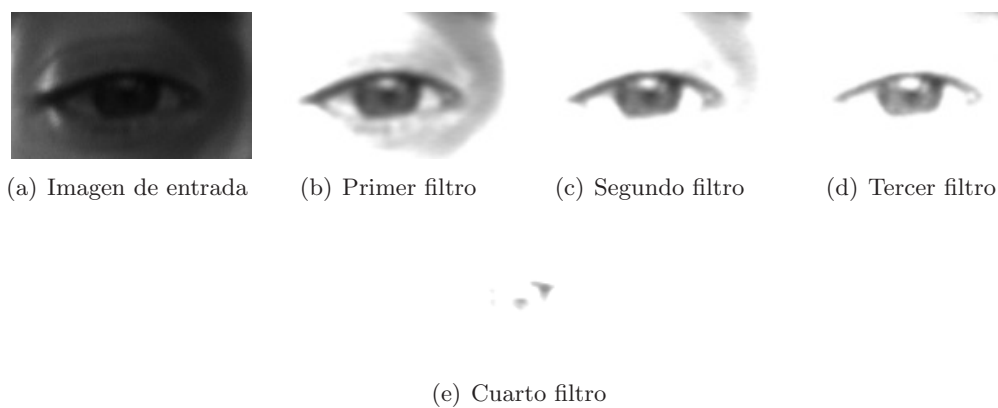


Figura 4.16: Filtros adaptativos aplicados a un ojo frontal

La aplicación de un filtro u otro, depende de la estimación de la apertura del ojo. Si la estimación es superior a un máximo determinado, se aplica el siguiente filtro de la lista, y si no lo es, se considera una medida válida. El valor máximo se ha fijado en 30 píxeles que, con el sistema de captura diseñado, es válido para todos los usuarios.

4.4.3. Evaluación de la apertura y posición del ojo en imágenes filtradas

Sobre la imagen filtrada se aplican los conceptos de integral proyectiva vertical, que corresponde a la suma del nivel de gris de todos los píxeles de una columna, aplicado a todas las columnas, y el de integral proyectiva horizontal que es la suma del nivel de gris de todos los píxeles de una fila, aplicado a todas las filas. Dada una imagen I , de dimensiones $W \times H$, las integrales proyectivas se encuentran definidas por las ecuaciones (4.3) y (4.4).

$$P_H(y) = \frac{1}{W} \cdot \sum_{x=0}^{W-1} I(x, y); \forall y = 0, \dots, H - 1 \quad (4.3)$$

$$P_V(x) = \frac{1}{H} \cdot \sum_{y=0}^{H-1} I(x, y); \forall x = 0, \dots, W - 1 \quad (4.4)$$

Aplicando estas ecuaciones sobre la imagen se obtienen señales continuas que se corresponden con la distribución de los niveles de gris en ambos ejes. En la figura 4.17 se representan algunos ejemplos.

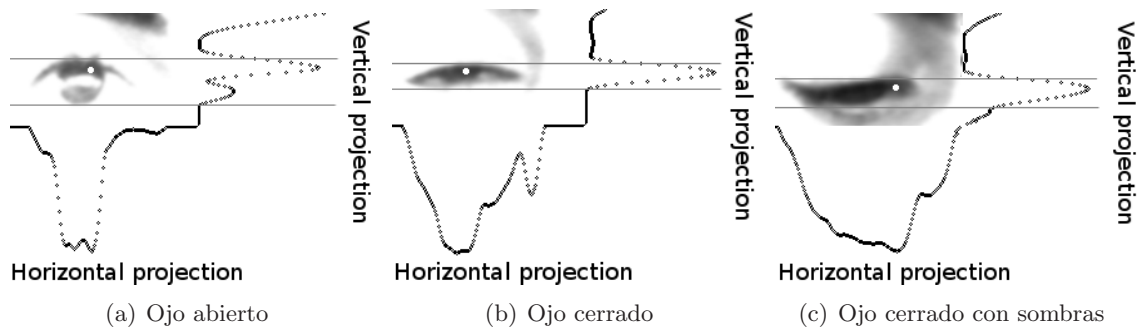


Figura 4.17: Cálculo de proyecciones integrales horizontales y verticales

Las proyecciones integrales horizontales y verticales del ojo, se asemejan a una campana de Gauss, como se indica en la figura 4.17.a. Estas señales se convolucionan con una Gaussiana patrón, de amplitud y desviación típica acordes a lo que se quiere encontrar, de modo que, el valor máximo de la convolución determina la posición del iris en la imagen, como se muestra en la figura 4.18.

Este proceso se ha revelado mucho más preciso que la búsqueda de máximos y mínimos sobre las funciones proyectivas ya que con el filtrado no siempre se eliminan todos los objetos no deseados de la imagen, habiendo ocasiones donde las cejas se introducen en la región clasificada como ojo, ocasionando medidas inexactas de su apertura.

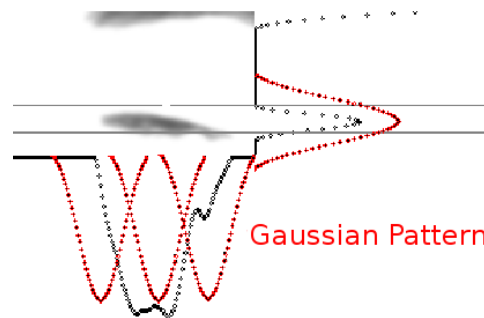


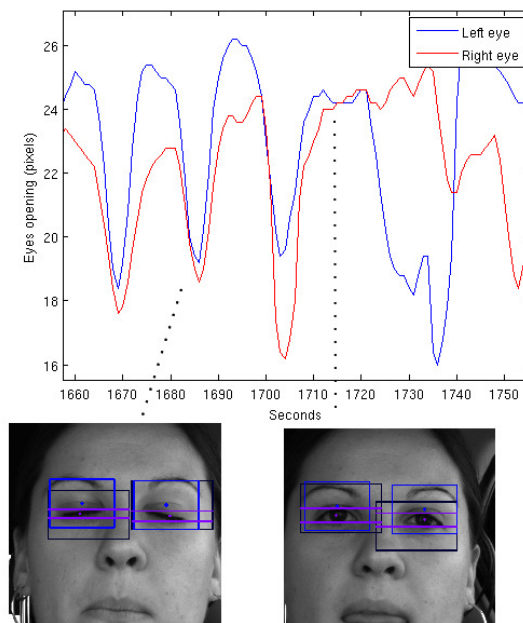
Figura 4.18: Convolución entre gaussiana patrón y proyecciones integrales

Para determinar la apertura del ojo, se busca el máximo en la curva de proyección vertical y desde él, hacia abajo, la consecución de tres ceros, que significaría que la imagen se encuentra saturada y no hay sombras, identificándose con el párpado inferior. Desde el máximo hacia arriba se busca la diferencia máxima entre niveles de gris que marcaría la localización del párpado superior. En la figura 4.17.b se ha aplicado este procedimiento. En la figura 4.17.c no se encontrarían valores nulos en la curva de proyección vertical al existir sombras por lo que la apertura se evalúa ajustando la curva próxima a la localización del ojo, a una Gaussiana dada

por la ecuación (4.5) donde x indica la posición en píxeles bien en el eje horizontal o vertical. La apertura del ojo coincide con la desviación típica de la Gaussiana.

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (4.5)$$

Del proceso anterior se obtiene la evolución temporal de la apertura del ojo, tanto en condiciones diurnas como nocturnas y la localización del iris. En la figura 4.19 se muestra un ejemplo de la evaluación temporal de la apertura del ojo para un ensayo en condiciones reales de iluminación, a partir de la cual se calcularán los parámetros de PERCLOS y frecuencia de parpadeo. En la figura 4.20, se muestra un ejemplo de la evolución de la apertura de los ojos en un parpadeo en condiciones de simulación. De esta señal se deducirá si ha existido parpadeo o no, para evaluar el número de parpadeos realizados por minuto. En la secuencia de figuras 4.20.a, 4.20.b, 4.20.c y 4.20.d se aprecia claramente la variación que sufre la señal cuando se produce un parpadeo en el conductor.



(a) Cierre de ojos

Figura 4.19: Evolución del cierre de ojos en ensayo real

4.5. Seguimiento de los ojos

Habiéndose determinado la posición del iris en una imagen, habría que repetir todo el proceso para la imagen siguiente, lo que sería computacionalmente costoso. Se ha empleado un filtro de Kalman como predictor para acotar la zona de búsqueda en la imagen siguiente. Sólo cuando se pierda definitivamente la imagen de los ojos, se repetirá el proceso en toda la imagen.

Para hacer más robusto el seguimiento, se ha implementado una máquina de tres estados: Medida, Predicción y Sin Medida. En el estado Medida si se tiene la posición en el instante k , se predice la posición para el instante $k+1$ mediante el filtro de Kalman, pasando al estado Predicción. Desde este estado, si se recibe una nueva medida se pasa al estado de Medida y se repite el proceso, pero si no se reciben medidas en cinco frames consecutivos (equivalente a 0.15

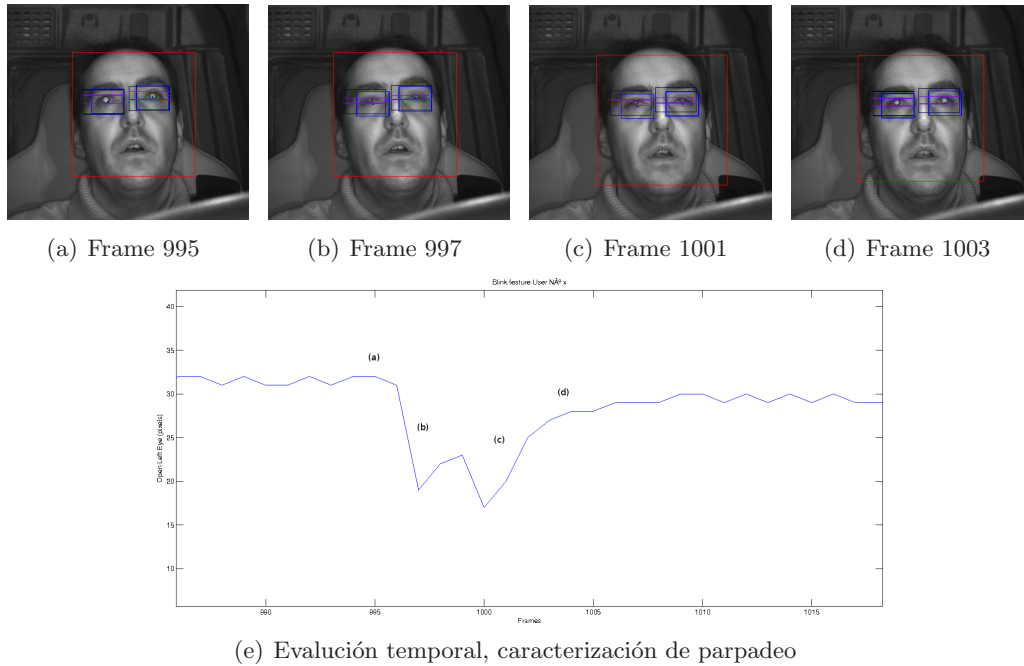


Figura 4.20: Evolución del cierre de ojos en ensayo de simulación

segundos) los ojos se dan por perdidos y se pasa al estado de Sin Medida iniciándose todo el proceso de búsqueda de los ojos. Mientras se está en el estado de Predicción se toma la posición estimada por el filtro de Kalman que suele ser válida durante unos pocos periodos de muestreo (se han tomado cinco como límite). El diagrama de estados se representa en la figura 4.21.

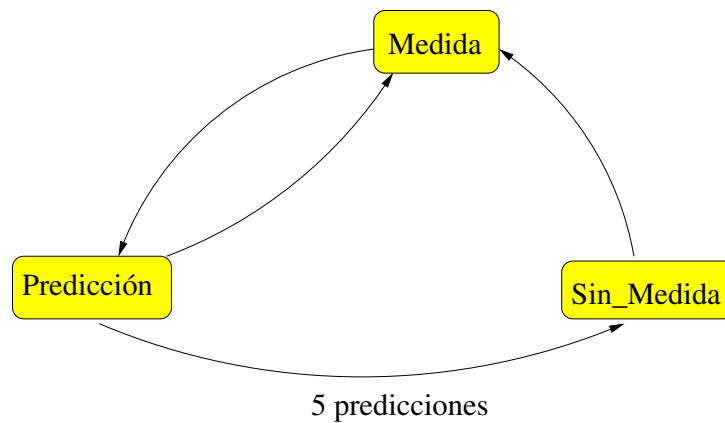


Figura 4.21: Máquina de estados aplicada al seguimiento de ojos

4.5.1. Filtro de Kalman como predictor

Estando en el periodo de muestreo k y habiendo realizado la medida en ese instante $\mathbf{y}(k)$, el filtro de Kalman, como predictor, obtiene:

$$\hat{\mathbf{x}}(k+1) \text{ minimizando } \mathbf{P}(k+1) = E\left[\left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1)\right]\left[\mathbf{x}(k+1) - \hat{\mathbf{x}}(k+1)\right]^T\right].$$

El objetivo es minimizar el valor medio de la matriz de covarianza del error $\mathbf{P}(k+1)$, entre el vector de estado $\mathbf{x}(k+1)$ y el vector de estado estimado $\hat{\mathbf{x}}(k+1)$

Habiendo medido en el instante k , el filtro proporciona la estimación del vector de estado para el siguiente instante de muestreo. Este es el modo más interesante de funcionamiento para el seguimiento de objetos en visión artificial. Anticipa donde va a estar situado, en el próximo periodo de muestreo, el objeto que se intenta seguir.

En el caso del seguimiento del ojo, no hay entradas que actúen sobre él, por lo que se trata de un sistema libre cuya representación en variables de estado viene dada por las ecuaciones (4.6), siendo \mathbf{A} la matriz del sistema, \mathbf{C} la matriz de salida o de sensores, $\mathbf{w}(k)$ el ruido de medida y $\boldsymbol{\nu}(k)$ el ruido del sistema.

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \boldsymbol{\nu}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{w}(k)\end{aligned}\quad (4.6)$$

La predicción dada por el filtro corresponde a la ecuación 4.7.

$$\begin{aligned}\hat{\mathbf{x}}(k+1) &= \mathbf{A}\hat{\mathbf{x}}(k) + \mathbf{G}(k)[\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k)] \\ \mathbf{G}(k) &= \mathbf{A}\mathbf{P}(k)\mathbf{C}^T[\mathbf{R} + \mathbf{C}\mathbf{P}(k)\mathbf{C}^T]^{-1} \\ \mathbf{P}(k+1) &= \mathbf{Q} + [\mathbf{A} - \mathbf{G}(k)\mathbf{C}]\mathbf{P}(k)\mathbf{A}^T\end{aligned}\quad (4.7)$$

En donde \mathbf{R} y \mathbf{Q} son las matrices de covarianza de los ruidos blancos de medida (salida) $\mathbf{w}(k)$ y del sistema $\boldsymbol{\nu}(k)$, respectivamente y $\mathbf{G}(k)$ la ganancia del filtro de Kalman.

4.5.1.1. Modelo de movimiento

Cuando se intenta seguir un punto (concretamente la pupila del ojo) que se está moviendo en una secuencia de imágenes, no existe posibilidad de actuar sobre la entrada del sistema, como ya se dijo, es un sistema libre, sin entrada $\mathbf{u}(k)=0$.

Como el periodo de muestreo es pequeño, puede considerarse que durante un periodo de muestreo la velocidad permanece prácticamente constante, lo que lleva a un modelo de desplazamiento del punto a seguir, lineal y uniforme. Esta consideración que es la empleada en la referencia [Ji & Yang, 2002a] y por otros autores, es válida en un periodo de muestreo. Es decir, siempre que se tengan medidas en el instante k , se puede tener una buena predicción de donde se encontrará el punto en el instante $k+1$.

Si una medida se interrumpe, puede tomarse la que da el modelo, que seguirá una línea recta con velocidad constante, pero esa posición no corresponderá con la del movimiento real de la pupila del ojo. En pocos periodos de muestreo el error puede que no sea muy elevado, pero crecerá con el tiempo y la estimación dada por el filtro de Kalman será totalmente inútil. Es decir, el modelo lineal y uniforme es bueno y dará estimaciones fiables siempre y cuando no se pierdan demasiadas medidas (se ha tomado el límite de 5).

Llamando \mathbf{x} y \mathbf{v} a los vectores de posición y velocidad respectivamente, y teniendo en cuenta que se trata de un movimiento sobre el plano, las ecuaciones del modelo serán:

$$\begin{aligned}\mathbf{v}(k+1) &= \mathbf{v}(k) \\ \mathbf{x}(k+1) &= \mathbf{x}(k) + \mathbf{v}(k)T\end{aligned}\quad (4.8)$$

Si se mide la posición en píxeles y la velocidad en píxeles que se ha movido el punto a seguir en cada frame consecutivo, el periodo de muestreo puede tomarse $T=1$ frame, por lo que las ecuaciones descomponiéndolas en sus dos componentes resultan:

$$\begin{aligned}v_u(k+1) &= v_u(k) \\ v_v(k+1) &= v_v(k) \\ u(k+1) &= u(k) + v_u(k) \\ v(k+1) &= v(k) + v_v(k)\end{aligned}\quad (4.9)$$

y representándolas en variables de estado:

$$\begin{bmatrix} u(k+1) \\ v(k+1) \\ v_u(k+1) \\ v_v(k+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u(k) \\ v(k) \\ v_u(k) \\ v_v(k) \end{bmatrix} \quad (4.10)$$

El modelo anterior puede transformarse para adquirir las medidas más fácilmente. Si se tiene en cuenta que:

$$\mathbf{v}(k)T = \mathbf{x}(k) - \mathbf{x}(k-1) \quad (4.11)$$

sustituyendo en las ecuaciones 4.8 se obtiene:

$$\mathbf{x}(k+1) = 2\mathbf{x}(k) - \mathbf{x}(k-1) \quad (4.12)$$

y representándola en variables de estado:

$$\begin{bmatrix} u(k+1) \\ v(k+1) \\ u(k) \\ v(k) \end{bmatrix} = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 2 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} u(k) \\ v(k) \\ u(k-1) \\ v(k-1) \end{bmatrix} \quad (4.13)$$

No es necesario utilizar al filtro de Kalman como observador. Todas las componentes del vector de estado pueden ser medidas por lo que la matriz de salida \mathbf{C} es igual a la matriz identidad de orden cuatro. Las únicas medidas que hay que ir haciendo y guardando son las coordenadas u y v de la pupila, en el instante actual y en el anterior. Las matrices del modelo calculado serían:

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & -1 & 0 \\ 0 & 2 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \quad \mathbf{B} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \mathbf{C} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

4.5.1.2. Parámetros e inicialización del filtro de Kalman

En sistemas libres, la ecuación de estado viene dada por (4.14):

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \boldsymbol{\nu}(k) \quad (4.14)$$

y el ruido del sistema, o de modelado, será la diferencia entre el verdadero vector de estado en el instante $k+1$ y el valor que daría el modelo $\mathbf{A}\mathbf{x}(k)$, según se desprende de la ecuación (4.15):

$$\boldsymbol{\nu}(k) = \mathbf{x}(k+1) - \mathbf{A}\mathbf{x}(k) \quad (4.15)$$

Este ruido se ha obtenido marcando, exactamente, la posición de la pupila en una secuencia de imágenes y restándole los valores que proporcionaría el modelo $\mathbf{A}\mathbf{x}(k)$. La representación de una de sus componentes se encuentra en la figura 4.22.

El ruido es blanco, de media nula, y su matriz de covarianza es de valor:

$$\mathbf{Q} = \begin{bmatrix} 8 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 8 \end{bmatrix}$$

Respecto al ruido de medida, el error típico que se comete es de 4 pixels en cada una de las direcciones u y v , por lo que la matriz de covarianza del ruido de medida será:

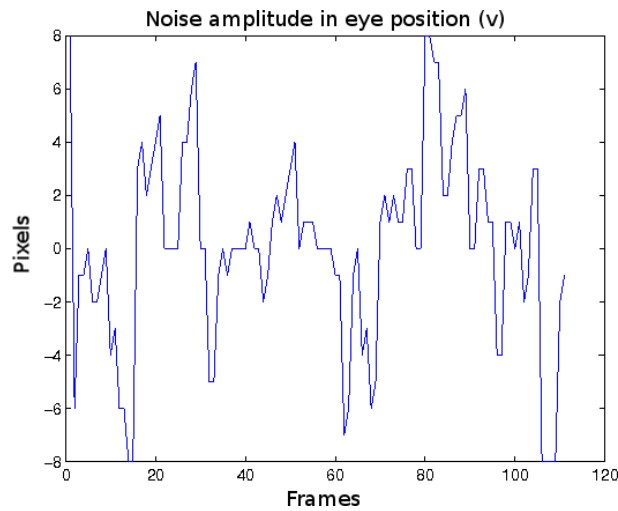


Figura 4.22: Ruido del sistema

$$\mathbf{R} = \begin{bmatrix} 16 & 0 & 0 & 0 \\ 0 & 16 & 0 & 0 \\ 0 & 0 & 16 & 0 \\ 0 & 0 & 0 & 16 \end{bmatrix}$$

Estas matrices se consideran estacionarias, es decir constantes. También es necesario inicializar la matriz de covarianza del error, pero como ha de converger pasado el régimen transitorio y su valor no es crítico, se ha tomado la matriz nula. Con estos valores iniciales, el algoritmo del filtro y su diagrama de bloques se indican en las figuras 4.23.a y 4.23.b.

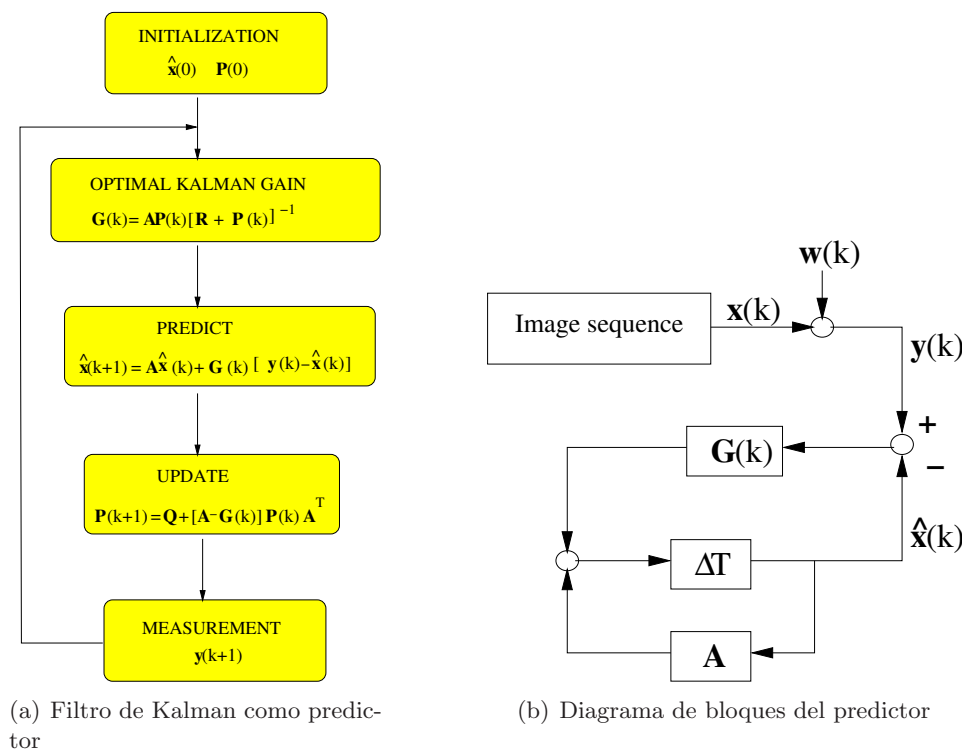


Figura 4.23: Filtro de Kalman

En la figura 4.24 se representa la coordenada v de la posición de la imagen de un ojo. En azul la posición real, en cian la medida afectada de ruido y en rojo la posición estimada por el predictor. La posición real se ha obtenido manualmente marcando con el cursor el centro del iris del ojo en todas las imágenes de una secuencia. La medida con ruido es la que suministra el sistema de visión y la posición estimada es la que adelanta el filtro de Kalman como estimación óptima de la posición real. Puede apreciarse que la estimación se acerca al verdadero vector de estado pero hay que tener en cuenta que se dispone de las medidas en todos los periodos de muestreo. Si falta alguna de las medidas la estimación empeora y con pocos periodos de muestreo perdidos, la predicción deja de ser útil.

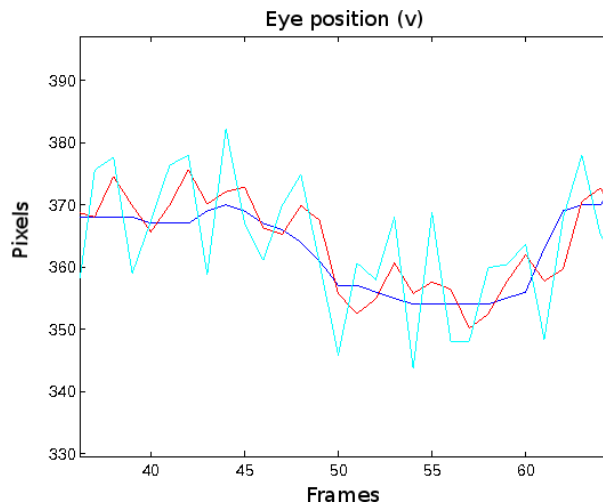


Figura 4.24: Coordenada v de la posición del ojo. Azul:real, Cian: medida con ruido y Rojo: estimación del FK.

4.6. Definición y cálculo práctico de parámetros visuales de interés

El PERCLOS y la frecuencia de parpadeo han sido ampliamente utilizados como indicadores de somnolencia pero no ha ocurrido lo mismo para el caso de la mirada perdida que también será un parámetro considerado en esta tesis. Cuando la somnolencia hace acto de presencia la frecuencia del parpadeo aumenta y el tiempo que los ojos permanecen semicerrados también aumenta (PERCLOS). Todos estos fenómenos son indicios claros de somnolencia. A continuación se van a explicar cada uno de ellos.

■ PERCLOS

Es el porcentaje de ojo cerrado durante un determinado periodo de tiempo. A la vista de la figura 4.25, los instantes t_1 y t_2 corresponden al paso de ojo abierto a ojo cerrado, mientras que los instantes t_3 y t_4 al paso de ojo cerrado a abierto. En el proceso de cierre de los ojos: t_1 es el instante en el que el ojo pasa a estar abierto menos del 80 %, y t_2 cuando el ojo se encuentra abierto menos del 20 %. En el proceso de apertura de los ojos: t_3 cuando pasa a estar más del 20 % abierto y t_4 cuando pasa a estar más del 80 % abierto. El 100 % corresponde a la apertura nominal sin estar fatigado.

En la práctica, se calculará el PERCLOS siempre que los ojos estén cerrados por debajo del 80 % de su valor nominal. Para ello, en cada periodo de muestreo, cuando se cumpla la

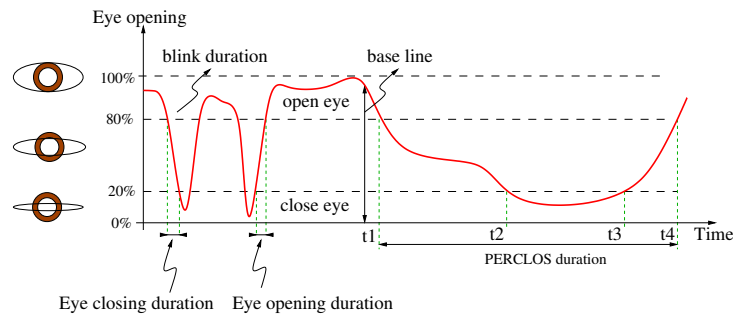


Figura 4.25: Evolución del cierre de ojos

condición anterior se asignará un 1 , y un 0 en caso contrario, como se indica en la figura 4.26. Para su evaluación se toma una ventana temporal deslizante, periodo a periodo, de 20 segundos que corresponde a 600 muestras, puesto que la frecuencia de muestreo es de 30 imágenes por segundo. Cada 1 que esté comprendido en esa ventana contribuye con un $1/600$ al valor del PERCLOS, resultando ser una función variable con el tiempo, que podrá tomar valores entre 0 y 100 %.

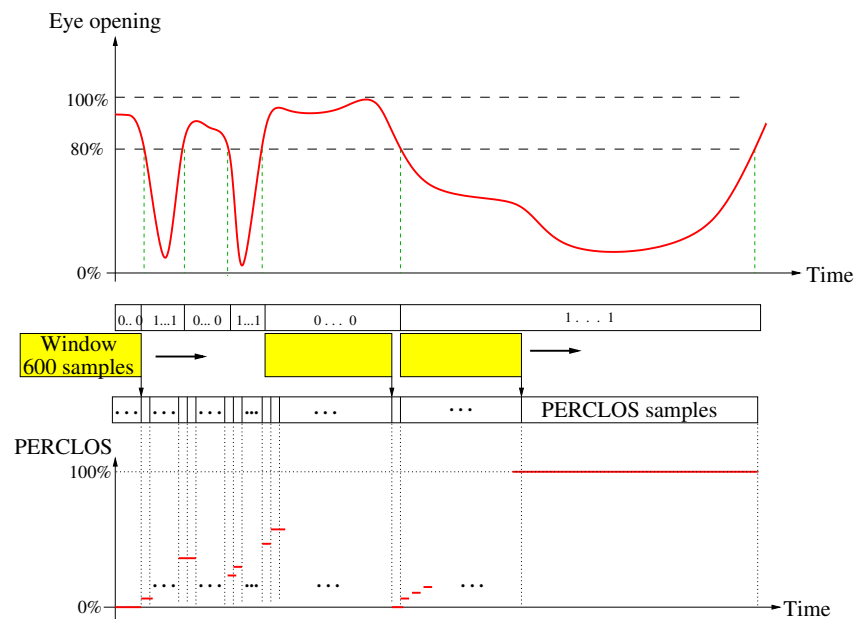


Figura 4.26: PERCLOS

El PERCLOS es evaluado para cada uno de los ojos puesto que la apertura no es la misma en ellos, siendo el valor final de PERCLOS el valor medio de ambos ojos.

■ Parpadeos

El parpadeo es el hecho de cerrar y abrir el ojo de forma rápida, su duración típica es de 200 ms, o lo que es lo mismo 6 muestras para nuestra frecuencia de muestreo. En la práctica se detectará cuando el cierre del ojo pase por el 50 %. La longitud de la ventana se tomará de 7 muestras y en caso de ser todos 1 , se considerará que es demasiado largo y correspondería a un fenómeno distinto al parpadeo. Se considerará que hay parpadeo cuando en la ventana haya de 3 a 6 unos seguidos. El proceso se indica en la figura 4.27.

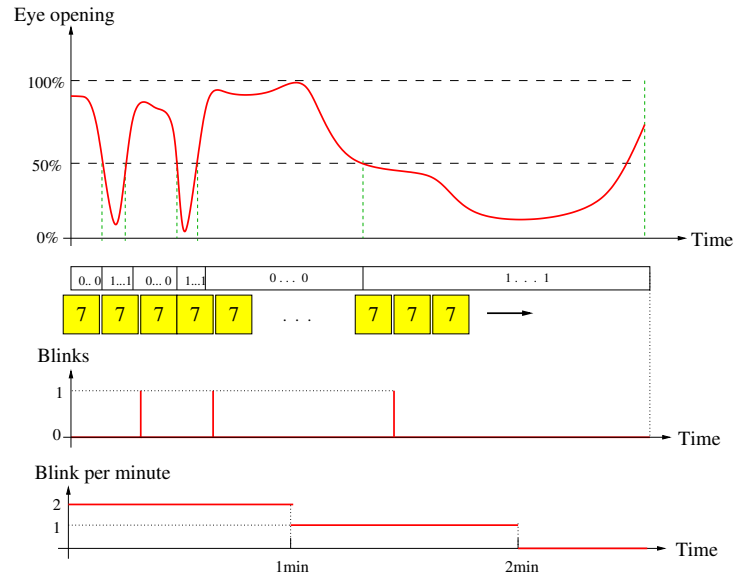


Figura 4.27: Parpadeo

En este caso el desplazamiento de la ventana no será de un periodo de muestreo sino de 7 periodos para detectar si hay parpadeo o no. Posteriormente se calculará el número de parpadeos en un minuto, obteniéndose la señal de la frecuencia de parpadeo. Los resultados obtenidos para un usuario determinado se muestran en la figura 4.28. Este usuario tiene la peculiaridad de que en el inicio del ensayo su estado es de vigilia y en el transcurso del mismo cambia a un estado de somnolencia, sin embargo, en la variable número de parpadeos realizados por minuto no se muestra cambio alguno de esta transición. De los experimentos realizados en esta tesis se ha concluido que en varios usuarios no hay relación clara entre la variación de la frecuencia de parpadeo y su nivel de fatiga. Es por ello que esta variable no será tomada en cuenta en el análisis de indicadores del conductor, siendo utilizada únicamente la señal de PERCLOS.

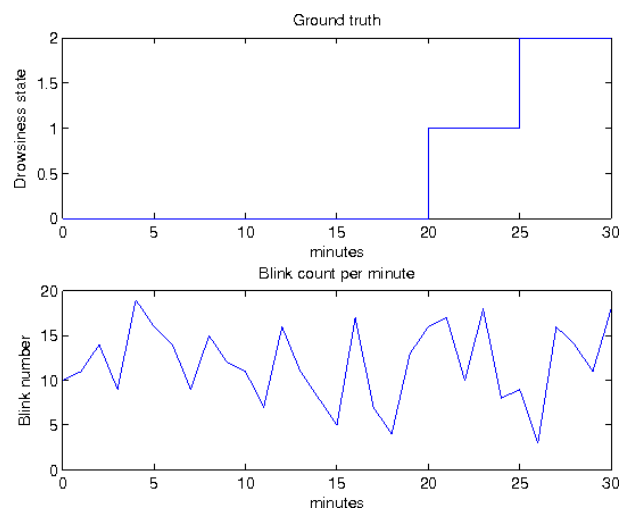


Figura 4.28: Parpadeos por minuto

En [Friedrichs & Yang, 2010b] y [Suzuki et al., 2006a] se utiliza la duración del parpadeo

para el estudio de la somnolencia. En nuestro caso esta variable no se ha analizado por estar incluida en la variable PERCLOS.

Capítulo 5

Método de medida y fusión de datos

En este capítulo, que está basado en los trabajos realizados por [Sandberg, 2011], se describen los diferentes indicadores que servirán para detectar la somnolencia, muchos de los cuales tienen parámetros que han de ser fijados de antemano y para no hacerlo de forma aleatoria se han empleado algoritmos genéticos que han permitido encontrar los valores óptimos. Cada uno de los indicadores constituye la entrada del clasificador que detectará el estado del conductor. Se lleva a cabo la fusión de varios indicadores mediante un clasificador neuronal.

5.1. Indicadores de fatiga en la conducción

La somnolencia afecta a la forma de conducir, disminuyendo la frecuencia de los pequeños movimientos correctores del volante, aumentando la frecuencia de los movimientos bruscos del mismo, incrementando la desviación del vehículo respecto a los límites del carril, aumentando la desviación típica de la posición lateral del vehículo, etc. Los movimientos del volante y la posición del vehículo en el carril son los indicadores más empleados en la detección de somnolencia [Liu et al., 2009].

El **ángulo de giro de volante** y su **velocidad de giro** son relativamente sencillos de medir desde el bus CAN del vehículo. Sin embargo, para medir la **posición lateral del vehículo** o su velocidad, se necesita un sistema *lane tracker* consistente en cámaras y algoritmos de visión para detectar los bordes del carril y calcular la distancia del vehículo al mismo. En el desarrollo de esta tesis estas variables se han obtenido como se indica en las secciones 3.1.3 y 3.2.3.

Además de los indicadores mencionados, es muy empleado el Time to Line Crossing (TLC), que es el tiempo necesario para rebasar los límites del carril sin que el conductor intervenga en los controles del vehículo, conociendo la velocidad, aceleración, dirección, posición lateral del vehículo y radio de curvatura de la carretera, como especifica [Paul et al., 2005]. De la información de TLC se derivarán nuevos indicadores como se mostrará en la sección 5.1.4.

Otro indicador evaluado en el desarrollo de esta tesis ha sido un indicador genérico de variabilidad [Sandberg et al., 2011], que por medio de procesos de optimización se pretende que esté muy correlado con el estado del conductor.

El término *indicador* anteriormente comentado es una función que reduce un segmento de una serie temporal a un valor escalar, empleando un tamaño determinado de ventana temporal para su evaluación. El deterioro en las características de conducción se cuantifica por medio de diferentes indicadores.

Por lo tanto, de las seis señales continuas de entrada del vehículo: posición y velocidad del vehículo en el carril, ángulo y velocidad de giro del volante y ángulo y velocidad de guiñada, se derivan los nuevos indicadores como se muestra en la figura 5.1.

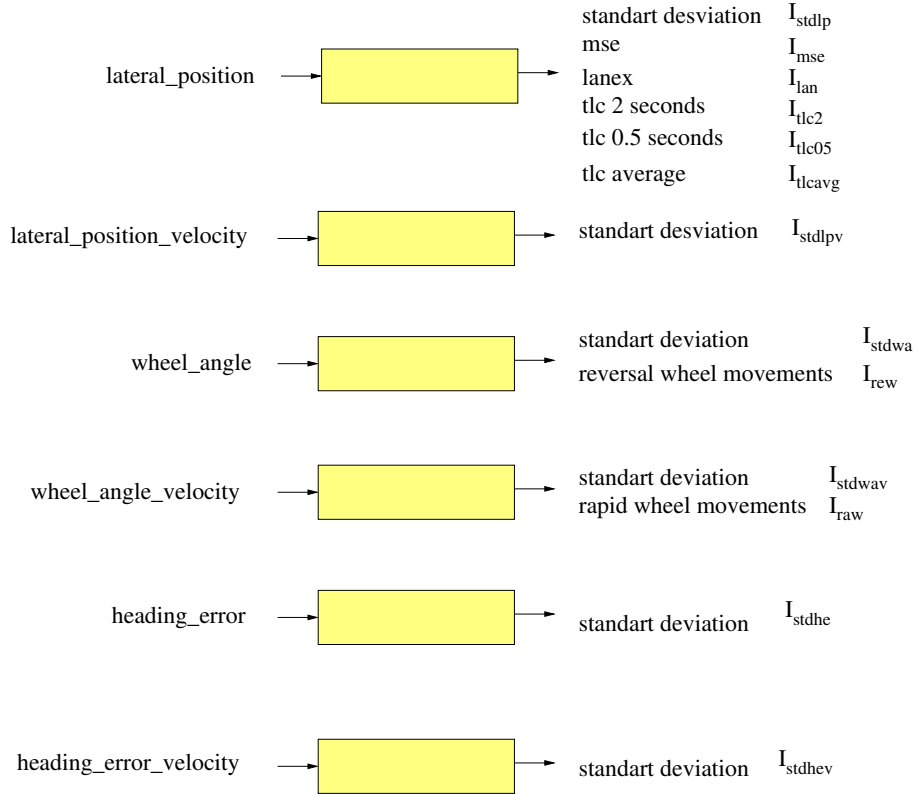


Figura 5.1: Indicadores evaluados en esta tesis

5.1.1. Desviación estándar sobre múltiples indicadores

La desviación estándar, σ , de una secuencia de datos x_i viene dada por la ecuación (5.1).

$$\sigma = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (5.1)$$

donde \bar{x} es la media aritmética de la secuencia y n el número de muestras que dependerá del tamaño de la ventana que se tome. La desviación estándar de la *posición lateral del vehículo* (I_{stdlp}), es el indicador de conducción más empleado en los estudios de somnolencia, pero también se aplica a la *velocidad lateral del vehículo* (I_{stdlpv}), al *ángulo de giro de volante* (I_{stdwa}), a la *velocidad de giro del volante* (I_{stdwav}), al *ángulo de guiñada* (I_{stdhe}) y a la *velocidad del ángulo de guiñada del vehículo* (I_{stdhev}), lo que da lugar a seis indicadores. Los dos últimos indicadores sólo serán evaluados para los ensayos de simulación.

5.1.2. Error cuadrático medio de la posición lateral

El error cuadrático medio de la posición lateral con respecto al centro del carril, genera el indicador I_{mse} , que se define igual que la desviación estándar de la ecuación (5.1) pero reemplazando el valor medio \bar{x} por el parámetro p que, en este caso, se toma el centro del carril que corresponde a 1,50m. En la figura 5.1 se ha representado este parámetro.

5.1.3. Lanex

Mide la tendencia que tiene el conductor a salirse del carril. Es la fracción de tiempo que el vehículo se encuentra fuera del carril. Se denota por I_{lan} y es evaluado, basándose en la posición

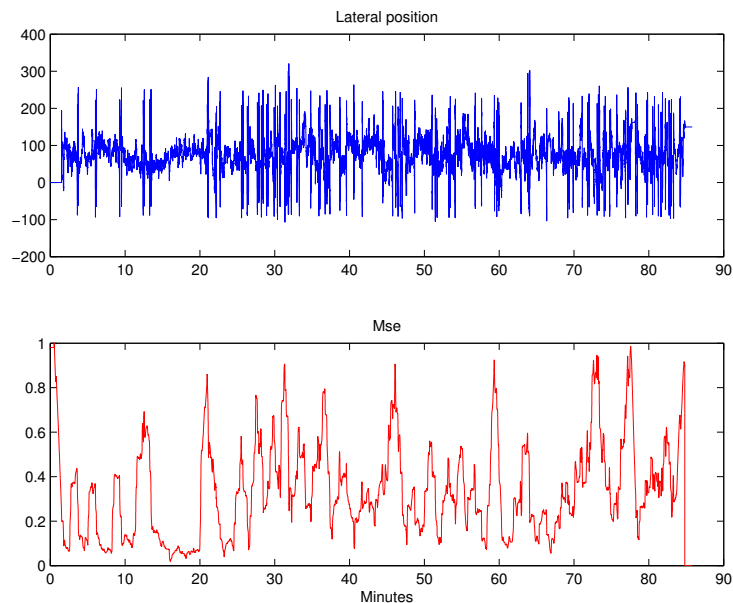


Figura 5.2: Error cuadrático medio

lateral, mediante la ecuación 5.2.

$$I_{lan} = \frac{\sum_{i=1}^n \theta(x_i)}{n} \quad (5.2)$$

donde $\theta(x_i)$ es la función dada por:

$$\theta(x_i) = \begin{cases} 1 & \text{si } x_i > x_L \\ 1 & \text{si } x_i < x_R \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.3)$$

Los parámetros $x_L = 10cm$ y $x_R = 340cm$ definen los límites izquierdo y derecho de la carretera, de tal forma que si $x_i = x_L$ las ruedas izquierdas del vehículo se encuentran tocando el límite virtual izquierdo del carril. Análogamente, cuando $x_i = x_R$, las ruedas derechas tocan el límite del carril derecho.

Es posible generalizar Lanex colocando los límites virtuales de la carretera x_L y x_R en otras posiciones dentro o fuera de la carretera. De esta forma, también se pueden incorporar tantas líneas virtuales como se quiera y la posición de estas líneas pueden estar sujetas a optimización. La generalización con múltiples líneas virtuales está dada por la ecuación 5.4.

$$I_{lang} = \frac{\sum_{i=1}^n \sum_{j=1}^m \theta_j(x_i)}{nm} \quad (5.4)$$

donde $\theta_j(x_i)$ está definido como:

$$\theta_j(x_i) = \begin{cases} \omega_{Lj} & \text{si } x_i > x_{Lj} \\ \omega_{Rj} & \text{si } x_i < x_{Rj} \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.5)$$

con m pares de líneas x_{Lj} y x_{Rj} y con pesos ω_{Lj} y ω_{Rj} tomando valores en el rango de $[0,1]$ desde $j = 1, \dots, m$. En la figura 5.3 se representa el indicador I_{lan} .

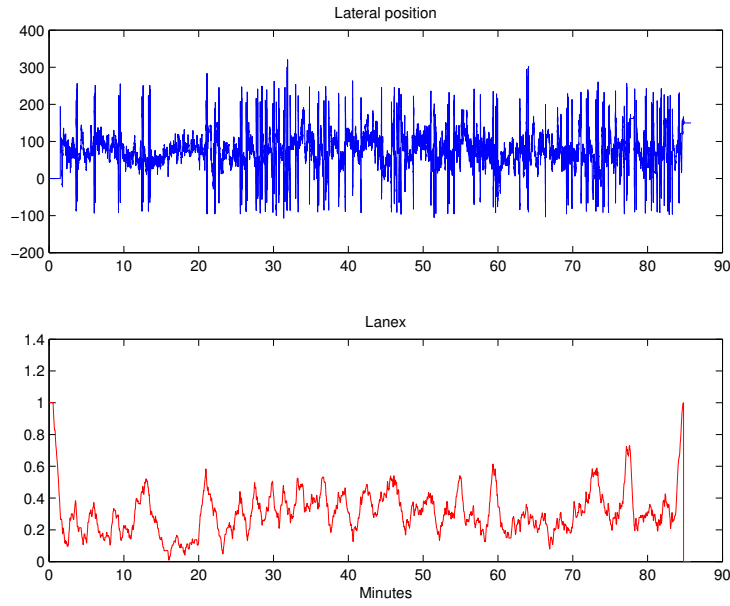


Figura 5.3: Indicador Lanex

5.1.4. Tiempo estimado en cruzar la línea del carril (TLC)

Conocidas la velocidad, aceleración, posición lateral y dirección del vehículo, y la curvatura de la carretera, con la presunción de que el conductor no realiza ninguna acción, es el tiempo necesario para cruzar algunos de los límites del carril. El TLC puede ser calculado estrictamente por trigonometría, pero la complejidad del cálculo ha llevado a realizar una aproximación como la expuesta en [van Winsum et al., 2000]. Llamando \dot{x} a la primera derivada de la posición lateral del vehículo, tal que los valores positivos corresponden a movimientos laterales hacia el límite del carril izquierdo y los valores negativos hacia el límite del carril derecho. Si la configuración del carril es la de la figura 5.4 y se dejan 50 cm de guarda a cada lado, el TLC vendrá dado por la ecuación (5.6).

$$TLC = \begin{cases} d_R/\dot{x} & \text{si } \dot{x} < 0 \\ d_L/\dot{x} & \text{si } \dot{x} > 0 \end{cases} \quad (5.6)$$

donde $d_R = -310 + x_i$ es la distancia desde el vehículo al límite derecho del carril y $d_L = -50 + x_i$ al límite izquierdo. Siempre que se esté posicionado dentro del carril, d_R tomará valores negativos mientras que d_L tomará valores positivos, por lo que el TLC será siempre positivo. El valor de TLC alcanza valores muy grandes cuando $\dot{x} = 0$ o tiene valores muy pequeños, y para evitarlo se acota a un valor máximo TLC_{max} . Si el vehículo se encuentra fuera del carril, $d_L < 0$ y $d_R > 0$ y el valor de TLC es cero.

Existen varios indicadores basados en el TLC. El número de veces que el TLC es menor de 2 segundos, en un intervalo de tiempo dado, es el indicador I_{tlc2} , que viene dado por:

$$I_{tlc2} = \sum_{i=1}^{n-1} f(TLC_{i+1}, TLC_i) \quad (5.7)$$

donde

$$f(TLC_{i+1}, TLC_i) = \begin{cases} 1 & \text{si } TLC_{i+1} < a \text{ y } TLC_i \geq a \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.8)$$



Figura 5.4: Configuración del carril

con $a=2$. El mismo indicador, pero con $a=0.5$ se denominará I_{tlc05} .

El valor medio del TLC, en un intervalo de tiempo dado, es I_{tlcavg} , que vendrá dado por:

$$I_{tlcavg} = -\frac{1}{n} \sum_{i=1}^n TLC_i \quad (5.9)$$

el signo menos se añade para que valores elevados del indicador correspondan con valores elevados de somnolencia. Todos los indicadores dependerán de los valores escogidos para TLC_{max} , d_R y d_L . En la figura 5.5 se muestra una representación particular de señal TLC.

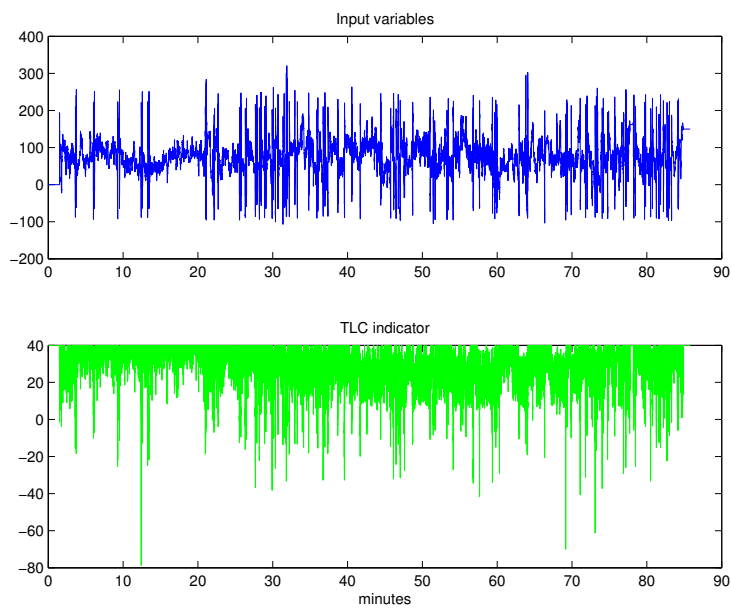


Figura 5.5: Señal de Time to Lane Crossing

5.1.5. Movimientos de volante

El indicador I_{rew} mide el número de veces que, en una ventana de tiempo determinada, el giro del volante está dentro del rango $c_1 < r < c_2$.

Se han propuesto varios rangos de evaluación: largos $c_1 = 15$ y $c_2 = \infty$, intermedios $c_1 = 5$ y $c_2 = 15$ y pequeños $c_1 = 1$ y $c_2 = 5$. Cuando el conductor está somnoliento el número de movimientos del volante pequeños es muy bajo. Los parámetros c_1 y c_2 son susceptibles de optimizarse.

5.1.6. Movimientos rápidos de volante

Mide la fracción de los movimientos del volante que exceden un umbral específico durante un intervalo de tiempo determinado y se denotará con I_{raw} . La ecuación 5.10 define al indicador.

$$I_{raw} = \frac{1}{n-1} \sum_{i=2}^n h(\dot{s}_i) \quad (5.10)$$

donde \dot{s} es la derivada del giro del volante y

$$h(\dot{s}_i) = \begin{cases} 1 & \text{si } |\dot{s}_i| > d \\ 0 & \text{en cualquier otro caso} \end{cases} \quad (5.11)$$

siendo d un umbral específico para el que Wierwille propone el valor de $d = 125^\circ/s$.

5.1.7. Ángulo de guiñada (Heading error)

El ángulo de guiñada es el formado por la tangente a la trayectoria de la carretera y el eje longitudinal del vehículo, como se indica en la figura 5.6. Sólo se dispone de él en condiciones de simulación por ser difícil de obtener en condiciones reales. La velocidad del ángulo de guiñada es la derivada del ángulo de guiñada.

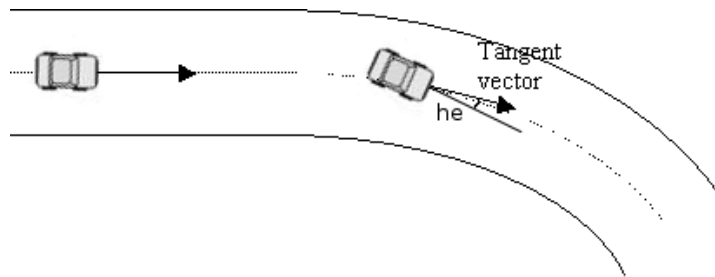


Figura 5.6: Ángulo de guiñada

5.1.8. Indicadores genéricos de variabilidad (GVI)

El indicador genérico I viene dado por la ecuación (5.12).

$$I = \frac{1}{n} \sum_{i=1}^n \omega(z_i) |z_i|^k \quad (5.12)$$

donde

$$z_i = x_i - [\delta \bar{x} + (1 - \delta)p] \quad (5.13)$$

con δ comprendido en el intervalo $[0,1]$, y p pudiendo tomar cualquier valor. Además:

$$\omega(z_i) = \frac{c_L}{1 + e^{-\alpha_L(z_i - \beta_L)}} + \frac{c_R}{1 + e^{-\alpha_R(z_i - \beta_R)}} \quad (5.14)$$

es una función peso con seis parámetros $\alpha_L, \beta_L, \alpha_R, \beta_R, c_L$ y c_R . Dependiendo de los valores que tomen los parámetros pueden obtenerse diferentes indicadores, por ejemplo con $\beta_L = \beta_R = 0, \alpha_L = \alpha_R = 0, c_L = c_R = 1, \delta = 1, \gamma k = 2$ se obtendría la desviación estandar.

La función I puede ser evaluada para cualquier señal de conducción. En esta tesis se aplica: a la posición lateral I_{glp} , a la velocidad lateral I_{glpv} , al ángulo de giro de volante I_{gwa} , a la velocidad de giro de volante I_{gwav} , al ángulo de guiñada I_{ghe} y a la velocidad del ángulo de guiñada I_{ghev} .

5.2. Optimización de indicadores

La mayoría de los indicadores anteriores dispone de parámetros que han de ser fijados. Para que la elección de los mismos no sea subjetiva, se ha empleado algoritmos genéticos para su optimización, generándose nuevos indicadores como se ilustra en la figura 5.7

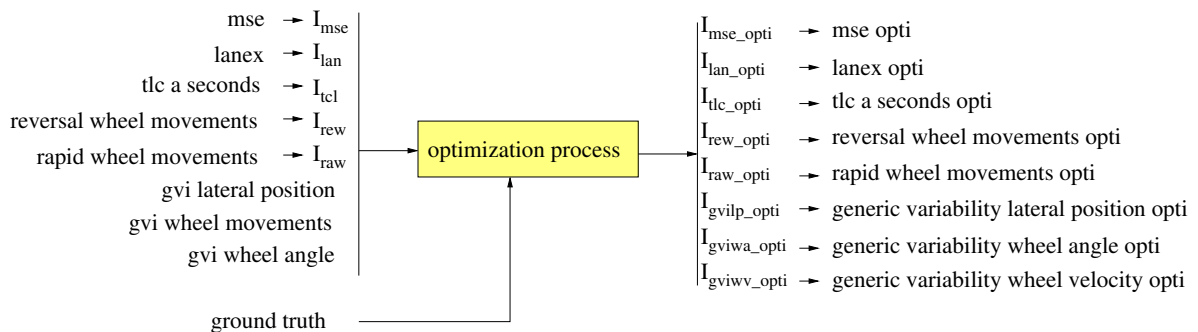


Figura 5.7: Indicadores optimizados

La optimización consiste en encontrar el valor de los parámetros de los indicadores que hace máxima una determinada función objetivo. Los métodos de optimización pueden clasificarse en clásicos, deterministas y estocásticos. La optimización clásica, basada en las derivadas, resuelve multitud de problemas pero no es aplicable a funciones objetivo no derivables como es nuestro caso; tampoco pueden aplicarse métodos determinísticos cuando se trabaja con señales aleatorias, por lo que hay que recurrir a métodos estocásticos.

5.2.1. Técnicas de optimización estocástica

Estos algoritmos de optimización, hacen uso de un operador estocástico para realizar la búsqueda del mínimo o máximo de una función objetivo. Los Algoritmos Genéticos (GA) no son simplemente una búsqueda aleatoria del mínimo sino que están basados en una población de soluciones y buscan iterativamente mejorar el conjunto de posibles soluciones.

Cuando se quiere optimizar indicadores de somnolencia, la **función objetivo** medirá el porcentaje de aciertos en la detección del estado del conductor, necesitándose comparar las soluciones del indicador con la señal de referencia del estado del conductor, por lo que será un problema de aprendizaje supervisado. Por tanto, el algoritmo de aprendizaje supervisado busca que la salida del indicador sea lo más parecida posible a la señal de referencia. La función objetivo empleada se describe en la sección 5.2.1.2 y la señal de referencia o ground truth en la sección 3.4.

En los procesos de aprendizaje supervisado hay que evitar el sobreentrenamiento, que se produce cuando el número de muestras de entrada es pequeño y el número de iteraciones elevado. Como los datos siempre están contaminados de ruido, si se produce sobreentrenamiento el sistema se adapta al ruido, sin filtrarlo. Un ejemplo hipotético de sobreentrenamiento se ilustra en la figura 5.8 en donde un entrenamiento correcto produciría la salida de color verde, mientras que el sobreentrenamiento produciría la de color rojo que se ajusta a las medidas realizadas contaminadas de ruido (puntos de color azul).

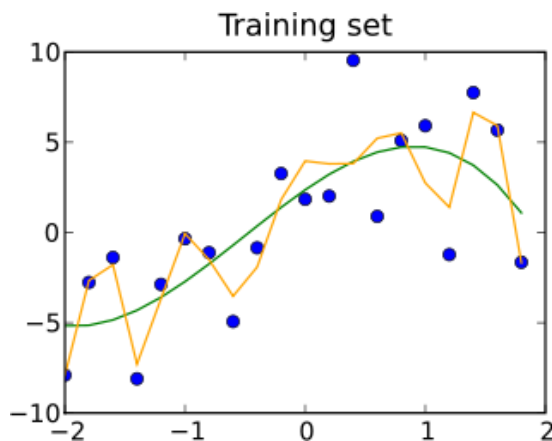


Figura 5.8: Sobreentrenamiento

Para evitar el sobreentrenamiento (*holdout validation*) se emplean conjuntos con un elevado número de muestras, de entrenamiento y de validación, de modo que al mismo tiempo que se entrena se van comprobando los resultados en un conjunto diferente de validación, que determinará cuando finaliza el entrenamiento. También es común utilizar un tercer conjunto de datos, conjunto de test, que se emplea para comprobar los resultados del parámetro optimizado sobre un conjunto de datos nuevos. En la sección 5.2.1.3 se describe cómo se han dividido los datos para los conjuntos de entrenamiento, validación y test.

5.2.1.1. Algoritmos genéticos

Los algoritmos genéticos están comprendidos en la familia de algoritmos evolutivos inspirados en la teoría de la evolución biológica de Darwin y su base genético-molecular. Se hace evolucionar una población de individuos sometiéndola a acciones aleatorias semejantes a las que actúan en la evolución biológica (mutaciones y recombinaciones genéticas), así como también a una Selección de acuerdo con algún criterio, en función del cual se decide cuáles son los individuos más adaptados, que sobreviven, y cuáles los menos aptos, que son descartados.

Los algoritmos genéticos establecen una analogía entre el conjunto de soluciones de un problema, llamado fenotipo, y el conjunto de individuos de una población natural, codificando la información de cada solución en una cadena, generalmente binaria, llamada cromosoma. Los símbolos que forman la cadena son llamados genes. Cuando la representación de los cromosomas se hace con cadenas de dígitos binarios se le conoce como genotipo. Los cromosomas evolucionan a través de iteraciones, llamadas generaciones. En cada generación, los cromosomas son evaluados usando alguna medida de aptitud. Las siguientes generaciones (nuevos cromosomas), llamada descendencia, se forman utilizando dos operadores genéticos, de sobrecruzamiento y de mutación.

El algoritmo genético puede presentar diversas variaciones, dependiendo de cómo se aplican los operadores genéticos (cruzamiento, mutación), de cómo se realiza la selección y de cómo se

5.2.1.2. Función objetivo

Para determinar la función objetivo son necesarias algunas consideraciones: primero, tener una señal de referencia (3.4) para medir la tasa de aciertos de la función objetivo y maximizar su valor y segundo, sincronizar los datos de entrada con la señal de referencia (5.2.1.3). Para medir la tasa de aciertos de la función objetivo se discretiza la señal de salida continua del indicador mediante un umbral de valor T , como se indica en la ecuación (5.15).

$$C = \begin{cases} \text{Somnolencia} & \text{si } \text{Indicador} > T \\ \text{Vigilia} & \text{en cualquier otro caso} \end{cases} \quad (5.15)$$

donde C es la clase asignada. En el proceso de optimización se evalúan todos los parámetros del indicador y el propio umbral T para que la salida sea óptima.

Si el número de elementos clasificados de una clase es mucho mayor que el de la otra, los parámetros del indicador se ajustarán para proporcionar la salida de la clase mayoritaria. Por ejemplo, si se tienen 30 muestras de las cuales 20 son clasificadas como vigilia y 10 como somnolencia, la tasa de aciertos si se clasificaran todas las muestras como somnolencia sería de $1/3$, mientras que si todas se clasificaran como vigilia sería de $2/3$. En un conjunto con el mismo número de muestras clasificadas para somnolencia y vigilia la tasa de aciertos, si se clasificaran todas las muestras en un estado, sería de $1/2$. Para evitar que la función objetivo dependa del número de elementos existentes en cada una de las clases, se utiliza la función objetivo de la ecuación (5.16) que obtendría una tasa de aciertos de $1/2$, independientemente del número de muestras, tanto si se clasifican todas como vigilia o como somnolencia.

$$f = \frac{\Gamma^{sens} + \Gamma^{spec}}{2} \quad (5.16)$$

siendo Γ^{sens} y Γ^{spec} la sensibilidad y especificidad respectivamente

Para demostrar la aseveración anterior se tendrá en cuenta que en una clasificación binaria sólo existen dos estados 0 y 1 y que existen cuatro posibilidades en la detección: Verdadero Negativo (TN) cuando siendo 0 se detecta 0 , Falso Positivo (FP) cuando siendo 0 se detecta 1 , Falso Negativo (FN) cuando siendo 1 se detecta 0 y Verdadero Positivo (TP) cuando siendo 1 se detecta 1 . Estas situaciones se reflejan en la figura 5.10.

T N $0 \rightarrow 0$	F P $0 \rightarrow 1$
F N $1 \rightarrow 0$	T P $1 \rightarrow 1$

Figura 5.10: Matriz de confusión

Cuando todas las muestras se clasifican como 0 la matriz adquiere la forma de la figura 5.11.a y si se clasifican todas como 1 tendrá la forma de la figura 5.11.b.

Teniendo en cuenta que la sensibilidad y la especificidad están definidas por la ecuación (5.17)

$$\Gamma^{sens} = \frac{TP}{TP + FN} \quad \Gamma^{spec} = \frac{TN}{TN + FP} \quad (5.17)$$

la función objetivo, aplicada al caso de clasificar todas las muestras como 0 , adquiere el valor $1/2$, igual que si se aplica al caso de clasificar todas las muestras como 1 .

T N	0
F N	0

(a) Clasificando todas las muestras a 0

0	F P
0	T P

(b) Clasificando todas las muestras a 1

Figura 5.11: Matrices de confusión

Si el número total de muestras entre usuarios es muy diferente también puede afectar a los resultados de la tasa de acierto de la función objetivo y sería necesario incluir la media como se muestra en la ecuación (5.18), donde N es el número de usuarios. En esta tesis no se ha considerado al no ser sustancial la diferencia del número de muestras entre diferentes usuarios.

$$f = \frac{1}{N} \sum_{i=1}^N f_i = \frac{1}{N} \sum_{i=1}^N \frac{\Gamma^{sens} + \Gamma^{spec}}{2} \quad (5.18)$$

5.2.1.3. Preparación de los datos

Las señales de entrada han de presentar la información a la misma frecuencia y en los mismos instantes de tiempo. En la señal de referencia se tiene una muestra cada cinco minutos que al ser tan baja no es válida para optimizar los indicadores debido al sobrentrenamiento. Por otro lado, las señales de conducción del vehículo, que como se ha comentado se registran a diferentes frecuencias, se adquieren para nuestra aplicación a una frecuencia de 10 Hz, por lo que se hace necesario interpolar la señal de referencia para igualar su frecuencia con la de las señales de conducción y tener el mismo número de muestras. El número de muestras por usuario es elevado lo que permite formar un conjunto de entrenamiento, otro de validación y un último de test, necesarios para la validación y testeo del sistema desarrollado.

En los ensayos reales, el número de muestras es de 1.080.000, lo que equivale a 1.800 minutos de conducción, y los conjuntos de entrenamiento, validación y test tienen 540.000, 270.000 y 270.000 muestras respectivamente.

En los ensayos de simulación, el número total de muestras es de 920.000, que equivalen a 1.700 minutos de conducción, y los conjuntos de entrenamiento, validación y test tienen 460.000, 230.000 y 230.000 muestras respectivamente. En la figura 5.12 se indica un ejemplo de la composición de estos grupos.

5.2.2. Parámetros de los indicadores que han sido optimizados

En el indicador del error cuadrático medio del desplazamiento en el carril, se ha optimizado el parámetro p , obteniéndose un valor de $p_{pt} = 90,54cm$ en escenarios reales.

En el indicador Lanex (con dos líneas virtuales), los valores óptimos encontrados son $x_{Lopt} = 41,67cm$ y $x_{Ropt} = 306,79cm$ que están dentro del carril y corresponden a la línea virtual izquierda situada inicialmente en $50cm$ y a la línea virtual derecha inicialmente situada en $310cm$. En caso de trabajar con cuatro u ocho líneas virtuales, los resultados obtenidos son similares a los obtenidos con parámetros sin optimizar.

En el indicador de TLC, se ha optimizado el número de veces que se encuentra por debajo de los límites determinados, obteniéndose $a_{pt} = 6,4426segundos$, frente a valores iniciales de 2 y 0.5 segundos.

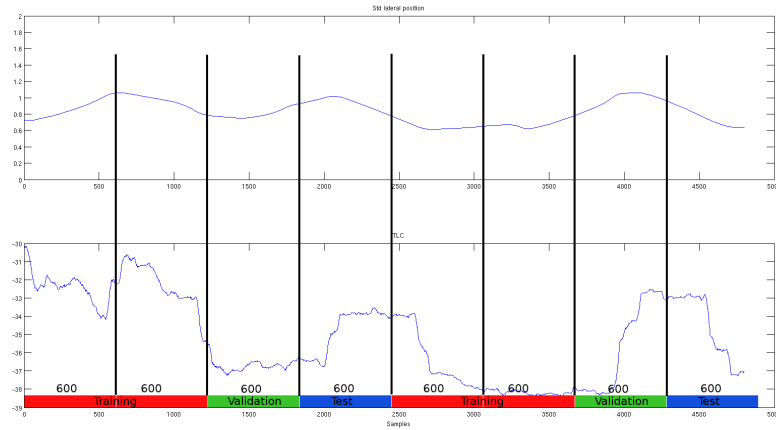


Figura 5.12: División de las muestras de entrada

En el indicador I_{rew} se han optimizado los parámetros c_1 y c_2 , obteniéndose $c_{1opt} = 0,3765^\circ$ y de $c_{2opt} = 1,0357^\circ$, los valores antes de la optimización son $c_1 = 0,9^\circ$ y $c_2 = 3,1^\circ$. En el indicador de movimientos rápidos del volante se ha optimizado el parámetro d y el resultado obtenido ha sido $d_{opt} = 10,3^\circ$ mientras que el valor utilizado sin optimizar es de $d = 20^\circ$.

Los parámetros optimizados de los indicadores basados en el índice de variabilidad genérica para los índices de posición del vehículo en el carril, ángulo de giro del volante y velocidad de giro del volante, se han expuesto en la tabla 5.1.

Tabla 5.1: Parámetros optimizados de indicadores genéricos

Indicador	GVI posición lateral	GVI ángulo volante	GVI velocidad ángulo volante
δ	0.9167	3.0687	1.3692
p	0.0349	1.3730	-1.0187
k	0.1419	2.7123	-0.8818
c_L	0.6098	-4.0468	0.9859
α_L	-0.2539	3.4330	0.2914
β_L	-1.4214	2.8663	0.8431
c_R	0.5633	1.9236	0.1808
α_R	0.4378	1.2922	1.6127
β_R	0.4725	2.8010	0.4613
T	0.6461	0.1054	0.2907

5.3. Fusión de indicadores, métodos de análisis

En la literatura científica, se utilizan cinco grandes grupos para clasificar los métodos de detección de inatención, atendiendo al criterio del tipo de medida que se realiza: medidas subjetivas de la inatención, medidas fisiológicas del conductor (descritas en 1.3.1), medidas físicas del conductor, medidas de la conducción realizada y medidas híbridas. Con la excepción del primer tipo de medida, en el resto se aplican técnicas de modelado no lineal.

El efecto de la somnolencia en el comportamiento humano no puede ser representado por un modelo lineal, por ello, las técnicas de modelado no lineal son ampliamente adoptadas en el área

de detección de fatiga en conductores. El modelado no lineal con técnicas de aprendizaje puede extraer información de datos muy ruidosos.

Las **Redes Neuronales Artificiales** (ANN) han sido estudiadas y utilizadas en numerosos campos de la ciencia y la ingeniería. Una de las principales ventajas de las redes neuronales es que infieren soluciones, de datos de los que a priori no se tiene conocimiento del patrón que siguen, o lo que es lo mismo, la red extrae empíricamente los patrones incluso si la ecuación entre las entradas y las salidas no existe. Esta característica es muy importante porque hay circunstancias en las que la relación entre la entrada y la salida del sistema no se conoce o es difícil de establecer. Las redes neuronales también tienen la posibilidad de generalizar (responden con una cierta exactitud a patrones que son similares a los patrones originales del aprendizaje), lo cual es muy útil porque los datos en el mundo real son ruidosos, distorsionados y frecuentemente incompletos. Las redes neuronales son no lineales, lo cual permite resolver algunos problemas complejos con más exactitud que las técnicas lineales [Eskandarian et al., 2007].

Los **Sistemas de Lógica Borrosa** (FIS) son muy prácticos por su concepto de modelo lingüístico. Las reglas borrosas se encuentran muy cercanas al lenguaje natural de un experto. Un sistema borroso gestiona el conocimiento dudoso e infiere el comportamiento de los datos observados. Por otro lado, la lógica borrosa puede ser usada para procesos de inducción de conocimiento [Bergasa et al., 2006].

Las **Máquinas de Soporte Virtual** (SVM) se basa en técnicas de aprendizaje estadístico y pueden ser usadas para clasificación de patrones y para la inferencia de relaciones no lineales entre variables. Estos métodos han sido satisfactoriamente aplicados en la detección, verificación y reconocimiento de caras [Liang et al., 2007]. La técnica de aprendizaje de SVM se realiza gracias a la información cognitiva de los humanos. SVM puede generar tanto modelos lineales como no lineales y es capaz de computar los modelos no lineales tan eficientemente como los lineales. Dado el conjunto de datos de entrada se cambia a un dominio transformado, y posteriormente se busca un hiper-plano que separa los datos con el error mínimo y ganancia máxima. El hiper-plano se transforma de nuevo al dominio de entradas para obtener los límites de decisión, los cuales deben ser potencialmente no lineales.

5.3.1. Técnicas de fusión con redes neuronales

Las medidas realizadas sobre el comportamiento del conductor pueden cambiar mucho entre usuarios debido a la complejidad, variabilidad y no linealidad del comportamiento humano. Establecer una dependencia entre las variables relacionadas con el conductor, las variables relacionadas con la conducción y sus indicadores asociados, y el estado en el que se encuentra el conductor sería muy complejo. Las técnicas adaptativas de aprendizaje, como las redes neuronales, pueden aprender de ejemplos y predecir la salida sin tener un conocimiento exacto de la relación entre entradas y salidas.

Un ejemplo de arquitectura de red neuronal es la mostrada en la figura 5.13.a, que está compuesta por una capa de neuronas de entrada cuyo número de neuronas, cuando sólo se emplea un indicador, es de 100, por las que se introducen 100 muestras consecutivas de la señal del indicador correspondiente, o 200 neuronas cuando se emplea fusión de dos indicadores. La arquitectura también se compone de otra capa de neuronas ocultas y finalmente una neurona de salida. El número de neuronas de la capa oculta influye en los resultados de la red, y se ha tomado igual a 10 por ser el valor que mejores resultados genera en el clasificador, como se muestra en la curva ROC (Véase la sección 6.2.1) de la figura 5.13.b, que está construída para un rango de número de neuronas de 5 a 35, en incrementos de 5. En la capa oculta, cada neurona recibe una serie de entradas a través de interconexiones y emite una salida, como se muestra en la figura 5.13.a. Esta salida viene dada por dos funciones: una *función de propagación* que por lo general consiste en el sumatorio de cada entrada multiplicada por el peso de su interconexión y

una *función de transferencia*, que acota la salida de la neurona. En la capa oculta se ha empleado la función tangente hiperbólica que proporciona valores en el intervalo $[-1,1]$ y en la de salida se ha empleado la función sigmoidea que proporciona valores en el intervalo $[0,1]$, considerándose que por debajo de 0,5 el conductor está en estado de vigilia, y en cualquier otro caso el estado es de somnolencia.

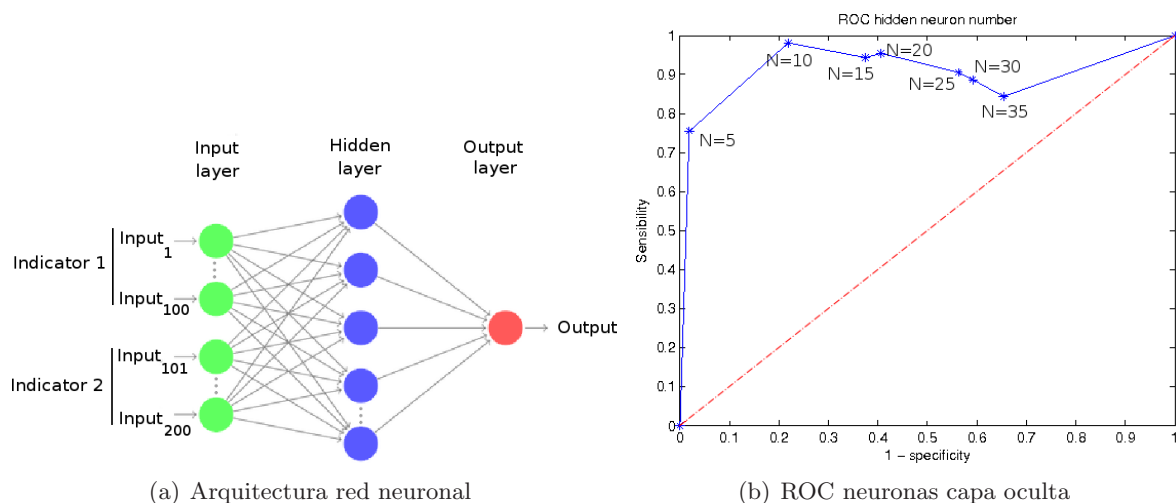


Figura 5.13: Red neuronal

El aprendizaje de la red es supervisado, es decir, necesita un conjunto de datos de entrada previamente clasificados que es la señal de referencia. La distribución de los datos de entrada es análoga a la explicada en la sección 5.2.1.3, datos de entrenamiento, validación y test. El estudio exhaustivo de diferentes topologías de ANN a utilizar para esta aplicación queda fuera de los objetivos de la tesis.

Capítulo 6

Resultados experimentales

Inicialmente, se exponen los resultados del sistema de evaluación de la apertura de los ojos por medio de secuencias de imágenes en simulación y condiciones reales, en situación de fatiga y vigilia. Posteriormente se exponen los resultados obtenidos en la detección de somnolencia, analizando la calidad de las clasificaciones mediante la matriz de confusión y los valores de la sensibilidad y especificidad. Los ensayos y los resultados alcanzados se han diferenciado dependiendo de si son en condiciones de simulación o en condiciones reales y, para los dos grupos, el clasificador ha consistido en una red neuronal cuyas entradas son muestras de la señal de un sólo indicador. Posteriormente, los indicadores se han optimizado y se han fusionado de dos en dos. Por último, se muestra una discusión sobre la comparativa de resultados de simulación y en condiciones reales.

6.1. Evaluación de la caracterización de la apertura del ojo

Para mostrar en detalle los resultados obtenidos con el sistema de evaluación de la apertura del ojo, medida fundamental para la evaluación del PERCLOS, y la relación existente con la somnolencia, se representan los resultados obtenidos para tres ensayos: simulación sin somnolencia, simulación con somnolencia y real con síntomas de fatiga.

Se han representado 25 minutos del ensayo de simulación sin somnolencia como se muestra en la figura 6.1. Durante todo el ensayo el estado del conductor es de vigilia, no mostrándose síntomas de fatiga o somnolencia ya que la apertura del ojo es próxima a la nominal como se muestra en la secuencia de imágenes de la figura 6.2.

En el ensayo de simulación con somnolencia se han representado 35 minutos, correspondiendo al mismo usuario, que se muestran en la figura 6.3, donde se aprecia un aumento del nivel del PERCLOS a partir del minuto 20 que corresponde con los apreciables síntomas de somnolencia de la secuencias de imágenes de la figura 6.4. La apertura del ojo en los primeros minutos, figuras 6.4.a y 6.4.b, corresponde a un estado de vigilia, sin embargo a medida que transcurre el tiempo se hace más acusado el cierre de ojos, como puede apreciarse en la secuencia de las figuras 6.4.d - 6.4.h, que corresponden con estados de fatiga y somnolencia del conductor.

Del ensayo real se han elegido dos sesiones donde existe fatiga, como se muestra en la figura 6.5. La fatiga aparece entre los minutos 110 y 140 del ensayo como puede apreciarse en la secuencia de imágenes 6.5.d - 6.5.g, en las que la apertura del ojo es bastante inferior a la de las imágenes 6.5.a y 6.5.b. En la figura 6.5.h, el conductor aparenta menos cansancio y corresponde a la fase final del ensayo.

A continuación se han elegido tres secuencias de imágenes que corresponde a un giro de cabeza, un deslumbramiento producido por el sol o por algún objeto de la vía y un paso por túnel que oscurece las imágenes para mostrar los resultados obtenidos por el sistema de evaluación de la apertura del ojo en condiciones especialmente difíciles de funcionamiento.

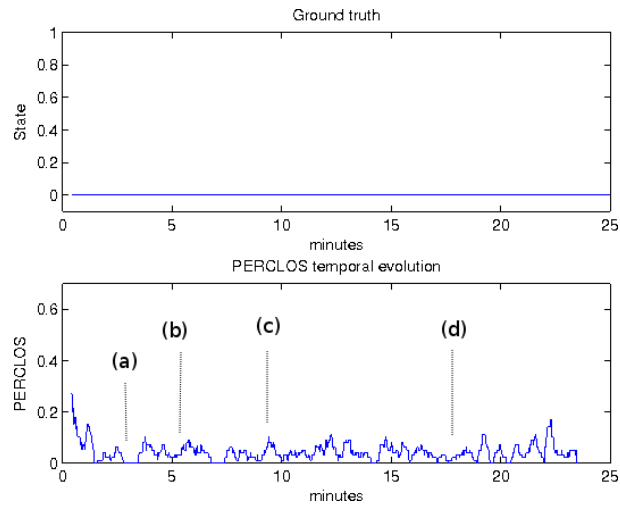


Figura 6.1: Evolución temporal del PERCLOS en ensayo de vigilia

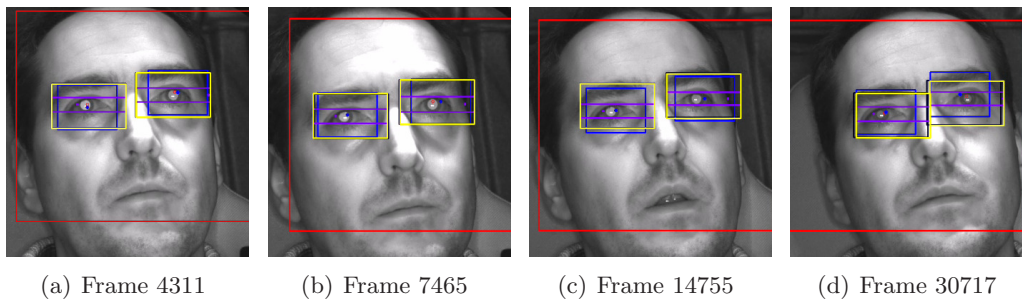


Figura 6.2: Secuencia de imágenes de ensayo de simulación sin somnolencia

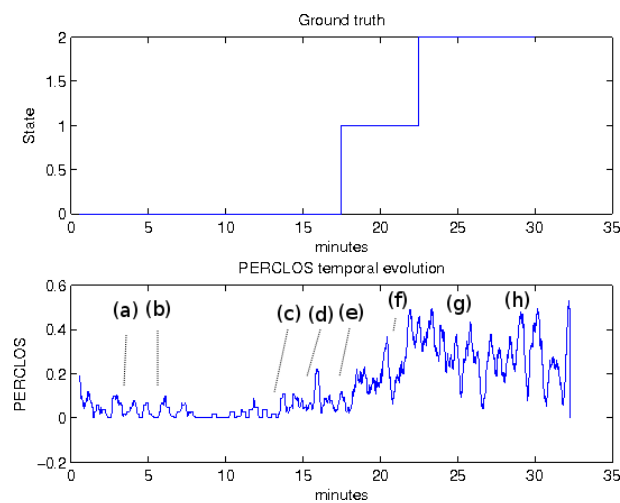


Figura 6.3: Evolución temporal del PERCLOS en ensayo de somnolencia

En la primera secuencia, se muestra como se detectan los ojos cuando gira la cabeza y la medida de la apertura de los mismos, logrando que el PERCLOS obtenido refleje el estado del conductor. En las primeras imágenes, 6.7.a - c, el conductor inicia el giro de cabeza, mantenién-

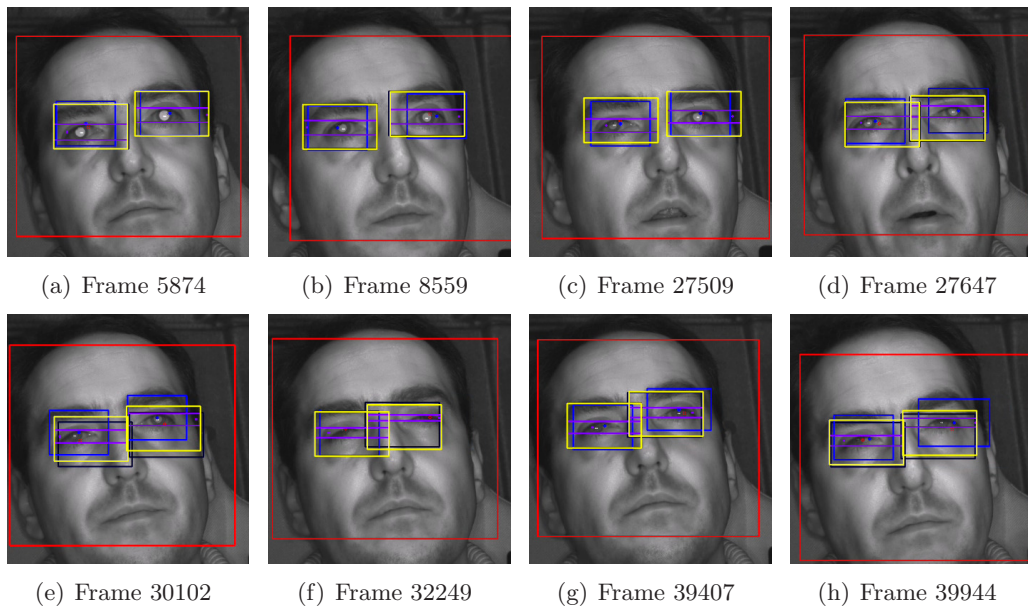


Figura 6.4: Secuencia de imágenes de ensayo de simulación con somnolencia

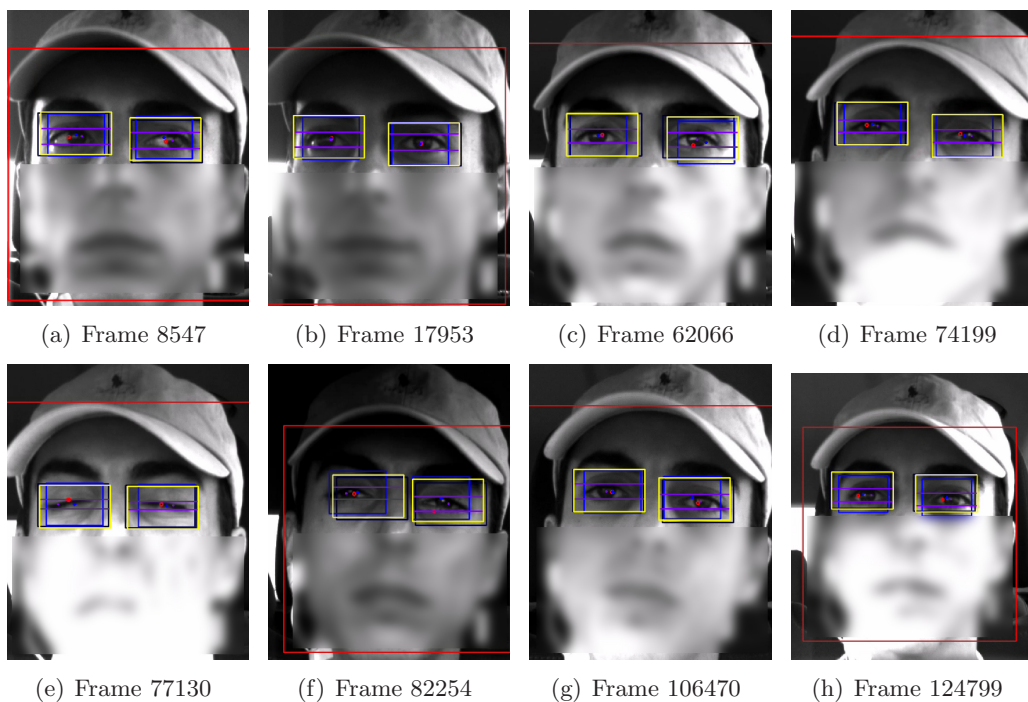


Figura 6.5: Secuencia de imágenes para ensayo real

dose con la cabeza girada en las imágenes 6.7.d - g y regresando a la postura de mirada al frente en la figura 6.7.h. En la imagen 6.7.e se detecta un error en la evaluación de la apertura, mientras que en el resto de imágenes la medida de la apertura es correcta.

En la segunda secuencia (figura 6.8), se muestra como se detectan los ojos y se evalúa la apertura cuando existe saturación a consecuencia del sol u objetos de la vía. La robustez del sistema puede ser verificada en la transición encontrada entre las imágenes 6.8.b y 6.8.c, donde

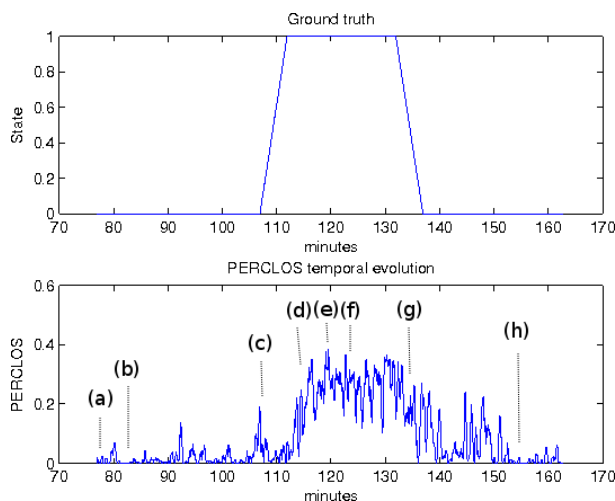


Figura 6.6: Evolución temporal del PERCLOS en ensayo real

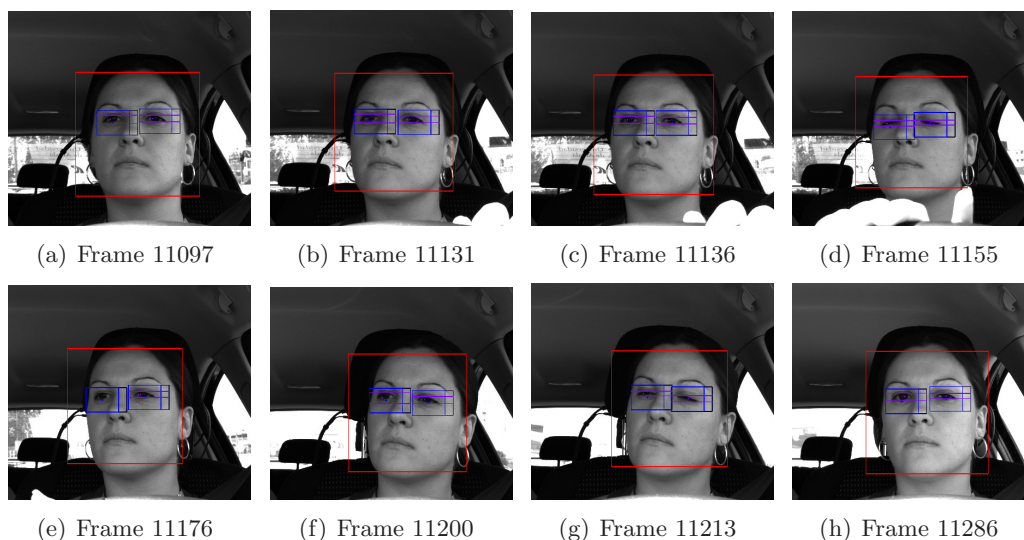


Figura 6.7: Caracterización en un movimiento de giro de cabeza

se aprecia un gran cambio en la luminosidad de la imagen y una correcta caracterización de la apertura de los ojos.

La tercera secuencia (figura 6.9), corresponde al paso del vehículo por un túnel, haciendo que la imagen se oscurezca y se pierdan detalles en ella, sin embargo el sistema de evaluación de apertura se ajusta al cambio de iluminación, como se muestra en las imágenes 6.9.c y 6.9.h.

6.2. Metodología de evaluación del funcionamiento del sistema

Los procedimientos para evaluar la calidad de los resultados obtenidos son diferentes si la clasificación es binaria o terciaria. Se exponen, en esta sección, los dos métodos, aunque se concluya que al clasificar en tres estados, los resultados no son suficientemente buenos y por lo tanto la propuesta final es implementar un clasificador con dos estados.

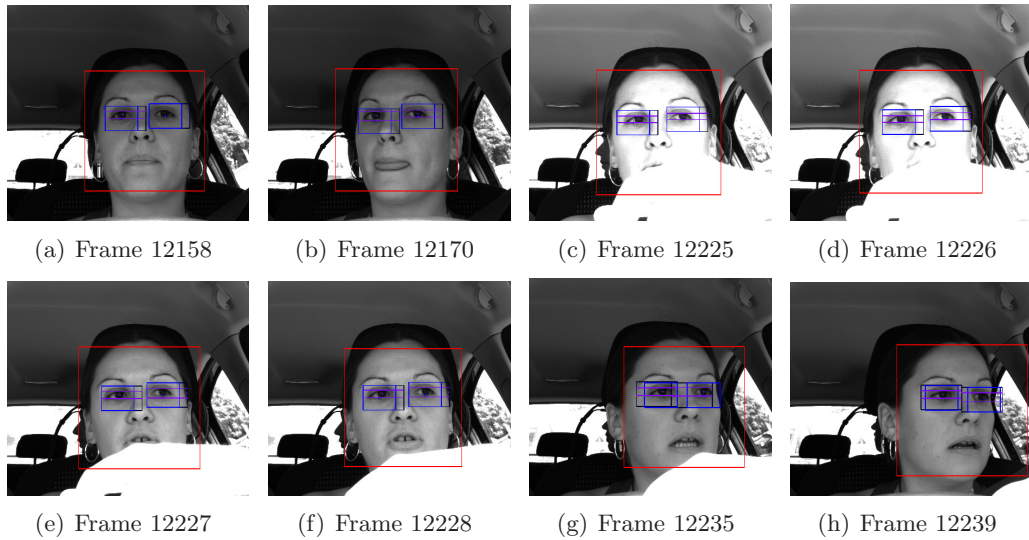


Figura 6.8: Caracterización tras el reflejo del sol

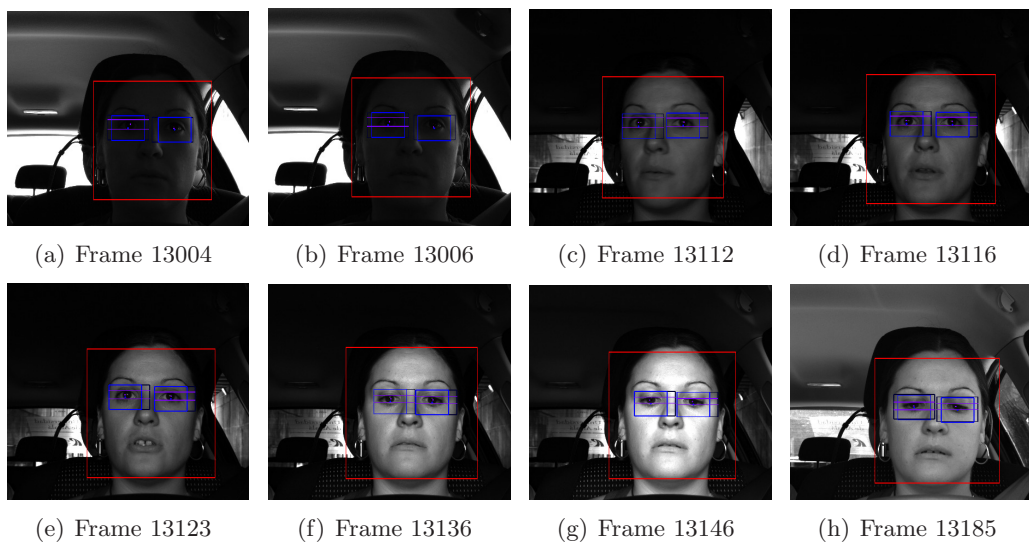


Figura 6.9: Caracterización tras la salida de un tunel

6.2.1. Evaluación de un clasificador binario

En el caso de clasificadores binarios existen dos estados de salida: vigilia (0) y somnolencia (1). Comparando la salida del clasificador con la señal de referencia se pueden dar las cuatro situaciones indicadas en la tabla 6.1.

Tabla 6.1: Matriz de confusión binaria

Señal de referencia \ Clasificador	Vigilia (0)	Somnolencia (1)
Vigilia (0)	Verdadero Negativo (TN) 0->0	Falso Negativo (FN) 0->1
Somnolencia (1)	Falso Positivo (FP) 1->0	Verdadero Positivo (TP) 1->1

- Verdadero Positivo (TP): Cuando el resultado de la clasificación es un estado de somnolencia y coincide con el estado que indica la señal de referencia.
- Falso Positivo (FP): Cuando el resultado de la clasificación es un estado de somnolencia y la señal de referencia indica que es un estado de vigilia.
- Verdadero Negativo (TN): Cuando el resultado de la clasificación es un estado de vigilia y coincide con el estado que indica la señal de referencia.
- Falso Negativo (FN): Cuando el resultado de la clasificación es un estado de vigilia y la señal de referencia indica que es un estado de somnolencia.

Se había definido la **sensibilidad** como la proporción de Verdaderos Positivos que son clasificados correctamente, y la **especificidad** como la proporción de Verdaderos Negativos que son clasificados correctamente. Sus expresiones vienen dadas en la ecuación (5.17).

Si el valor de la sensibilidad fuese la unidad, entonces no habría falsos negativos, es decir, todos los 0 (vigilia) estarían bien clasificados. Si la especificidad fuese la unidad indicaría que no habría falsos positivos, o lo que es lo mismo, todos los 1 (somnolencia) estarían bien clasificados. Lo ideal sería que tanto la sensibilidad como la especificidad fueran la unidad.

Se define el espacio ROC (Receiver Operating Characteristic) como el plano cuyo eje de ordenadas es la sensibilidad y el eje de abscisas 1-especificidad. A cada matriz de confusión le corresponde un punto en el espacio ROC. Para un mismo conjunto de N datos de entrada, diferentes clasificadores proporcionan matrices de confusión diferentes y la representación en el espacio ROC de todas ellas forman la curva ROC. El clasificador óptimo es aquel que no tiene falsos negativos (sensibilidad=1) ni falsos positivos (especificidad=1) y que correspondería con el punto (0,1) del espacio ROC, mientras que el peor clasificador sería el que tuviera la sensibilidad y la especificidad iguales a cero y que correspondería con el punto (1,0). Un clasificador aleatorio (inservible) daría el mismo número de ceros que de unos y en ambos casos la mitad serían verdaderos y la otra mitad falsos. Su matriz de confusión estaría rellena en cada uno de sus elementos con $N/4$ muestras, representándose en el punto (1/2, 1/2) del espacio ROC.

Los clasificadores válidos han de estar representados por encima de la diagonal (0,0) (1,1) y lo más cercanos posible al punto (0,1), como se representa gráficamente en la figura 6.10.

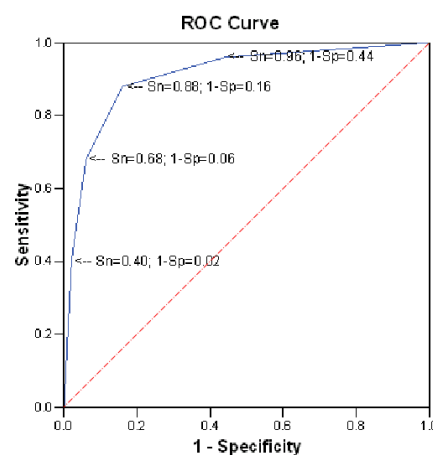


Figura 6.10: Curva ROC

La diferencia entre la señal de referencia y la salida del clasificador es la señal de error. Su valor cuadrático medio indica la *tasa de error* por lo que la tasa de aciertos será $1 - \text{tasa de error}$.

6.2.2. Evaluación de un clasificador terciario

En el caso de clasificadores ternarios existen tres estados de salida: vigilia, fatiga y somnolencia y las curvas ROC de sensibilidad y especificidad no pueden aplicarse directamente. En [Patel & Markey, 2005] se han evaluado la bondad de clasificadores ternarios, mediante el método de tela de araña. En general, un problema de clasificación de N clases puede ser representado por una matriz de confusión de $N \times N$ elementos, de los que $N^*(N-1)$ corresponderán a elementos mal clasificados.

La interpretación de la matriz de confusión ternaria se representa en la tabla 6.2 donde, por ejemplo, la nomenclatura Vigilia(0) \rightarrow Somnolencia(2) es el porcentaje de estados de Vigilia(0) que se han clasificado como Somnolencia(2).

Tabla 6.2: Matriz de confusión de tres estados

Señal de referencia \ Clasificador	Vigilia(0)	Fatiga(1)	Somnolencia(2)
Vigilia(0)	Vig(0) \rightarrow Vig(0)	Vig(0) \rightarrow Fat(1)	Vig(0) \rightarrow Som(2)
Fatiga(1)	Fat(1) \rightarrow Vig(0)	Fat(1) \rightarrow Fat(1)	Fat(1) \rightarrow Som(2)
Somnolencia(2)	Som(2) \rightarrow Vig(0)	Som(2) \rightarrow Fat(1)	Som(2) \rightarrow Som(2)

Obtenida la matriz de confusión, de $N^*(N-1)$ elementos, se representan los porcentajes de mal-clasificados de la matriz de confusión en un gráfico radial correspondiente a un hexágono, cuyos vértices corresponden a los valores representados en verde en la tabla 6.2. Por otro lado se representan los porcentajes de aciertos-clasificados de la matriz de confusión en un gráfico radial correspondiendo a un triángulo, cuyos vértices corresponden a los valores representados en azul por la tabla 6.2.

Un ejemplo de matriz de confusión se muestra en la tabla 6.3 en donde se han marcado en verde las seis casillas correspondientes a malas-clasificaciones, su representación de gráfico radial se representa en la figura 6.11.a. Por otro lado, en la tabla 6.3 las tres casillas representadas de color azul corresponden a los aciertos-clasificados, su representación de gráfico radial se representa en la figura 6.11.b.

Tabla 6.3: Ejemplo de matriz de confusión de tres estados

Señal de referencia \ Clasificador	Vigilia(0)	Fatiga(1)	Somnolencia(2)
Vigilia(0)	0.8400	0.3500	0.1000
Fatiga(1)	0.1300	0.3700	0.2000
Somnolencia(2)	0.0300	0.2800	0.7000

El peor caso de clasificación correspondería a una clasificación aleatoria de los tres estados y de todas las posibles confusiones con una probabilidad de $1/3$. El hexágono de clasificación aleatoria está representado en rojo. Las representaciones de un buen clasificador se acercarán a hexágonos regulares y estarán dentro del hexágono de clasificación aleatoria. Cuanto menor sea el área del hexágono, mayor será la especificidad. El triángulo de color rojo representa el polígono de clasificación aleatoria. Un buen clasificador se parecerá lo máximo posible a un triángulo equilátero y estará fuera del triángulo de clasificación aleatoria. Cuanto mayor sea el área, mayor será la sensibilidad del clasificador.

El cálculo de la matriz de confusión para cada uno de los ensayos se ha realizado en Matlab y los parámetros de entrada han sido la señal de referencia de somnolencia y la señal de salida de la red neuronal.

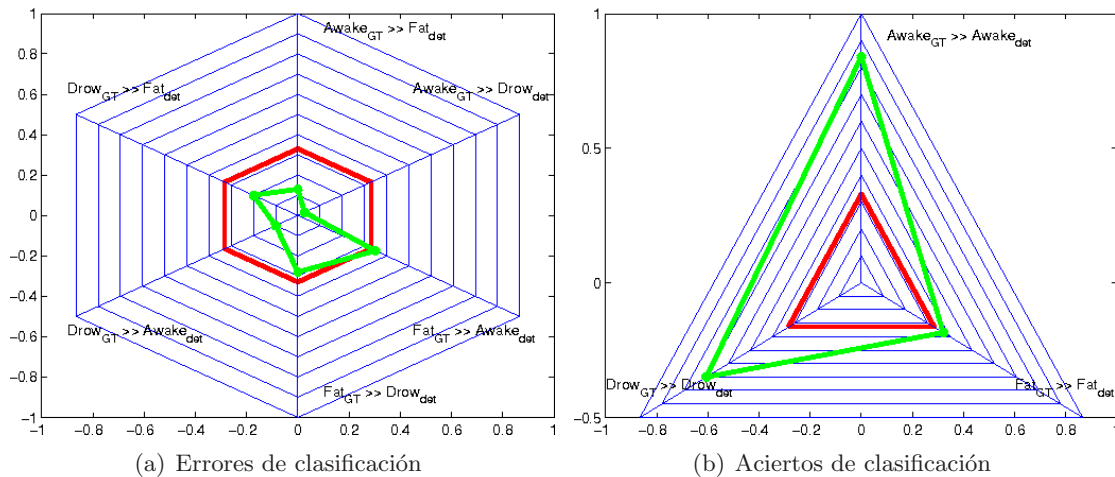


Figura 6.11: Representación gráfica de la matriz de confusión

6.3. Resultados en condiciones de simulación

En la evaluación de los resultados de esta tesis, relacionados con la simulación, se dispone aproximadamente de 900 minutos referenciados como vigilia y aproximadamente 800 minutos referenciados como fatiga, presentando algunos de ellos evidentes síntomas de somnolencia (7%), tanto en las señales e indicadores de conducción como en los parámetros faciales del conductor.

Respecto a la señal de referencia y considerando una clasificación de tres estados (vigilia, fatiga y somnolencia), el porcentaje de minutos que tiene cada uno de los estados es el mostrado en la figura 6.12.a, mientras que si se considera una clasificación de dos estados (vigilia y somnolencia), el porcentaje de minutos es el mostrado en la figura 6.12.b.

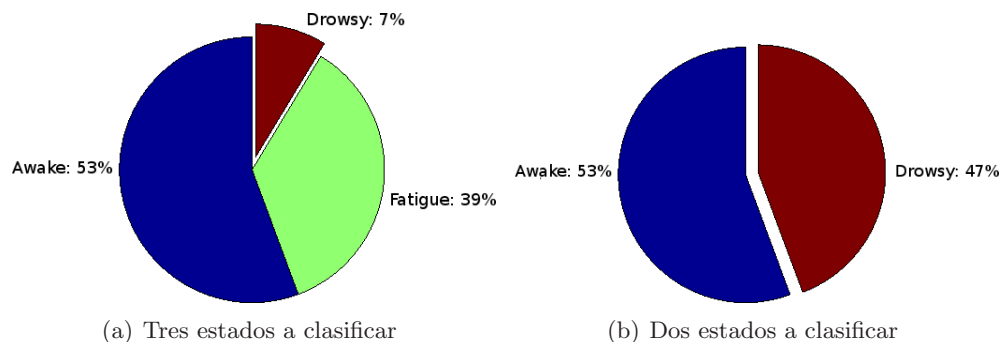


Figura 6.12: Porcentajes de clasificación del total de ensayos

En la figura 6.12.a se muestra que sólo existe un 7% de las muestras, clasificadas como somnolencia, siendo complicado que los sistemas de aprendizaje descritos en la tesis sean capaces de generalizar y obtener las características que determinan este estado. Por esta razón y dado que para el diseño de un sistema práctico es suficiente con determinar un estado de fatiga claro que permita avisar al conductor, se opta por fusionar los porcentajes de clasificación de los estados de fatiga y somnolencia, obteniendo un porcentaje del 47%, como se muestra en la figura 6.12.b. En el desarrollo de este capítulo se mostrarán resultados para ambos esquemas de clasificación.

En cada uno de los ejercicios realizados se ha evaluado, con las técnicas descritas en la sección 4.6, el porcentaje de tiempo que los ojos se encuentran cerrados (PERCLOS), obteniéndose

señales como las mostradas en la figura 6.13.

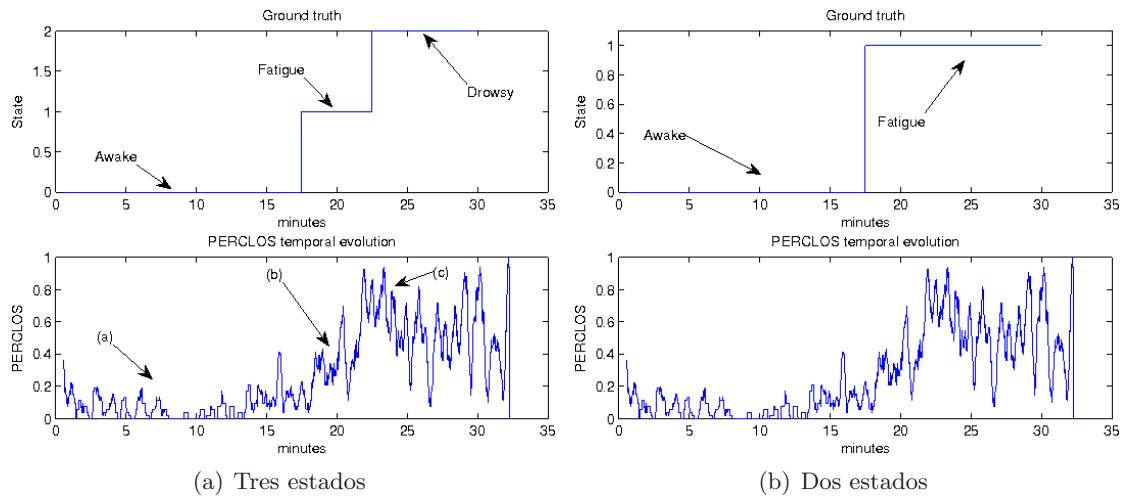


Figura 6.13: Señal de referencia y evolución temporal del PERCLOS

Con respecto a una clasificación de tres estados, representado por la gráfica 6.13.a, en los primeros 15 minutos de conducción el conductor se encuentra en un estado de vigilia, correspondiendo con valores de PERCLOS bajos; a medida que el conductor se encuentra más fatigado, los valores de PERCLOS aumentan y entre los minutos 22 y 32 se alcanza el estado de somnolencia.

En la gráfica 6.14.a se muestra, en diagramas de barras de error, el valor del PERCLOS relativo al estado en el que se encuentra un conductor concreto referenciado por “N° 1”. En el estado de vigilia el valor medio del PERCLOS es de 0,05 y la desviación típica de 0,03, representada por la longitud de la barra asociada a ese estado. Para el estado de somnolencia se tiene un valor medio de 0.3 y una desviación típica de 0.08 que ha aumentado porque la apertura de los ojos, cuando se está entrando en un proceso de somnolencia, está sometida a mayor variabilidad que cuando se encuentran en estado de vigilia. En la figura 6.14.b se muestran los mismos conceptos pero teniendo en cuenta a todos los usuarios. La desviación estándar de cada uno de los estados aumenta con respecto al valor obtenido para un usuario.

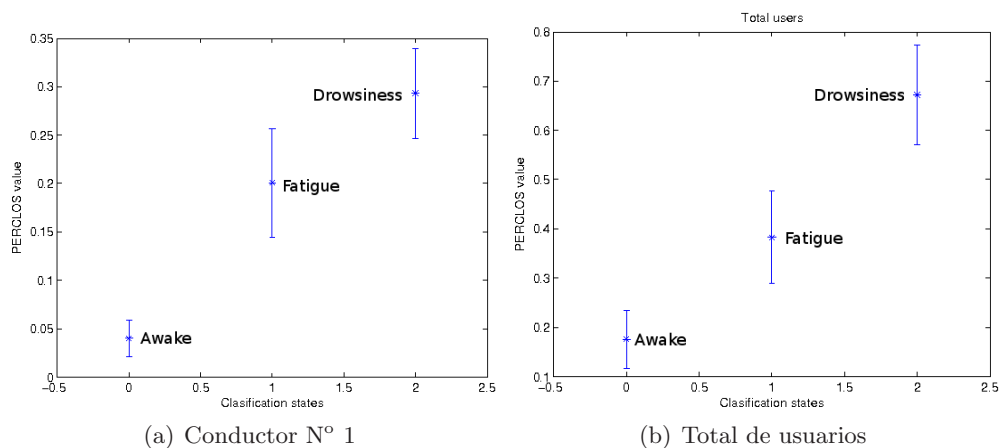


Figura 6.14: Clasificación en tres estados por medio del PERCLOS

En el caso de considerar sólo dos estados de clasificación, se muestran los resultados en la figura 6.15 y puede apreciarse que se distingue mejor donde está ubicada cada una de las clases

ya que la diferencia en el valor medio es elevada y las desviaciones típicas no se solapan.

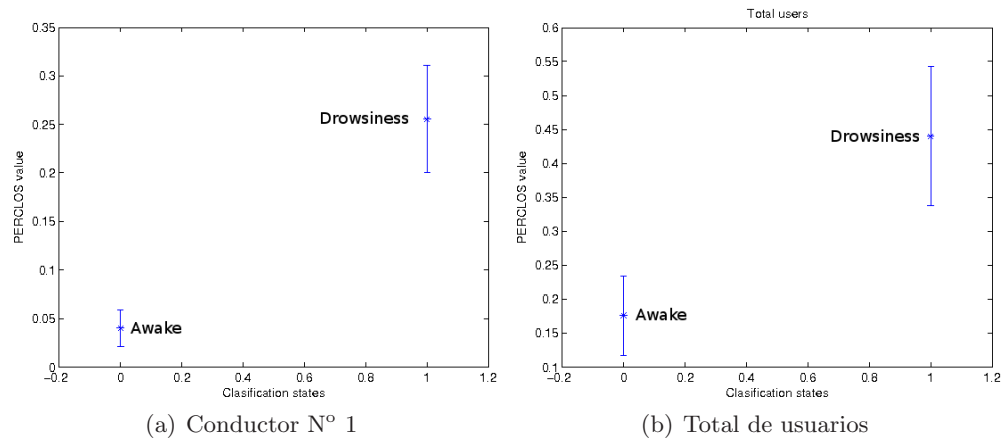


Figura 6.15: Clasificación en dos estados por medio del PERCLOS

En la figura 6.16 se representa la desviación estándar de la posición del vehículo en el carril, evaluada en una ventana temporal de 1 minuto, para cada uno de los tres estados. Analizando los resultados de un usuario, figura 6.16.a, es difícil apreciar diferencias entre el estado de vigilia y el estado de fatiga, aumentando la desviación típica asociada a cada uno de los estados, respecto a los valores obtenidos con la variable PERCLOS. La causa es que la posición del vehículo en el carril está menos correlada con la somnolencia del conductor que el PERCLOS. Cuando se analizan todos los ensayos, figura 6.16.b, se aprecia diferencia entre los valores medios encontrados en el estado de vigilia y fatiga, aunque los valores de desviación típica siguen siendo elevados.

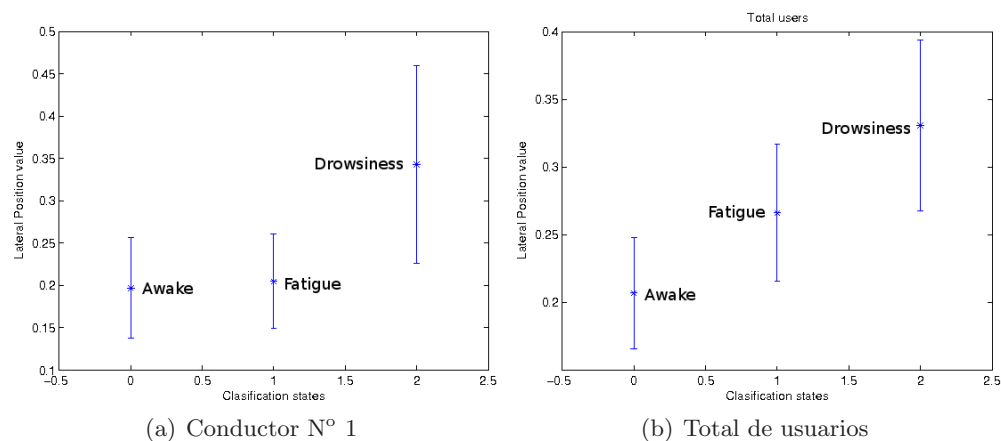


Figura 6.16: Clasificación en tres estados por medio de la desviación estándar de la posición en el carril

En el caso de clasificación binaria (figura 6.17), se diferencia mejor donde están ubicadas cada una de las clases pero las desviación típica, sobre todo en el estado de somnolencia, sigue siendo elevada. Esto demuestra que con este indicador es difícil discriminar el estado del conductor.

Las representaciones anteriores también pueden realizarse para ver las diferencias existentes entre los ensayos realizados sin privación de sueño (0 en el eje de abscisas) y con privación de sueño (1 en el eje de abscisas). Para el total de usuarios, en la figura 6.18.a se representa el PERCLOS en ambas condiciones y en la figura 6.18.b la desviación estándar de la posición del

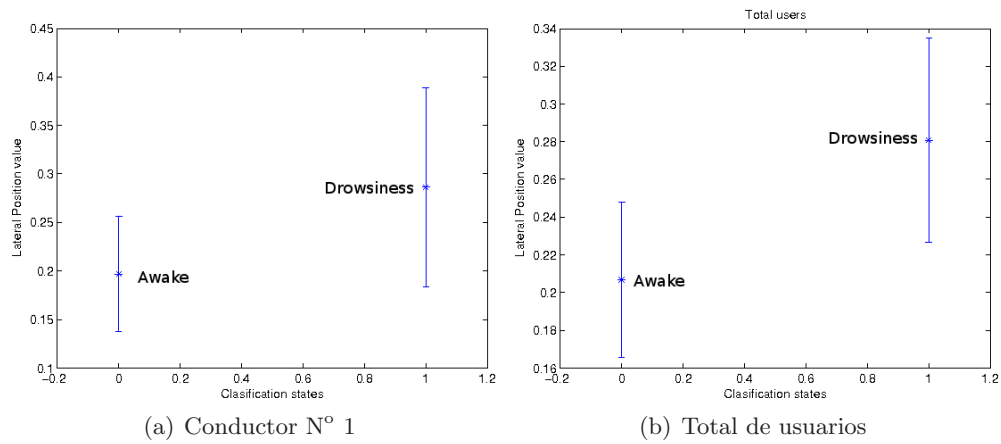


Figura 6.17: Clasificación en dos estados por medio de la desviación estándar de la posición en el carril

vehículo en el carril. La desviación estándar, en los ejercicios de privación de sueño, es grande ya que existen situaciones en las que el conductor se encuentra en vigilia y somnolencia.

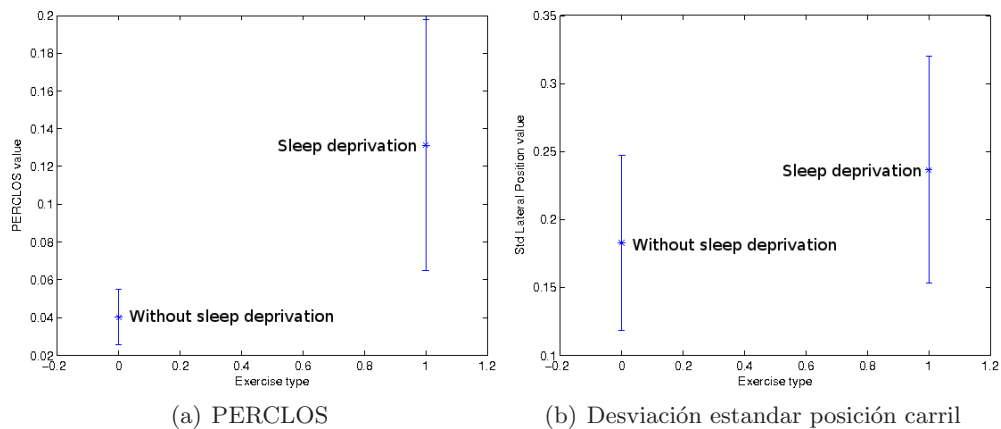


Figura 6.18: Indicadores sin y con privación de sueño

En la figura 6.19 se representa el PERCLOS y la posición del vehículo en el carril, para el mismo tramo de conducción. Las señales representadas se obtuvieron de dos ejercicios diferentes, uno con privación de sueño y otro sin privación de sueño. La posición del vehículo en el carril tiene más variabilidad cuando se está privado de sueño que cuando no se está.

6.3.1. Análisis individual de los indicadores

Los indicadores descritos en la sección 5.1 van a ser aplicados a los ensayos realizados por los usuarios en condiciones de privación y sin privación de sueño, y posteriormente serán utilizados como entradas en un clasificador basado en redes neuronales.

Se han evaluado todos los indicadores descritos para todos los usuarios y durante toda la duración del ensayo. En la figura 6.20 se representan tres indicadores (desviación estándar de la posición del vehículo sobre el carril, lanex y tlc medio) aplicados a un usuario concreto y durante 5 km de conducción. Según la señal de referencia, al ensayo con privación de sueño le corresponde el estado de somnolencia, y sin privación de sueño, el estado de vigilia.

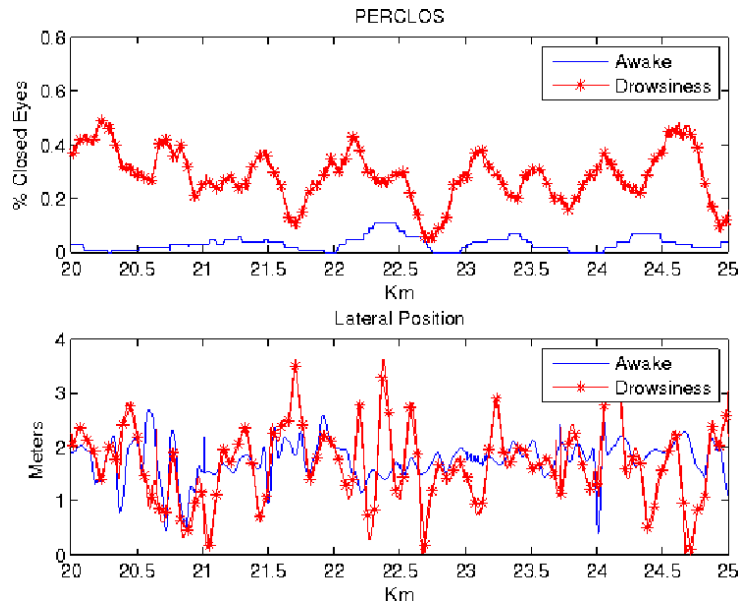


Figura 6.19: Señales de PERCLOS y posición del vehículo en el carril en condiciones de vigilia y somnolencia, para un mismo tramo de conducción

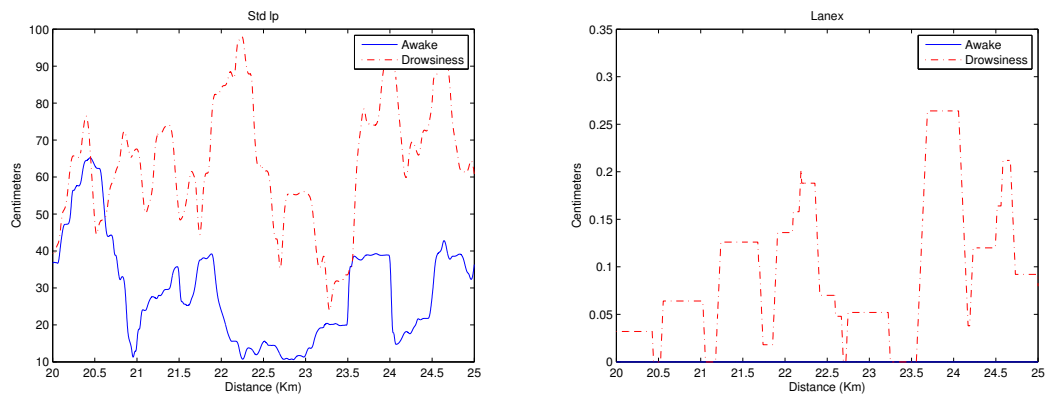
La figura 6.20.a representa la desviación estándar, y su valor medio es mayor en el estado de somnolencia que en el de vigilia. Por otro lado, en la figura 6.20.b se representa el valor lanex, que es mayor en el estado de somnolencia y por último en la figura 6.20.c se aprecia un valor medio de TLC más elevado con somnolencia que sin ella.

Generalizando para todos los usuarios, durante toda la duración del ensayo, y guiándose por la señal de referencia, se han evaluado con un diagrama de barras de error el valor medio y la desviación típica de todos los indicadores en los tres posibles estados a clasificar. La figura 6.21 representa dos de ellos.

Se aprecia en la figura 6.21.b. que los valores medios del TLC para estados diferentes son muy parecidos, por lo que la clasificación será complicada de realizar con ese indicador. Aplicando el diseño de la red neuronal descrito en la sección 5.3.1, y las entradas y salidas empleadas para el entrenamiento, validación y test comentadas en la sección 5.2.1.3 se obtiene una tasa de aciertos como la mostrada en la tabla 6.4 del clasificador ternario empleado individualmente con cada uno de los siguientes indicadores: PERCLOS, posición del vehículo en el carril, su media y su desviación estándar, movimiento del volante, su media y desviación estándar, heading error, su media y desviación estándar, error cuadrático medio de la posición del vehículo en el carril (mse), lanex, tlc superior a 2 s, tlc superior a 0.5 segundos, tlc medio y movimientos rápidos del volante. Además, se representa en la figura 6.22, los diagramas de tela de araña correspondientes a la matriz de confusión de la ecuación (6.1) del indicador TLC inferior a 2 segundos, para analizar la calidad del clasificador.

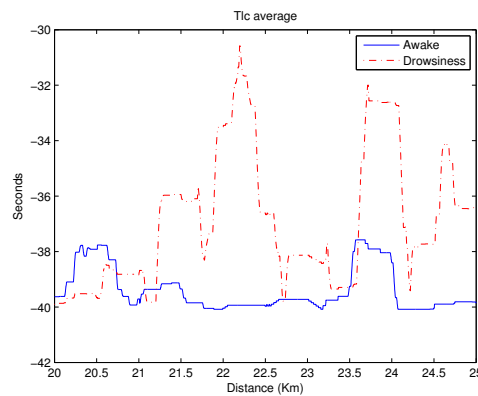
$$matriz\ confusion\ tlc\ 2 = \begin{bmatrix} 0,86 & 0,46 & 0,10 \\ 0,13 & 0,53 & 0,47 \\ 0,20 & 0,15 & 0,42 \end{bmatrix} \quad (6.1)$$

En la figura 6.22.a se puede apreciar que las clasificaciones correctas del estado somnolencia se encuentran próximas al vértice de clasificación aleatoria pero el triángulo correspondiente a la matriz de confusión es externo al aleatorio y la calidad de los aciertos puede considerarse aceptable. Por otro lado, en la figura 6.22.b existen vértices del hexágono de clasificación que



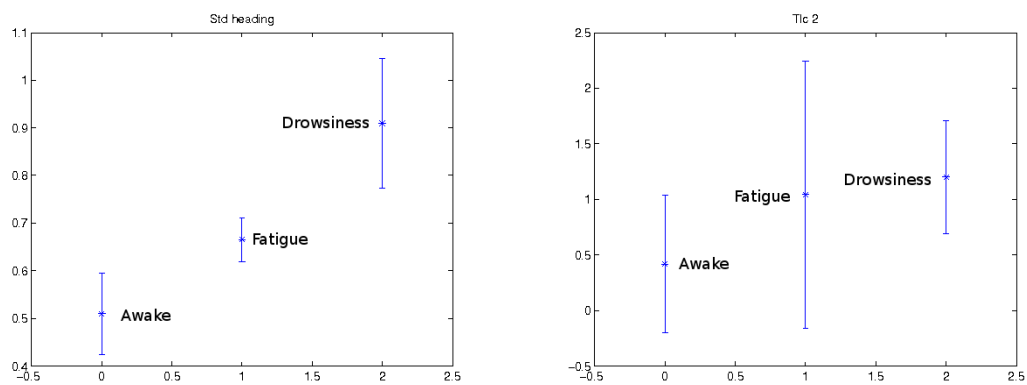
(a) Desviación estándar posición vehículo en el carril

(b) Lanex



(c) TLC medio

Figura 6.20: Indicadores en simulación



(a) Desviación estándar del error del ángulo de guiñada

(b) TLC inferior a 2 segundos

Figura 6.21: Representación en diagrama de barras de error para tres estados de clasificación y dos indicadores

se encuentran fuera del hexágono de clasificación aleatoria, lo que indica una baja calidad del clasificador.

En la tabla 6.4 se aprecia que la tasa de aciertos, en todos los indicadores relativos a la

Tabla 6.4: Clasificación tres estados

Indicador	Tasa de acierto
Lateral Position	18.75
Std Lateral Position	59.69
Steering Wheel	10.13
Std Steering Wheel	57.02
Heading	8.54
Std Heading	59.56
Mse	61.92
Lanex	52.90
Tlc 2 seg.	68.71
Tlc 0.5 seg.	54.47
Tlc medio	31.72
Rápidos mov. volante	9.37
PERCLOS	94.38

conducción, es inferior al 65 %, excepto en el TLC inferior a 2 segundos que se alcanza el 68 %. Estos resultados son insuficientes para cualquier sistema de evaluación del estado del conductor por lo que en lo sucesivo se optará por una clasificación binaria con objeto de elevar estos ratios.

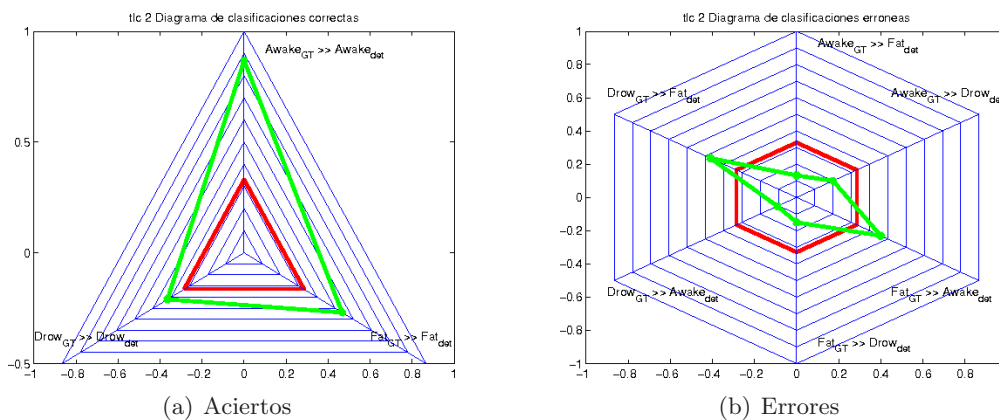
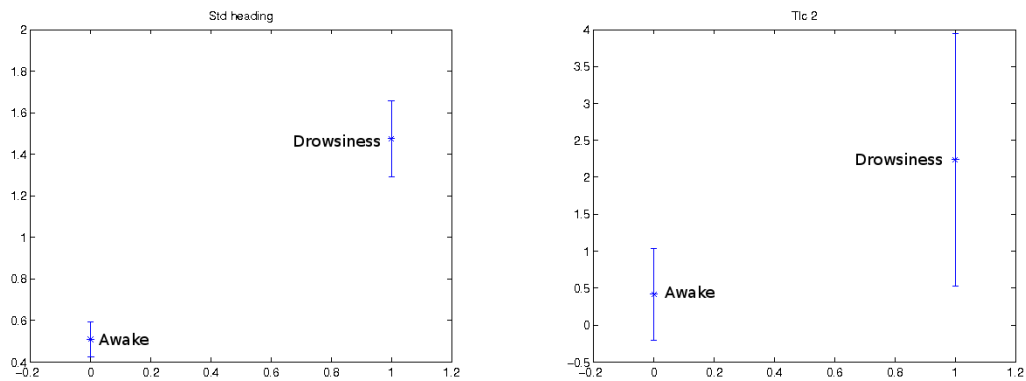


Figura 6.22: Representación en tela de araña para el indicador TLC inferior a 2 segundos

La figura 6.23 corresponde a la representación en diagramas de barras de error, del valor medio y la desviación típica de los indicadores anteriores cuando la clasificación es binaria, apreciándose que respecto a la clasificación terciaria de la figura 6.21 la diferencia encontrada entre clases aumenta.

En la tabla 6.5 la tasa de aciertos se ha calculado según se indicó en la sección 6.2.1.

Valores del 50 % de la función objetivo f , explicada en la sección 5.2.1.2, representan una clasificación aleatoria carente de interés, por lo que sólo se tendrán en consideración valores superiores al 70 %. A la vista de la tabla los indicadores PERCLOS, TLC medio, mse, desviación estándar del error del ángulo de guiñada y desviación estándar de la posición del vehículo en el carril son los más adecuados, notándose que son los indicadores más complejos los que proporcionan valores de la función objetivo más elevados. Los valores de los indicadores de la



(a) Desviación estándar del error del ángulo de guiñada

(b) TLC inferior a 2 segundos

Figura 6.23: Representación en diagrama de barras de dos indicadores con clasificación binaria

Tabla 6.5: Indicadores sin optimizar

Indicador	Tasa de acierto	Especificidad	Sensibilidad	f
Lateral Position	66.71	92.45	18.75	55.6
Std Lateral Position	71.46	74.68	68.50	71.59
Steering Wheel	43.70	20.75	75.00	47.87
Std Steering Wheel	62.92	98.11	0	49.05
Heading	51.15	45.28	56.25	50.76
Std Heading	86.45	98.11	62.50	80.30
Mse	76.25	84.91	59.38	72.14
Lanex	69.98	100	15.63	57.81
Tlc 2 seg.	74.68	100	28.13	64.06
Tlc 0.5 seg.	66.45	100	6.25	53.12
Tlc medio	82.92	86.79	71.88	79.33
Rápidos mov. volante	55.86	84.91	3.13	44.02
PERCLOS	97.61	92.56	89.38	90.97

tabla 6.5 prácticamente coincide con los obtenidos en el trabajo de [Sandberg, 2011], a pesar de que el conjunto de datos analizado es totalmente diferente, lo cual valida la metodología empleada.

6.3.2. Análisis de los indicadores optimizados

Aplicando las técnicas de optimización descritas en la sección 5.2, se han obtenido los parámetros optimizados de los indicadores: error cuadrático medio (mse), Lanex, TLC, movimientos del volante y velocidad de giro del volante. En la tabla 6.6 se muestran los resultados del proceso de optimización.

Los indicadores genéricos de variabilidad, expuestos en la sección 5.1.8, también pueden ser optimizados. El indicador genérico puede ser aplicado a cualquiera de los indicadores expuestos aunque se ha optado por aplicarlo a la desviación estándar de la posición del vehículo sobre

Tabla 6.6: Parámetros de indicadores optimizados

mse	lanex	tlc	wheel	wheel_velocity
$p = 0,48$	$c_L = 0,6386$	$a = 0,6168$	$c_1 = -3,22$	$d = 0,24$
	$c_R = 0,1544$		$c_2 = 10,35$	

el carril, a la desviación estándar del ángulo de giro del volante y a la desviación estándar del error del ángulo de guiñada, por ser los indicadores más adecuados para mejorar los resultados obtenidos anteriormente. Tras el proceso de optimización se han obtenido los parámetros de la tabla 6.7.

Tabla 6.7: Parámetros de indicadores genéricos

Parametros	Std lp	Std wheel	Std heading error
σ	0.7888	1.693	0.4477
p	0.2890	0.8688	0.9340
k	3.5951	-1.3110	0.7952
c_L	0.7000	0.4023	0.8909
α_L	0.3271	-1.6880	0.9340
β_L	-1.4606	0.1837	0.1656
c_R	1.7514	2.2573	1.2461
α_R	-1.1867	1.9283	0.8530
β_R	0.7104	0.9899	-1.1849

Determinados los parámetros de los indicadores optimizados se evalúan para cada uno de los ensayos de simulación. En la figura 6.24 se representan, en un intervalo de conducción de 5 Km y para un mismo usuario, dos indicadores optimizados. Las representaciones de Lanex y TLC indican grados diferencias entre el estado de vigilia y el estado de somnolencia, por lo que se demuestra que estos indicadores son apropiados para la detección de la somnolencia del conductor.

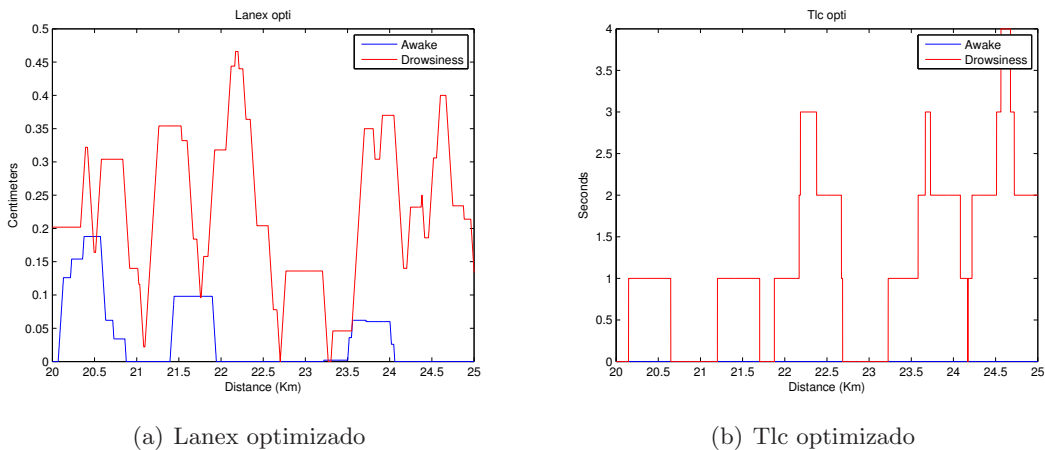


Figura 6.24: Representación temporal de indicadores optimizados

Generalizando para todos los usuarios, durante todo el tiempo del ensayo, y guiándose por la señal de referencia, se representan en la figura 6.25 dos indicadores optimizados.

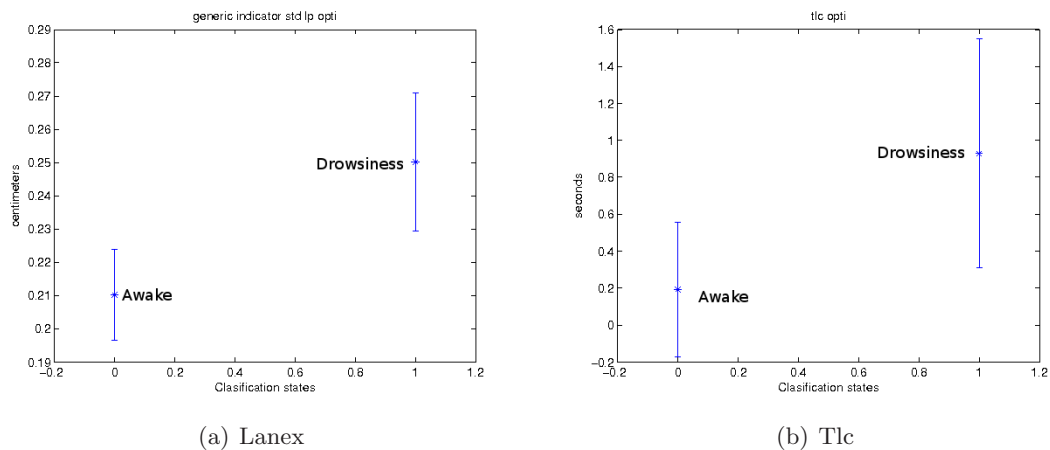


Figura 6.25: Diagrama de barras de error de dos indicadores optimizados para dos estados de clasificación.

En la tabla 6.8 se muestra la tasa de aciertos, especificidad, sensibilidad y valor de la función objetivo de la salida de la red neuronal cuando la entrada es cada uno de los siguientes indicadores optimizados: genérico de variabilidad de la desviación estándar de la posición del vehículo en el carril (GVI std lp), del ángulo de giro del volante (GVI std wheel) y del error de ángulo de guiñada (GVI std he), error cuadrático medio de la posición del vehículo en el carril (mse), Lanex, TLC y movimientos rápidos de volante (MRV).

Comparando las tablas 6.5 y 6.8 se aprecia que los resultados obtenidos con los indicadores optimizados mejoran la tasa de aciertos.

Tabla 6.8: Indicadores optimizados

Indicadores	Tasa de acierto (%)	Especificidad	Sensibilidad	f
GVI std lp	80.57	88.68	62.50	75.59
GVI std wheel	75.86	96.23	37.50	66.86
GVI std he	87.36	92.45	81.23	86.84
mse	74.68	100	48.13	74.06
Lanex	88.80	100	65.63	82.81
TLC	78.21	86.79	59.38	73.08
MRV	74.68	84.91	53.13	69.02

El mayor valor de la función objetivo es la del indicador genérico de variabilidad de la desviación estándar del error del ángulo de guiñada con un 86.84 % que mejora los resultados de la desviación estándar del error del ángulo de guiñada sin optimizar con un valor del 80.30 %. El error de ángulo de guiñada sólo depende del estado del conductor ya que es independiente de la trayectoria de la carretera, por lo que es lógico que la relación con la señal de referencia sea alta. Todos los indicadores de la tabla 6.8 superan el 70 % excepto GVI std wheel y MRV, por lo tanto se demuestra que los cuatro principales parámetros de la conducción que proporcionan mayor información para la detección de somnolencia son: el indicador GVI aplicado al error del ángulo de guiñada con un valor de la función objetivo del 86.84 %, el indicador Lanex con un valor de la función objetivo del 82.81 %, el indicador GVI aplicado a la desviación estándar de la posición del vehículo en el carril con un valor de la función objetivo del 75.59 % y el indicador

mse con un valor de la función objetivo del 74.06 %.

6.3.3. Análisis de la fusión de indicadores

Hasta ahora se han considerado los indicadores individualmente. En la tabla 6.9 se han combinado varios de ellos, atendiendo a los valores más altos de la función objetivo f.

Tabla 6.9: Combinación de indicadores simulación

Indicadores	Tasa de acierto (%)	Especificidad	Sensibilidad	f
PERCLOS - GVI std he	99.13	100	99	99.5
PERCLOS - GVI std lp	97.95	100	90.63	95.3
PERCLOS - Lanex opti	97.95	100	90.63	95.3
PERCLOS - TLC opti	95.60	100	84.38	92.19
PERCLOS - mse opti	99.00	100	96.88	98.44
Std tlc avg - GVI std he	73.12	88.68	46.88	67.78
Std tlc avg - GVI std lp	80.57	98.11	46.88	72.49
Std tlc avg - Lanex opti	86.45	88.68	78.13	83.40
Std tlc avg - TLC opti	86.45	100	59.38	79.69
Std tlc avg - mse opti	85.27	100	56.25	78.12
Std tlc avg - GVI std he - mse opti	93.25	96.23	90.62	93.42

Prácticamente todos los indicadores presentan mejores resultados cuando se combinan con otros, de dos en dos. La combinación del PERCLOS con otros indicadores obtienen los mejores resultados y mejoran su detección individual. La combinación entre el PERCLOS y el indicador de variabilidad genérico de la desviación estándar del error del ángulo de guiñada o la combinación con el mse optimizado generan una tasa de acierto superior al 98 %, superando los resultados obtenidos en [Sandberg et al., 2011]. Para combinaciones de más de 2 indicadores los resultados obtenidos siempre son peores que para combinaciones de dos en dos. A modo de ejemplo se muestra la combinación de los tres mejores indicadores donde se puede observar que el valor de la función objetivo es del 93.42 %, probablemente por el excesivo número de entradas que hace que la red no sea capaz de clasificar.

6.4. Resultados en condiciones reales

En la evaluación de los resultados de esta tesis, relacionados con ensayos reales, se dispone de aproximadamente de 1296 minutos de vigilia y 504 minutos de fatiga en los que en algunos de ellos hay síntomas muy acusados de fatiga, tanto en las señales de conducción, como en los parámetros faciales del conductor.

Respecto a la señal de referencia sólo se considera una clasificación de dos estados (vigilia y fatiga) porque en los ensayos reales no se ha alcanzado el estado de somnolencia por ser peligroso para el conductor. El porcentaje de minutos que tiene cada uno de los estados, para la totalidad de los ensayos, es el mostrado en la figura 6.26.

En cada uno de los ejercicios realizados se ha evaluado, con las técnicas descritas en la sección 4.6, el porcentaje de tiempo que los ojos se encuentran cerrados (PERCLOS), obteniéndose señales como las mostradas en la figura 6.27. En los primeros 100 minutos de conducción el conductor se encuentra en estado de vigilia, correspondiendo con valores de PERCLOS bajos; a medida que el conductor se encuentra más fatigado, los valores de PERCLOS aumentan y entre los minutos 110 y 135 se alcanza el estado de fatiga, regresando al estado de vigilia en la última parte del ensayo.

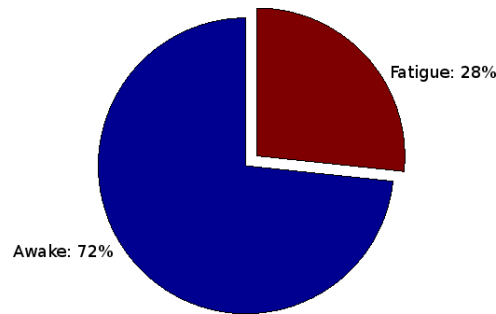


Figura 6.26: Porcentajes de clasificación del total de ensayos

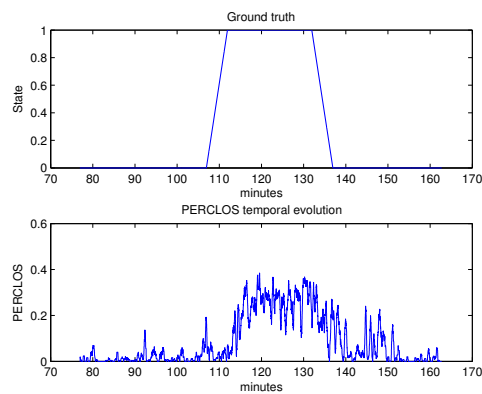


Figura 6.27: Señal de referencia y evolución temporal de PERCLOS

En la gráfica 6.28.a se muestra, en diagramas de barras de error, el valor del PERCLOS relativo al estado en el que se encuentra el conductor. En el estado de vigilia el valor medio del PERCLOS es de 0,1 y la desviación típica de 0.15, representada por la longitud de la barra asociada a ese estado. Para el estado de fatiga se tiene un valor medio de 0.6 y una desviación típica de 0.2 que ha aumentado porque la apertura de los ojos, cuando se está entrando en un proceso de somnolencia, está sometida a mayor variabilidad que cuando se está en estado de vigilia. En la figura 6.28.b se ha tenido en cuenta a todos los usuarios.

En la figura 6.29 se representa la desviación estándar de la posición del vehículo en el carril para cada uno de los dos estados. Analizando los resultados de un usuario, figura 6.29.a, es difícil apreciar diferencias entre el estado de vigilia y el estado de fatiga, aumentando la desviación típica asociada a cada uno de los estados, respecto a los valores obtenidos con la variable PERCLOS. La causa es que la posición del vehículo en el carril está menos correlada con la somnolencia del conductor que el PERCLOS y depende más de la forma de la carretera y del tipo de conducción. Cuando se analizan todos los ensayos, figura 6.29.b, se aprecia diferencia entre los valores medios encontrados en el estado de vigilia y fatiga aunque la diferencia sigue siendo pequeña, además, los valores de desviación típica siguen siendo elevados.

6.4.1. Análisis individuales de los indicadores

Se han representado diferentes variables registradas de uno de los usuarios para relacionarlas con la señal de referencia. Las señales se han generado en el tramo de carretera de conducción expuesto en la sección 3.2. En la figura 6.30 se representan las señales de entrada de PERCLOS, desviación del vehículo con respecto a la línea derecha de la carretera y el ángulo de giro del

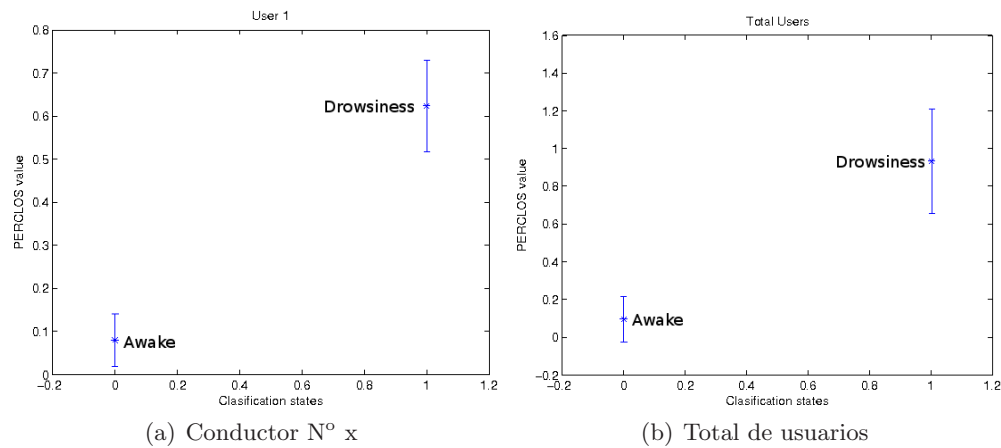


Figura 6.28: PERCLOS con clasificación binaria

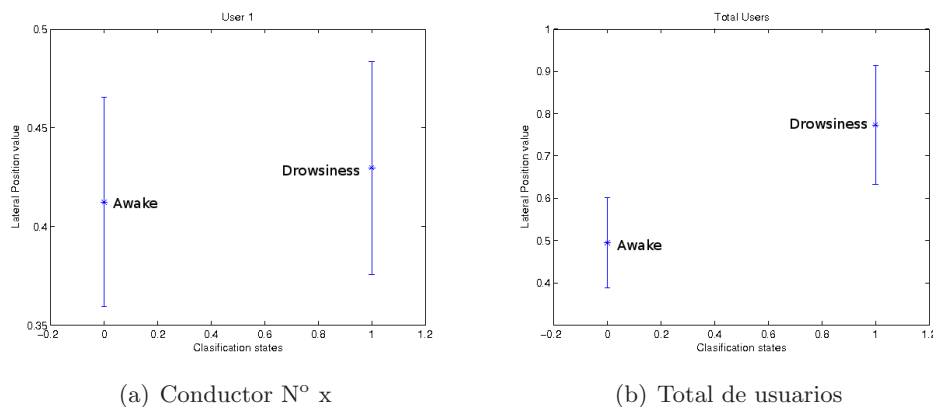


Figura 6.29: Porcentaje STD LP dos estados

volante, junto a la señal de referencia con la que habrán de ser comparadas, notándose que no existe mucha correlación entre ellas.

La señal de PERCLOS es bastante ruidosa debido a que el sistema de medida de visión artificial se ve sometido a errores por deslumbramientos o sombras en ambientes diurnos. Por otro lado, en la señal del desplazamiento del vehículo se aprecia que en los primeros 25 minutos no existe mucha dispersión en la medida, que va en aumento según transcurre el tiempo. Aunque el comienzo de la dispersión coincide con el indicado por la señal de referencia, no parece que tengan relación ya que el final de la dispersión no coincide con el final de la señal de referencia. En la señal del ángulo de giro del volante no se aprecia información relacionada con la fatiga.

La figura 6.30 denota que no hay correlación entre las señales de conducción y la señal de referencia, pero no ocurre así con la señal de PERCLOS, por lo que se obtendrán los indicadores descritos en la sección 5.1 para posibilitar la detección de somnolencia.

Aplicando el diseño de la red neuronal descrito en la sección 5.3.1, y las entradas y salidas empleadas para el entrenamiento, validación y test comentadas en la sección 5.2.1.3, en la tabla 6.10 se muestra la tasa de aciertos del clasificador binario, empleado individualmente con cada uno de los siguientes indicadores: PERCLOS, posición del vehículo en el carril, su media y su desviación estándar, movimiento del volante, su media y desviación estándar, velocidad del ángulo de giro del volante, su media y desviación estándar, error cuadrático medio de la posición del vehículo en el carril (mse), lanex, tlc inferior a 2 s, tlc inferior a 0.5 segundos, tlc medio y

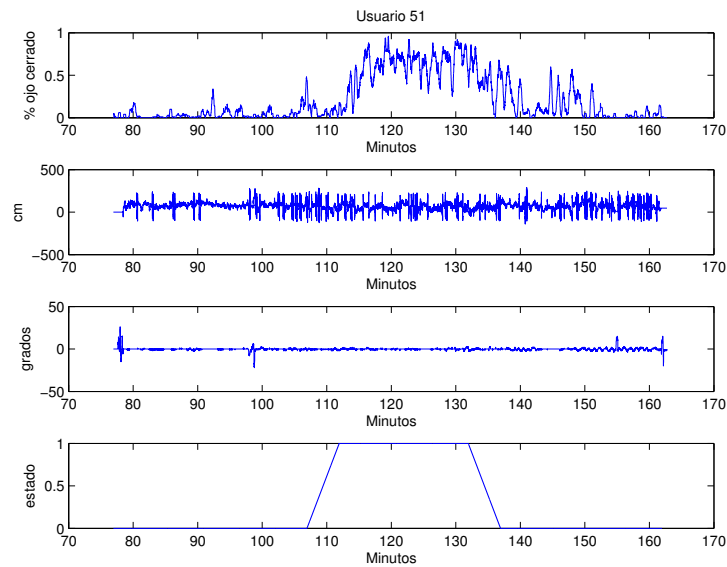


Figura 6.30: Indicadores sin optimizar

movimientos rápidos del volante.

Tabla 6.10: Indicadores sin optimizar

Indicador	Tasa de acierto	Sensibilidad	Especificidad	f
Lateral Position	37.71	5.26	100	52.63
Std Lateral Position	61.24	57.89	57.14	57.51
Steering Wheel	54.24	52.63	42.86	47.74
Std Steering Wheel	66.33	84.21	0	42.1
Velocidad volante	50.82	60.53	0	30.26
Std velocidad volante	71.51	92.11	0	46.05
Mse	68.0	86.84	0	43.42
Lanex	63.46	71.05	21.43	46.24
Tlc 2 seg.	63.11	65.53	55	60.26
Tlc 0.5 seg.	45.37	31.58	50	40.79
Tlc medio	66.33	84.21	0	42.1
Rápidos mov. volante	76.68	48.15	37.28	42.71
PERCLOS	94.05	92.21	89	90.6

De los indicadores evaluados, los más favorables para el diseño de un sistema de detección de somnolencia serían el PERCLOS, el TLC inferior a 2 segundos y la desviación estándar de la posición del vehículo en el carril, con valores de la función objetivo superiores al 55%. Aunque hay otros indicadores con tasas de acierto superiores al 55%, tienen la especificidad o la sensibilidad muy bajas, lo que hace que no sean adecuados como detectores de somnolencia ya que, una sensibilidad nula indica que no se detecta ningún estado de fatiga y todas las salidas del clasificador son vigilia. Como el 72% de todas las muestras corresponden a vigilia, la tasa de aciertos coincide con el 72% pero el clasificador es malo al no detectar ningún estado de fatiga.

Por ese motivo es más fiable fijarse en la función objetivo f que en la tasa de aciertos.

6.4.2. Análisis de los indicadores optimizados

Muchos de los indicadores evaluados necesitan fijar ciertas constantes que pueden ser optimizadas para mejorar los resultados del clasificador. La optimización se ha realizado empleando los algoritmos genéticos descritos en la sección 5.2. Con esta técnica se aportan 7 indicadores de entrada al clasificador, que permitirán evaluar la calidad de los resultados de detección de somnolencia. Los parámetros optimizados son los correspondientes al indicador genérico de variabilidad de la desviación estándar de la posición del vehículo en el carril, del ángulo de giro del volante y de la velocidad del ángulo de giro del volante. Por otro lado, se ha obtenido indicadores optimizados del error cuadrático medio (mse), lanex, tlc, movimientos del volante y velocidad de giro del volante. En la tabla 6.11 se exponen los resultados del proceso de optimización para estos últimos indicadores.

Tabla 6.11: Parámetros de indicadores optimizados

mse	lanex	tlc	wheel	wheel_velocity
$p = 175,87$	$c_L = 268$	$a = 6,10$	$c_1 = -1,81$	$d = 1,92$
-	$c_R = 52$	-	$c_2 = 0,57$	-

Los indicadores genéricos de variabilidad también pueden ser optimizados, y ser aplicados a cualquiera de los indicadores expuestos, aunque en la tabla 6.12 se han aplicado a la desviación estándar de la posición del vehículo sobre el carril, a la desviación estándar del ángulo de giro del volante y a la desviación estándar de la velocidad del ángulo de giro del volante.

Tabla 6.12: Parámetros de indicadores optimizados genéricos

Parametros	Std lp	Std wheel	Std wheel velocity
σ	0.9167	3.0687	1.3692
p	0.0349	1.3730	-1.0187
k	0.1419	2.7123	-0.8818
c_L	0.6098	-4.0468	0.9859
α_L	-0.2539	3.4330	0.2914
β_L	-1.4214	2.8663	0.8431
c_R	0.5633	1.9236	0.1808
α_R	0.4378	1.2922	1.6127
β_R	0.4725	2.8010	0.4613

La tabla 6.13 muestra la tasa de aciertos del clasificador binario cuando las entradas son los indicadores anteriores, aplicados individualmente.

De los indicadores optimizados, los que mejor detectan la somnolencia son el TLC optimizado, el indicador genérico de variabilidad de la desviación estándar del vehículo en el carril y el mse, con valores superiores al 60%. Se aprecia como los indicadores optimizados mejoran el resultado de la función objetivo sin optimizar. Muchos de los indicadores optimizados no pueden ser empleados por no tener la sensibilidad adecuada, ya que una sensibilidad nula es indicativa de haber realizado una clasificación constante a un único estado de salida: vigilia.

Tabla 6.13: Indicadores optimizados

Indicadores	Tasa acierto(%)	Especificidad	Sensibilidad	f
GVI Lateral Position	76.27	62.89	62.14	62.51
GVI Steering Wheel	74.51	97.30	0	48.65
GVI velocidad volante	76.27	96.75	20	58.37
Mse	58.86	44.74	78.57	61.65
Lanex	73.35	92.11	0	46.05
Tlc	81.39	92.11	42.86	67.48
Rapidos movimientos	54.42	52.63	35.71	44.17

6.4.3. Análisis de la fusión de indicadores

De todos los indicadores relacionados con la conducción son los indicadores optimizados los que proporciona la función objetivo más elevada, por esa razón en la fusión de indicadores siempre se va a utilizar uno de los indicadores optimizados. De los indicadores sin optimizar los que mejores tasas de acierto obtienen son: PERCLOS y tlc inferior a 2 segundos. Las combinaciones elegidas se indican en la tabla 6.14.

Tabla 6.14: Combinación de indicadores

Indicadores	Tasa acierto(%)	Especificidad	Sensibilidad	f
PERCLOS - tlc opti	93.48	97.32	90.00	93.66
PERCLOS - mse opti	95.93	100	90.00	95
PERCLOS - GVI lateral position	90.76	95.54	87.50	91.52
TLC 2 - tlc opti	72.15	94.64	5.00	49.82
TLC 2 - mse opti	76.38	71.43	85.00	78.21
TLC 2 - GVI lateral position	77.70	92.86	40.00	66.43

Prácticamente todos los indicadores presentan mejores resultados cuando se combinan con otros. La combinación del PERCLOS con otros indicadores obtiene los mejores resultados y mejoran su detección individual. La combinación entre el PERCLOS y el mse optimizado genera tasas de aciertos del 95%. Cuando se fusiona con indicadores de conducción, el valor del PERCLOS es determinante para obtener resultados de la función objetivo superiores al 90%, mientras que la fusión entre indicadores de conducción no supera el 80%.

6.5. Simulación vs Condiciones reales

Tanto en el estado de vigilia como en el de somnolencia existen diferencias entre los ensayos simulados y reales, que se reflejan en los indicadores del conductor y de la conducción.

Respecto al indicador PERCLOS se han obtenido mejores resultados en simulación que en condiciones reales, debido al ruido introducido en condiciones reales por la iluminación, aun así los resultados obtenidos para ambos ensayos han sido muy similares en cuanto a tanto por ciento de detección.

El indicador del error del ángulo de guiñada, relacionado con la conducción en los ejercicios de simulación, proporciona buenos resultados en la estimación de la somnolencia pero no ha podido compararse con las condiciones reales por no disponer de él en estas condiciones. Se considera una señal de gran interés para ser medida en un vehículo real.

En la evaluación de los indicadores optimizados se obtienen mejores resultados en los ensayos de simulación, puesto que en los ensayos reales se cometen errores a la hora de estimar la posición del vehículo en el carril por utilizar un sistema de visión artificial basado en un Lane Departure Warning que está sometido a ruido, lo que no ocurre en los ensayos de simulación. Lo mismo ocurre con los indicadores derivados del ángulo de giro del volante.

En ambas condiciones, la somnolencia está asociada con el incremento de la desviación del vehículo en el carril, y en simulación los pequeños movimientos de corrección del volante que se hacen en estado de vigilia, desaparecen cuando el conductor está en estado de somnolencia.

En simulación, la mejor combinación es la fusión entre el PERCLOS y la desviación estándar del error del ángulo de guiñada con un valor de la función objetivo del 99.5%, la segunda mejor combinación es el PERCLOS y el indicador mse optimizado con un valor del 98.44%. En condiciones reales, la mejor combinación es el PERCLOS y mse optimizado con un valor de 95.00%, segunda mejor combinación en simulación. Si se hubiera tenido el error del ángulo de guiñada en condiciones reales quizás hubiera dado el mejor porcentaje en la clasificación, superando inclusive la fusión con el indicador mse optimizado en condiciones reales.

Los resultados obtenidos avalan la metodología empleada que traslada los resultados del simulador naturalista a las condiciones reales, extrayendo conclusiones válidas para las condiciones reales con porcentajes de detección menores debido al mayor ruido asociado a las señales.

Capítulo 7

Conclusiones y trabajos futuros

La inatención debida a la fatiga o somnolencia al volante es bien conocida como un factor de riesgo determinante en accidentes de tráfico. Muchos conductores combinan la privación de sueño con excesivas horas de conducción existiendo un grave conflicto entre las necesidades fisiológicas y las actividades sociales o profesionales del conductor. Por lo tanto, la clave para conseguir la prevención de accidentes es el entendimiento de los límites humanos entre la fatiga, la somnolencia y la privación de sueño. La fatiga y somnolencia en conductores es un factor muy importante a tener en cuenta por ocasionar muchos accidentes de tráfico. Por esta razón durante la última década el problema de la somnolencia en conductores ha recibido un incremento de atención por parte de la comunidad científica, las compañías de vehículos y en general por la sociedad.

Uno de los objetivos de esta tesis ha sido la implementación de un sistema de visión computacional para la estimación de la apertura de los ojos del conductor utilizando técnicas de reconocimiento de objetos y caracterización de parámetros de los objetos detectados, tanto en condiciones de simulación naturalista como en condiciones reales de conducción, para conductores sin y con privación de sueño en la realización de los ensayos. El sistema desarrollado puede ser empleado en tiempo real y generar un aviso cuando se detecta que el conductor entra en estado de somnolencia.

La estimación de la apertura del ojo se realiza de forma automática e independiente del usuario y está basada en la caracterización y localización de los ojos del conductor mientras conduce, haciendo uso de técnicas de detección de objetos por apariencia, detector de Viola and Jones, junto con las modificaciones realizadas por Rainer Lienhart y Jochen Maydt para la detección de la cara y los ojos en la imagen. Para disminuir fallos en la detección de los ojos se han empleado técnicas de clasificación K-mean a los candidatos generados por Viola and Jones, haciendo que la detección sea más robusta. Para la evaluación de la posición y apertura del ojo se ha realizado un filtrado adaptativo y un modelo Gaussiano. La posición medida en el instante k sirve de entrada a un filtro de Kalman para predecir la posición en el instante $k+1$, lo que facilita su detección en ese instante y disminuye el tiempo de proceso.

Tres de los problemas que surgen en la detección de los ojos son: las condiciones cambiantes de iluminación, el tiempo que transcurre en la localización de los mismos y el complejo proceso de evaluación de la apertura del ojo [Yuille et al., 1989b]. Este trabajo se ha apoyado en técnicas efectivas que se adaptan a los cambios de iluminación y en proyecciones integrales de la imagen y en un modelo Gaussiano con objeto de estimar la apertura y posición del ojo a la frecuencia de 30 imágenes por segundo. De esta forma, a diferencia de otros sistemas existentes en el estado del arte [Ryan et al., 2008, D’Orazio et al., 2007a, Senaratne et al., 2007, Suzuki et al., 2006b], se obtiene un sistema en tiempo real y robusto que ha sido probado en condiciones reales durante largos procesos de evaluación. Por otro lado, a diferencia de otros sistemas comerciales cerrados [SeeingMachines, 2011, SmartEye, 2011], nuestro sistema es totalmente abierto, demostrando

unos resultados iguales o superiores a los comerciales, proporcionando detalles técnicos de funcionamiento y un exhaustivo plan de test para el análisis de su funcionamiento.

Por otro lado, se han evaluado parámetros relacionados con información del vehículo proporcionado por sistemas ADAS como: la posición del vehículo sobre el carril, el ángulo de giro del volante y el error del ángulo de guiñada para determinar indicadores que puedan estar relacionados con la somnolencia. Los indicadores obtenidos han sido las desviaciones estándar de las variables anteriores, el error cuadrático medio de la posición del vehículo en el carril, el tiempo para rebasar los límites del carril, e información de los movimientos rápidos realizados en el volante, como ya se recoge en [Sandberg et al., 2011] para otro conjunto de datos diferentes.

El sistema se ha evaluado tanto en simulación, bajo condiciones controladas de iluminación, como en condiciones reales, con niveles de iluminación muy dispares. El total de usuarios evaluados es de 20, habiéndose obtenido gran cantidad de horas de conducción y vídeos de la cara del conductor, así como la señal ground-truth de somnolencia y señales relacionadas con la conducción. En ambos escenarios se obtienen resultados superiores al 90 % en la función objetivo de un clasificador binario que utiliza como entrada la información de la apertura del ojo, mejorándose los resultados de otros trabajos del estado del arte [Friedrichs & Yang, 2010a, Caterpillar, 2008].

Se ha validado una metodología consistente en analizar comportamientos en un entorno controlado mediante un simulador naturalista y validarlo en condiciones reales. Los resultados obtenidos en simulación y en condiciones reales se comportan de la misma forma si bien los porcentajes de detección disminuyen un poco en las condiciones reales, debido al mayor ruido de medida de las señales de entrada.

El algoritmo desarrollado alcanza elevadas tasas de acierto de detección de somnolencia. Esta información, junto con la de señales del vehículo como: la desviación estándar del vehículo en el carril, el ángulo de giro del volante o el error del ángulo de guiñada; puede ser empleada por expertos para analizar resultados sobre el comportamiento de un conductor con privación de sueño.

En la última década se ha investigado mucho en aspectos de seguridad vial con sistemas que ayudan al conductor (ADAS) interpretando las señales de tráfico, evaluando la distancia con el vehículo precedente para evitar golpes traseros en retenciones inesperadas, detectando peatones en condiciones diurnas y nocturnas para prevenir atropellos, etc, sin embargo los sistemas relacionados con el estado del conductor se encontraban en un punto de inflexión a la hora de determinar características como la distracción o el nivel de somnolencia y ha sido la incorporación de la visión artificial lo que ha transmitido un decidido impulso de progreso e innovación que les permitirá hacerse un hueco en el difícil mundo del automóvil. El objetivo final de que el conductor cese la conducción cuando se haya alcanzado un determinado nivel de somnolencia tiene todavía recorrido hasta su implantación comercial ya que los falsos positivos que alertan al conductor son de momento elevados y más que ayudar pueden generar distracción e incluso incomodidad en el conductor. Esta tesis ha pretendido contribuir a validar estas hipótesis.

7.1. Principales contribuciones

Del desarrollo y análisis de los resultados obtenidos en los capítulos anteriores se considera que las principales contribuciones de esta tesis son las siguientes:

1. **No hay una regla de oro de cómo medir la somnolencia durante la conducción.** Existen muchos indicadores potencialmente prometedores, pero todos ellos tienen limitaciones al proporcionar señales muy ruidosas en las que a veces no se manifiesta claramente la relación con la somnolencia y sobre todo en condiciones reales.
2. **Implementación de un sistema de visión artificial.** Se ha implementado un sistema

de visión artificial abierto y propio capaz de evaluar la señal de PERCLOS en tiempo real sobre un simulador naturalista y en condiciones reales de conducción.

3. **Se han analizado tres niveles de clasificación** en la detección de la fatiga del conductor. De forma experimental se ha concluido que los mejores resultados de clasificación se obtienen con dos niveles.
4. **Inclusión de una nueva metodología de generación de ground truth**, consistente en fusionar la información KSS con la proporcionada por tres expertos.
5. **Desarrollo de un algoritmo robusto de cálculo de PERCLOS**. El algoritmo implementado, basado en un modelo Gaussiano, no necesita umbrales fijos para su funcionamiento, ni procesos de calibración, responde adecuadamente ante cambios de iluminación y es apto para condiciones reales por lo que podría formar parte de cualquier sistema comercial. El sistema es monocular y de coste reducido frente a otros sistemas del estado del arte [SmartEye, 2011], [Ryan et al., 2008], [Zhu & Ji, 2005]. Utiliza un filtro de Kalman como predictor para mejorar el seguimiento.
6. **Alto nivel de detección con la señal de PERCLOS**. La correlación entre la señal de referencia y la que proviene de la apertura de los ojos es muy alta por lo que, cuando la clasificación es binaria, alcanza valores de la función objetivo por encima del 90 %. Este indicador, relacionado con las características faciales del conductor, es el que proporciona los mejores resultados, bastante mejores que los indicadores que provienen de la información del vehículo. En parte, la elevada tasa de aciertos puede venir dada por la influencia que sobre los expertos ejerce la apertura de los ojos cuando analizan los vídeos para construir la señal de referencia.
7. **Análisis de indicadores derivados de la conducción**. Las señales relativas a la conducción no parecen tener correlación con la somnolencia, pero no ocurre así con los indicadores derivados de estas señales que se describen en [Sandberg et al., 2011], sobre todo después de optimizarlos ya que obtienen valores de la función objetivo superiores al 75 %. El indicador del ángulo de guiñada es independiente de la trayectoria del vehículo, obteniendo buenos resultados con tasas de acierto por encima del 80 %. No ocurre así con la posición del vehículo en el carril o con el ángulo de giro del volante, en donde a los movimientos necesarios para seguir la trayectoria se superponen los debidos a efectos de somnolencia.
8. **Optimización y fusión de indicadores**. Optimizando los indicadores con técnicas de optimización estocásticas y fusionando los mismos mediante una red perceptrón multicapa se obtienen mejores resultados que si se tiene en cuenta un sólo indicador. La fusión del PERCLOS con el indicador genérico de variabilidad de la desviación estándar de la posición del vehículo sobre el carril obtiene valores de la función objetivo del 99.5 % en simulación, mejorando el resultado en casi un 10 % respecto a cuando se emplea el PERCLOS individualmente. En condiciones reales se obtiene un valor de la función objetivo del 93.6 % fusionando el indicador de PERCLOS con el indicador del mse optimizado.
9. **Resultados simulación vs. real**. En simulación las dos mejores combinaciones son la fusión entre el PERCLOS y la desviación estándar del error del ángulo de guiñada y el PERCLOS y el indicador mse optimizado. En condiciones reales la mejor combinación es el PERCLOS y mse optimizado. Si se hubiera tenido el error del ángulo de guiñada en condiciones reales quizás hubiera dado el mejor porcentaje en la clasificación, superando inclusive la fusión del PERCLOS con el indicador mse optimizado en condiciones reales,

por lo tanto, los resultados obtenidos avalan la metodología empleada que traslada los resultados del simulador naturalista a las condiciones reales, extrayendo conclusiones válidas para las condiciones reales con porcentajes de detección menores debido al mayor ruido asociado a las señales.

10. **Resultados obtenidos acorde a otros trabajos de interés.** En términos generales, los resultados obtenidos están en concordancia con otros importantes trabajos sobre detección de somnolencia [Sandberg, 2011], a excepción de la discusión sobre la importancia de la variable PERCLOS ya que, en esta tesis, se concluye que es el mejor indicador de somnolencia debido, probablemente, a que el sistema de visión desarrollado para su obtención es propietario y no comercial por lo que los resultados obtenidos son mejores que en otros trabajos referenciados del estado del arte [Friedrichs & Yang, 2010a, Caterpillar, 2008].

7.2. Trabajos futuros

En el desarrollo de esta tesis se han abierto nuevas líneas de investigación y de trabajo que podrían ser exploradas en el futuro.

- **Indicador de fiabilidad para medidas relacionadas con el conductor.** En condiciones reales, sería posible estimar la fiabilidad de las medidas relacionadas con la cara del conductor, evaluando los posibles deslumbramientos en la imagen, o premiando la posición de mirada al frente de la cabeza.
- **Mejorar la fiabilidad del cálculo del PERCLOS para usuarios con gafas.**
- **Modelo del filtro de Kalman.** La predicción de la posición de los ojos podría mejorarse si se abandonase el modelo de velocidad rectilínea y uniforme y se adoptase otro más cercano al movimiento real como podría ser un movimiento circular y uniforme, o considerando aceleración constante durante el periodo de muestreo.
- **Uso de clasificadores SVM.** La aplicación de técnicas de apariencia, de los clasificadores SVM, pueden ser complementarias a las técnicas de características empleadas aquí para la evaluación de la apertura del ojo. También podría ser utilizada la clasificación SVM para realizar medidas precisas de los parpadeos del conductor.
- **Codificación de las señales relacionadas con la conducción.** Se podrían codificar las señales relacionadas con la conducción basándose en tramos de un minuto, de modo que el valor que se asignara fuese la diferencia entre el valor máximo y mínimo ocurridos en el citado tramo de tiempo. De esta forma, en fase de somnolencia los valores de la posición serían más elevados que en la fase de vigilia.
- **Estudio frecuencial de los indicadores de conducción.** Sería necesario realizar un estudio en el dominio de la frecuencia de los indicadores de conducción para encontrar posibles parámetros relacionados con el estado de somnolencia del conductor.
- **Evaluación del ángulo de guiñada en ensayos reales.** Como se ha demostrado, los indicadores relacionados con el error del ángulo de guiñada obtienen buenos resultados al clasificar el estado del conductor, por tanto, sería interesante que trabajos relacionados con la información de GPS y trayectorias en mapas pudieran evaluar con precisión el error del ángulo de guiñada del vehículo en condiciones reales.
- **Clasificación por medio de otros métodos.** Aunque se han hecho pruebas con otros tipos de redes para clasificar, y se ha optado por la red perceptrón multicapa, queda abierto el trabajo a otros tipos de redes y a otros métodos de clasificación.

-
- **Ensayos en condiciones reales con somnolencia.** Siempre es complicado realizar estudios de somnolencia cuando se está conduciendo por el riesgo que conlleva quedarse dormido, aunque este riesgo puede minimizarse con un conductor adicional que controle al vehículo cuando se pierda el control por somnolencia.
 - **Sistema comercial.** El algoritmo desarrollado de evaluación del PERCLOS podría ser comercializado añadiendo la implementación de la fusión con información relacionada con el error del ángulo de guiñada, en tiempo real.

Conclusions and further works

Inattention due to fatigue or drowsiness while driving is well known as a fundamental factor that causes traffic accidents. Many drivers combine sleepiness with excessive periods of driving, that causes a serious conflict between physiological needs and social or professional activities of the driver. As consequence, the key to reach accident prevention the understanding of the human limits about fatigue, drowsiness and sleepness. Fatigue and drowsiness in drivers are important factors to be considered due to the amount of traffic accidents they cause. Because of that, during the last decade, driver drowsiness has received an increase of attention on the part of the scientific community, the car companies and society in general became more aware of this problem.

One of the goals of this Thesis is the implementation of a computer vision system that estimates the driver's eye aperture using object recognition techniques and characterization of the parameters from the detected objects. It works even if either realistic simulation or real conditions are present, with awake and sleepy users during the tests. The system developed can be run in real time and warn the user when a drowsiness state is detected.

The eye closure estimation is automatic and user-independent. The algorithm is based on localization and characterization of driver's eyes while they are driving, using appearance-based techniques like Viola and Jones [Viola & Jones, 2001] detector modified by Rainer Lienhart and Jochen Maydt [Lienhart & Maydt, 2002] in order to detect not only the face but the eyes on the image. In order to reduce eye misdetection, a K-means classifier is applied to the candidates given by the detection algorithm. For the purpose of eye closure and position measurement some adaptive filtering is performed before applying a Gaussian model. The detected position at time k is the input of a linear Kalman filter to predict the position in time $k+1$, that makes detection easier and reduces computation time.

The three main problems that appear when detecting eyes on the image are: illumination variation, time to locate them and the complex evaluation process of eye closure [Yuille et al., 1989b]. This work is based on effective techniques that can deal with illumination changes. Projective integration and Gaussian modelling is applied for the purpose of estimating the eye closure and its position at a framerate of 30 fps. This way, unlike other systems of the state-of-the-art [Ryan et al., 2008, D'Orazio et al., 2007a, Senaratne et al., 2007, Suzuki et al., 2006b], a real-time robust system is achieved, tested under real operation conditions in long evaluation trials. On the other hand, unlike other closed commercial systems [SeeingMachines, 2011, SmartEye, 2011] the one presented here is opensource, and it can be clearly seen that results are at least equal, or even better, than the commercial ones, providing technical details and an exhaustive testbench to test its operation features.

On the other hand, parameters given by ADAS systems and related to vehicle information have been evaluated. Such parameters are: relative lane-vehicle position, steering wheel angle and heading angle error to determine signals that can be correlated to drowsiness. Obtained indicators are the standard deviation of the previous variables, root mean square error of the relative lane-vehicle position, time to line crossing, and information about quick steering wheel movements, as it is gathered in [Sandberg et al., 2011] applied to other dataset.

The whole system has been evaluated either in simulation, under controlled illumination conditions, or real operation, under illumination conditions out of control. The acquired database is over 20 users, having acquiring hundreds of driving operation hours, and videos in which the face of each driver can be located, as well as a ground truth drowsiness signal and other signals related to the driving operation. In both scenarios, the binary classifier recall rate is greater than 90 %, using as input the eye closure information, improving other results of the state-of-the art [Friedrichs & Yang, 2010a, Caterpillar, 2008]

The proposed methodology has been validated. It consists of analyzing the driver's behavior in a realistic simulator, which is an environment under control, and the validation takes place in real driving conditions. Results obtained in both, simulation and real conditions are approximately the same apart from the detection rate, that is decreased a bit in real operation, due to the signal to noise ratio, inherent to the input signal.

The developed algorithm reaches high recall rates when detecting drowsiness. This information, combined to other vehicle signals like: standard deviation of the relative vehicle-lane position, steering wheel angle and heading error angle, can be used for the experts to analyze the behaviour of a sleepy driver.

In the last decade, there has been many research investing on different aspects of traffic safety by means of ADAS: interpreting traffic signs, evaluating distance to the preceding vehicle in order to avoid rear crashes in traffic jams, detecting pedestrians in day and night conditions to avoid accidents, etc. However, systems related to driver's state were held in a turning point when trying to determine characteristics such as inattention or drowsiness level. It has been the introduction of computer vision techniques what really boosted the progression and innovation, and finally will enable them to introduce in the automotive field. The final goal is preventing the user to keep driving when a specific drowsiness level has been reached. There is still a long way for this systems to be finally included in commercial vehicles due to the current high false positive rates, warning the user so many times, generating more distractions and creating discomfort in the user. The present Thesis contributed to validated this hypothesis.

7.3. Main Contributions

Taking into account the analysis of the results obtained in the previous chapters, the main contributions of this Thesis are the following:

1. **There is no golden rule about measuring drowsiness while driving.** There are several promising signals, however, all of them have some limitations due to the measurement noise. Therefore, in some cases there is no clear relation between those signals and drowsiness, and this effect is even greater in real driving conditions.
2. **Implementation of a computer vision system.** An opensource computer vision system has been implemented. It is capable of evaluating PERCLOS signal in real time, in a realistic simulator and in real operation conditions. Results obtained with our system are equal or better than other commercial ones [SeeingMachines, 2011, SmartEye, 2011], being more flexible in its working.
3. Even though **three classification levels** are taken initially, it has been demonstrated that experimentally is better using just two of them.
4. **Robust PERCLOS estimation algorithm development.** The implemented algorithm, based on a Gaussian model, does not require fixed thresholds nor calibration to work properly. In addition, it can deal with illumination variations and it has been tested that it is able to work with outdoor conditions and different users that do not wearing

glasses. Therefore, the algorithm can be integrated in a commercial on board system. This system is low cost, compared to other systems in the state-of-the-art [SmartEye, 2011], [Ryan et al., 2008], [Zhu & Ji, 2005], because it is monocular. A Kalman filter is used to improve the tracking and prevent detection failures.

5. **High detection rate of the PERCLOS signal.** The correlation between ground truth and the eye closure has been shown to be is very high, so that, when using a binary classifier, it can reach a recall rate higher than 90 % in simulation and real conditions. This indicator, related to facial characteristics, is the one that gives better results, even better than the ones obtained using car-related signals. This figure is higher than in other important works of the state of the art [Sandberg, 2011]. This is different could be justified by the better working of our vision system and the ground truth generation method. It has been shown that experts mainly base their decisions in PERCLOS because there is a high correlation between this visual indicator and the driver drowsiness.
6. **Driving-related signal analysis.** Signals related to driving do not appear to have high correlation with drowsiness, although it does not seem to happen when we take some derived signs, as the described in [Sandberg et al., 2011]. After some optimization process, the recall rate can reach values over 75 %. The yaw angle is the only one independent from the vehicle trajectory, so it yields good results, with recall rate over 80 %. It does not happen when using other indicators like: lane-vehicle position or steering wheel angle, because all actions required to keep the vehicle in the lane are hiding the drowsiness effects.
7. **Optimization and signal merging** Fusing indicators by means of stochastic optimization techniques, and using a multilayer perceptron neural network, better results are obtained than using only one indicator. Merging PERCLOS with the standard deviation of the relative vehicle position on the lane, a recall rate of 99.5 % can be obtained, which is almost 10 % better than taking PERCLOS signal alone. This fusion has been shown to be the best for a driver sleepiness system.
8. **Simulation vs real results.** During the simulation, the best possible combinations are the following: the fusion of PERCLOS and the standard deviation of the heading angle; and PERCLOS and the optimized MSE indicator. During real conditions, the best possible combination is the last one. If heading angle had had taken into account during real operation, the recall rate would have been improved. Therefore, the obtained results guarantee the methodology used, and can be easily extrapolated from the realistic simulator to real driving operation. Conclusions obtained using this methodology are valid for real conditions even though the detection rate is lower since the input signals are noisier.
9. **Results obtained compared with other related works.** Results are, in general terms, in line with other important works about this subject [Sandberg, 2011] except in the consideration of the PERCLOS, as the best indicator for a driver drowsiness system, maybe due to we use our own vision system and not a commercial one. Then, results are better than other important works of the state-of-the-art [Friedrichs & Yang, 2010a, Caterpillar, 2008].

7.4. Further work

During the development of this Thesis, some new research lines have been opened and could be studied in the future.

- **Confidence indicator of driver-related measurements** In real operation conditions, it would be possible to estimate the confidence given by the driver's face measurements,

evaluating possible glares on the image or encouraging gaze direction when looking at the front.

- **Users wearing glasses.** In this case our proposal does not work correctly. More research effort will be necessary to implement an algorithm capable of working with these users.
- **Kalman filter model.** Eye position estimation could be improved by changing the linear model implemented. There are some models that are better to estimate eye movements, like circular movements or considering a motion model with constant acceleration during the sampling period.
- **SVM classifier.** The application of appearance-based techniques on SVM classifiers could complement the feature based one implemented in this thesis, in order to improve the eye closure. In addition, SVM classifier can be used to improve eye blinking detection.
- **Driving signal coding.** Driving related signals can be coded in windows of 1 minute length, so that the coded value would be the difference between the maximum and the minimum in that period of time. Doing this, during drowsiness state, the obtained position values would be higher than the awakesness state.
- **Frequency study of driving signals.** It would be necessary to study of the signals given by the car driving operation and in the frequency domain, in order to find some correspondences between the drowsiness state.
- **Heading angle evaluation in real operation conditions.** As it has been demonstrated, signals related to heading angle error yields good results when classifying driver's state. As a consequence it would be very interesting if works related to GPS and trajectories can evaluate accurately heading error in a real vehicle.
- **Other classification methods.** Even thought other neural network topologies have been tested, and the final implementation uses a multilayer perceptron network, the problem is still open and other network topologies and classification algorithms could be applied.
- **Commercial system.** The developed drowsiness evaluation algorithm could be commercialized as soon as the merging of PERCLOS with heading angle error is integrated in the whole system. We propose to implement a commercial system using this indicator fusion.
- **Real operation conditions with drowsiness.** It is clearly seen that real drowsiness tests are difficult to perform and risky, even though a team drives the vehicle when it would be out of control. More long tests in real conditions should be necessary to validate our proposal.

Bibliografía

- [Agustin et al., 2006] Agustin, J. S., Villanueva, A., & Cabeza, R. (2006). Pupil brightness variation as a function of gaze direction. In *Proceedings of the 2006 symposium on Eye tracking research applications*, ETRA '06 (pp. 49–49). New York, NY, USA: ACM.
- [Åkerstedt, 1995] Åkerstedt, T. (1995). Work hours, sleepiness and accidents introduction and summary. *Journal of Sleep Research*, 4, 1–3.
- [Åkerstedt & Gillberg, 1990] Åkerstedt, T. & Gillberg, M. (1990). Subjective and objective sleepiness in the active individual. *Int J Neurosci*, 52(1-2), 29–37.
- [Åkerstedt et al., 2001] Åkerstedt, T., Kecklund, G., & Hörte, L. (2001). Night driving, season, and the risk of highway accidents. *Slee*, 24, 401–406.
- [Åkerstedt et al., 2005] Åkerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2005). Impaired alertness and performance driving home from the night shift: a driving simulator study. *Journal of Sleep Research*, 14(1), 17–20.
- [Alcantarilla et al., 2008] Alcantarilla, P., Bergasa, L., Jimenez, P., Sotelo, M., Parra, I., Fernandez, D., & Mayoral, S. (2008). Night time vehicle detection for driving assistance lightbeam controller. In *Intelligent Vehicles Symposium, 2008 IEEE* (pp. 291–296).
- [Bala, 1997] Bala, L.-P. T. K. L. J. (1997). Automatic detection and tracking of faces and facial features in video sequences. *International Picture Coding Symposium (PCS) 1997, Berlin*, 0.
- [Bergasa et al., 2006] Bergasa, L., Nuevo, J., Sotelo, M., Barea, R., & Lopez, M. (2006). Real-time system for monitoring driver vigilance. *Intelligent Transportation Systems, IEEE Transactions on*, 7(1), 63–77.
- [Berka et al., 2007] Berka, C., Levendowski, D. J., Lumicao, M. N., Yau, A., Davis, G., Zivkovic, V. T., Olmstead, R. E., & Tremoulet, P. D. C. P. L. (2007). Eeg correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviation, Space, and Environmental Medicine*, 78(1), 231–244.
- [Beymer & Flickner, 2003] Beymer, D. & Flickner, M. (2003). Eye gaze tracking using an active stereo head. In *Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on*, volume 2 (pp. II–451–8 vol.2).
- [Bos et al., 2008] Bos, J., Bles, W., & Groen, E. (2008). A theory on visually induced motion sickness. *Displays*, 29(2), 47–57.
- [Boyraz et al., 2008] Boyraz, P., Acar, M., & Kerr, D. (2008). Multi-sensor driver drowsiness monitoring. *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, 222(11), 2041–2062.
- [Breiman, 2001] Breiman, L. (2001). Random forests. In *Machine Learning* (pp. 5–32).

- [Bretzner & Krantz, 2005] Bretzner, L. & Krantz, M. (2005). Towards low-cost systems for measuring visual cues of driver fatigue and inattention in automotive applications. *Vehicular Electronics and Safety. IEEE International Conference*, (pp. 161–164).
- [Campagne et al., 2004] Campagne, A., Pebayle, T., & Muzet, A. (2004). Correlation between driving errors and vigilance level: influence of the driver's age. *Physiology & Behavior*, 80(4), 515–524.
- [Caterpillar, 2008] Caterpillar (2008). *Operator Fatigue Detection Technology Review*. Technical report.
- [CEIT, 2011] CEIT (2011). Centro de estudios e investigaciones técnicas de gipuzkoa. <http://www.ceit.es> Last view, July 2011.
- [Commission, 2003] Commission, E. (2003). Communication From the Commission - European Road Safety Action Programme - Halving the number of road accident victims in the European Union by 2010: A shared responsibility. online. <http://tinyurl.com/mtad98>.
- [Commission, 2009] Commission, E. (2009). Communication on the future on transport, adopted by the commission on 17th june 2009. online. http://ec.europa.eu/transport/strategies/2009_future_of_transport_en.htm.
- [Connor et al., 2002] Connor, J., Norton, R., Ameratunga, S., Robinson, E., I.Civil, Dunn, R., Bailey, J., & Jackson, R. (2002). Driver sleepiness and risk of serious injury to car occupants: Population based control study. *British Medical Journal*, 324, 1125–1129.
- [Cristinacce & Cootes, 2006] Cristinacce, D. & Cootes, T. (2006). : (pp. 929–938).
- [Croo et al., 2001] Croo, H. D., Bandmann, M., Mackay, G. M., Rumar, K., & Vollenhoven, P. (2001). *The role of driver fatigue in commercial road transport crashes*. Technical report, European Transport Safety Council, Brussels.
- [Crowley & Berard, 1997] Crowley, J. L. & Berard, F. (1997). Multi-modal tracking of faces for video communications. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 0, 640.
- [Dinges, 1995] Dinges, D. (1995). An overview of sleepiness and accidents. *Sleep research*, 2, 4–14.
- [D'Orazio et al., 2004] D'Orazio, T., Leo, M., Cicirelli, G., & Distanto, A. (2004). An algorithm for real time eye detection in face images. In *ICPR '04: Proceedings of the Pattern Recognition, 17th International Conference on (ICPR'04) Volume 3* (pp. 278–281). Washington, DC, USA: IEEE Computer Society.
- [D'Orazio et al., 2007a] D'Orazio, T., Leo, M., Guaragnella, C., & Distanto, A. (2007a). A visual approach for driver inattention detection. *Pattern Recognition*, 40(8), 2341–2355. Part Special Issue on Visual Information Processing.
- [D'Orazio et al., 2007b] D'Orazio, T., Leo, M., Guaragnella, C., & Distanto, A. (2007b). A visual approach for driver inattention detection. *Pattern Recognition*, 40(8), 2341–2355. Part Special Issue on Visual Information Processing.
- [Dougherty, 1992] Dougherty, E. (1992). *An Introduction to Morphological Image Processing*.

- [DriverFatigueMonitor, 2011] DriverFatigueMonitor (2011). Driver fatigue, lane management & warning systems. <http://www.driverfatiguemonitor.com/dfm/dfm.html>
Last view, July 2011.
- [Edenborough et al., 2005] Edenborough, N., Hammoud, R., Harbach, A., Ingold, A., Kisacanin, B., Malawey, P., Newman, T., Scharenbroch, G., Skiver, S., Smith, M., Wilhelm, A., Witt, G., Yoder, E., & Zhang, H. (2005). Driver state monitor from delphi. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 2 (pp. 1206–1207 vol. 2).
- [Ersal et al., 2010] Ersal, T., Fuller, H., Tsimhoni, O., Stein, J., & Fathy, H. (2010). Model-based analysis and classification of driver distraction under secondary tasks. *Intelligent Transportation Systems, IEEE Transactions on*, 11(3), 692–701.
- [Eskandarian & Mortazavi, 2007] Eskandarian, A. & Mortazavi, A. (2007). Evaluation of a smart algorithm for commercial vehicle driver drowsiness detection. *2007 IEEE Intelligent Vehicles Symposium*, (pp. 553–559).
- [Eskandarian et al., 2007] Eskandarian, A., Sayed, R., Delaigue, P., Blum, J., & Mortazavi, A. (2007). *Advanced driver fatigue research*. Technical Report FMCSA-RRR-07-001, U.S Department of Transportation Federal Motor Carrier Safety Administration.
- [ESM, 2011] ESM (2011). Investigación en seguridad y factores humanos. <http://www.esm.es>
Last view, July 2011.
- [F. & Richard, 1998] F., D. D. & Richard, G. (1998). *PERCLOS, a valid psychophysiological measure of alertness as assessed by psychomotor vigilance*. Technical Report FHWA-MCRT-98-006, Federal Highway Administration, Office of Motor Carriers, Office of Motor Carrier Research and Standards, Washington, DC.
- [Faber, 2004] Faber, J. (2004). Detection of different levels of vigilance by eeg pseudo spectra. *Neural Network World*, 14(3-4), 285–290.
- [Fairclough & Graham, 1999] Fairclough, S. H. & Graham, R. (1999). Impairment of driving performance caused by sleep deprivation or alcohol: A comparative study. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 41(1), 118–128.
- [Fan et al., 2010] Fan, X., Sun, Y., Yin, B., & Guo, X. (2010). Gabor-based dynamic representation for human fatigue monitoring in facial image sequences. *Pattern Recognition Letters*, 31(3), 234–243.
- [Farid et al., 2006a] Farid, M., Essaili, A. E., Kopf, M., & Bub, H. (2006a). Methods to develop a driver observation system used in an active safety system. In *22. VDI/VW International Conference on Active Safety and Driver Assistance Systems* Wolfsburg, Germany.
- [Farid et al., 2006b] Farid, M., Kopf, M., Bubb, H., & Essaili, A. (2006b). Methods to develop a driver observation system used in an active safety system. *International Conference on Active Safety and Driver Assistance Systems*, (1960).
- [Fatigue & Accidents, 2001] Fatigue, T. R. S. F. T. P. O. A. D. & Accidents, R. (2001). Driver fatigue and road accidents: a literature review and positin paper. The Royal Society for the Prevention of Accidents.
- [Feris et al., 2000] Feris, R. S., de Campos, T. E., Junior, R. M. C., Marcondes, R., & Junior, C. (2000). Detection and tracking of facial features in video sequences. In *Lecture Notes in Artificial Intelligence* (pp. 129–137): Springer-Verlag Press.

- [Foundation for Traffic Safety, 2010] Foundation for Traffic Safety, A. (2010). Asleep at the wheel: The prevalence and impact of drowsy driving.
- [Friedrichs & Yang, 2010a] Friedrichs, F. & Yang, B. (2010a). Camera-based drowsiness reference for driver state classification under real driving conditions. In *Intelligent Vehicles Symposium (IV), 2010 IEEE* (pp. 101–106).
- [Friedrichs & Yang, 2010b] Friedrichs, F. & Yang, B. (2010b). Camera-based drowsiness reference for driver state classification under real driving conditions. In *Intelligent Vehicles Symposium (IV), 2010 IEEE* (pp. 101–106).
- [Friedrichs & Yang, 2010c] Friedrichs, F. & Yang, B. (2010c). Drowsiness monitoring by steering and lane data based features under real driving conditions. *European Signal Processing Conference*, 18, 23–27.
- [Furugori et al., 2005] Furugori, S., Yoshizawa, N., Iname, C., & Miura, Y. (2005). Estimation of driver fatigue by pressure distribution on seat in long term driving. *Review of Automotive Engineering*, 26(1), 53–58.
- [Gander & James, 1999] Gander, P. & James, I. (1999). Investigating fatigue in truck crashes. Wellington School of Medicine and Commercial Vehicle Investigation.
- [Grauman et al., 2001] Grauman, K., Betke, M., Gips, J., & Bradski, G. R. (2001). Communication via eye blinks - detection and duration analysis in real time. *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 1, 1010.
- [Greivenkamp & Guyton, 1994] Greivenkamp, J. E. & Guyton, D. L. (1994). A quantification of the bruckner test for strabismus.
- [Hallinan, 1991] Hallinan, P. W. (1991). Recognizing human eyes. *Geometric Methods in Computer Vision*, (pp. 212–226).
- [Hansen et al., 2002a] Hansen, D., Hansen, J., Nielsen, M., Johansen, A., & Stegmann, M. (2002a). Eye typing using markov and active appearance models. In *Applications of Computer Vision, 2002. (WACV 2002). Proceedings. Sixth IEEE Workshop on* (pp. 132–136).
- [Hansen & Hansen, 2006] Hansen, D. W. & Hansen, J. P. (2006). Robustifying eye interaction. *Computer Vision and Pattern Recognition Workshop*, 0, 152.
- [Hansen et al., 2002b] Hansen, D. W., Hansen, J. P., Niels, M., & Stegmann, M. B. (2002b). Eye typing using markov and active appearance models. In *WACV '02: Proceedings of the Sixth IEEE Workshop on Applications of Computer Vision* (pp. 132). Washington, DC, USA: IEEE Computer Society.
- [Hansen & Ji, 2009] Hansen, D. W. & Ji, Q. (2009). In the eye of the beholder: A survey of models for eyes and gaze. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, (1).
- [Hansen & Pece, 2005] Hansen, D. W. & Pece, A. E. C. (2005). Eye tracking in the wild. *Comput. Vis. Image Underst.*, 98, 155–181.
- [Haro et al., 2000] Haro, A., Flickner, M., & Essa, I. (2000). : (pp. 163–168).
- [Heinzmann et al., 2008] Heinzmann, J., Tate, D., & Scott, R. (April 2008). Using technology to eliminate drowsy driving. *SPE International Conference on Health, Safety, and Environment in Oil and Gas Exploration and Production*.

- [Herpers et al., 1996] Herpers, R., Michaelis, M., Lichtenauer, K.-H., & Sommer, G. (1996). Edge and keypoint detection in facial regions. *Automatic Face and Gesture Recognition, IEEE International Conference on*, 0, 212.
- [Hoddes et al., 1973] Hoddes, E., Zarcone, V., Smythe, H., Phillips, R., & Dement, W. C. (1973). Quantification of sleepiness: A new approach. *Psychophysiology*, 10(4), 431–436.
- [Huang & Wechsler, 1999] Huang, J. & Wechsler, H. (1999). Eye detection using optimal wavelet packets and radial basis functions (rbfs). *International Journal of Pattern Recognition and Artificial Intelligence*, 13, 1009–1026.
- [Ingre et al., 2006a] Ingre, M., Akerstedt, T., Peters, B., Anund, A., & Kecklund, G. (2006a). Subjective sleepiness, simulated driving performance and blink duration: examining individual differences. *J Sleep Res*, 15(1), 47–53.
- [Ingre et al., 2006b] Ingre, M., Akerstedt, T., Peters, B., Anund, A., Kecklund, G., & Pickles, A. (2006b). Subjective sleepiness and accident risk avoiding the ecological fallacy. *Sleep*, 15(2), 142–148.
- [Ishikawa et al., 2004] Ishikawa, T., Baker, S., Matthews, I., & Kanade, T. (2004). Passive driver gaze tracking with active appearance models. In *In Proceedings of the 11th World Congress on Intelligent Transportation Systems*.
- [Ji & Yang, 2002a] Ji, Q. & Yang, X. (2002a). Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 8(5), 357–377.
- [Ji & Yang, 2002b] Ji, Q. & Yang, X. J. (2002b). Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imaging*, 8(5), 357–377.
- [Jiménez, 2011] Jiménez, P. (2011). *Face Pose Estimation with Automatic 3D Model Creation for a Driver Inattention Monitoring Application*. PhD thesis, University of Alcalá, Department of Electronics.
- [Jovanis et al., 1991] Jovanis, P., Kaneko, T., & Lin, T. (1991). Exploratory analysis of motor carrier accident risk and daily driving patterns. *Transportation Research Record*, 1332(1), 34–43.
- [Kaida et al., 2006] Kaida, K., Takahashi, M., Åkerstedt, T., Nakata, A., Otsuka, Y., Haratani, T., & Fukasawa, K. (2006). Validation of the karolinska sleepiness scale against performance and eeg variables. *Clinical Neurophysiology*, 117(7), 1574–1581.
- [Kecklund & Akerstedt, 1993] Kecklund, G. & Akerstedt, T. (1993). Sleepiness in long distance truck driving: an ambulatory eeg study of night driving. *Ergonomics*, 36(9), 1007–1017.
- [Kim & Ramakrishna, 1999] Kim, K.-N. & Ramakrishna, R. (1999). Vision-based eye-gaze tracking for human computer interface.
- [Klauer et al., 2006] Klauer, S. G., Dingus, T. A., Neale, V. L., , & Sudweeks, J. D. (2006). The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data. *National Highway Traffic Safety Administration, DC, DOT HS*, 810.
- [Krajewski et al., 2009] Krajewski, J., Sommer, D., Trutschel, U., Edwards, D., & Golz, M. (2009). *Steering wheel behavior based estimation of fatigue*, (pp. 118–124).

- [Lexus, 2011] Lexus (2011). Lexus. <http://www.testdriven.co.uk/lexus-ls-600h>
Last view, July 2011.
- [Li et al., 2005] Li, D., Winfield, D., & Parkhurst, D. (2005). Starburst: A hybrid algorithm for video-based eye tracking combining feature-based and model-based approaches. In *Computer Vision and Pattern Recognition - Workshops, 2005. CVPR Workshops. IEEE Computer Society Conference on* (pp.79).
- [Liang et al., 2007] Liang, Y., Reyes, M., & Lee, J. (2007). Real-time detection of driver cognitive distraction using support vector machines. *Intelligent Transportation Systems, IEEE Transactions on*, 8(2), 340–350.
- [Lienhart & Maydt, 2002] Lienhart, R. & Maydt, J. (2002). An extended set of haar-like features for rapid object detection. In *IEEE ICIP 2002* (pp. 900–903).
- [Liu et al., 2009] Liu, C. C., Hosking, S. G., & Lenné, M. G. (2009). Predicting driver drowsiness using vehicle measures: Recent insights and future challenges. *Journal of Safety Research*, 40(4), 239–245.
- [Loy & Zelinsky, 2003] Loy, G. & Zelinsky, A. (2003). Fast radial symmetry for detecting points of interest. *IEEE Trans. Pattern Anal. Mach. Intell.*, 25(8), 959–973.
- [Mahieu, 2009] Mahieu, Y. (2009). *Highlights of the Panorama of Transport*. Technical report, Eurostats.
- [Matsumoto & Zelinsky, 2000] Matsumoto, Y. & Zelinsky, A. (2000). An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *FG '00: Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000* (pp. 499). Washington, DC, USA: IEEE Computer Society.
- [McCartt et al., 1996] McCartt, A. T., Ribner, S. A., Pack, A. I., & Hammer, M. C. (1996). The scope and nature of the drowsy driving problem in new york state. *Accident Analysis & Prevention*, 28(4), 511–517.
- [Mercedes, 2011] Mercedes (2011). Mercedes-benz. <http://www.emercedesbenz.com/Aug08>
Last view, July 2011.
- [Miyaji et al., 2009] Miyaji, M., Kawanaka, H., & Oguri, K. (2009). Driver's cognitive distraction detection using physiological features by the adaboost. In *Intelligent Transportation Systems, 2009. ITSC '09. 12th International IEEE Conference on* (pp. 1–6).
- [Nguyen et al., 2002] Nguyen, K., Wagner, C., Koons, D., & Flickner, M. (2002). Differences in the infrared bright pupil response of human eyes. In *Proceedings of the 2002 symposium on Eye tracking research & applications*, ETRA '02 (pp. 133–138). New York, NY, USA: ACM.
- [Nixon, 1985] Nixon, M. (1985). Eye spacing measurement for facial recognition. In *Proc. SPIE Applications of Digital Image Processing*, volume 575 (pp. 279–283).
- [Nuevo, 2009] Nuevo, J. (2009). *Face Tracking with Active Models for a Driver Monitoring Application*. PhD thesis, University of Alcalá, Department of Electronics.
- [Oron-Gilad et al., 2008] Oron-Gilad, T., Ronen, A., & Shinar, D. (May 2008). Alertness maintaining tasks (amts) while driving. *Accident Analysis & Prevention*, 40(3), 851–860.

- [Otmani et al., 2005] Otmani, S., Pebayle, T., Roge, J., & Muzet, A. (2005). Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. *Physiology & Behavior*, 84(5), 715–724.
- [Papadelis et al., 2007] Papadelis, C., Chen, Z., Kourtidou-Papadeli, C., Bamidis, P., Chouvarda, I., Bekiaris, E., & Maglaveras, N. (September 2007). Monitoring sleepiness with on-board electrophysiological recordings for preventing sleep-deprived traffic accidents. *Clinical Neurophysiology*, 118(9), 1906–1922.
- [Patel & Markey, 2005] Patel, A. C. & Markey, M. K. (2005). : (pp. 581–589).
- [Paul et al., 2005] Paul, A., Boyle, L., Tippin, J., & Rizzo, M. (2005). : University of Iowa, Iowa City.
- [Pérez et al., 2003] Pérez, A., Córdoba, M. L., García, A., Méndez, R., Muñoz, M. L., Pedraza, J. L., & Sánchez, F. (2003). A precise eye-gaze detection and tracking system. In *International Conference in Central Europe on Computer Graphics and Visualization*.
- [Philip et al., 2005a] Philip, P., Sagaspe, P., Moore, N., Taillard, J., Charles, A., Guilleminault, C., & Bioulac, B. (2005a). Fatigue, sleep restriction and driving performance. *Accident Analysis and Prevention*, 37(3), 473–478.
- [Philip et al., 2005b] Philip, P., Sagaspe, P., Taillard, J., Valtat, C., Moore, N., Akerstedt, T., Charles, A., & Bioulac, B. (2005b). Fatigue, sleepiness, and performance in simulated versus real driving conditions. *Sleep*, 28(12), 1511–1516.
- [Ranney et al., 2001] Ranney, T. A., Mazzai, E., Garrott, R., & Goodman, M. J. (2001). *NHTSA Driver Distraction Research: Past, Present, and Future*. Technical report, NHTSA.
- [Reinders et al., 1996] Reinders, M., Koch, R., & Gerbrands, J. (1996). Locating facial features in image sequences using neural networks. In *Automatic Face and Gesture Recognition, 1996., Proceedings of the Second International Conference on* (pp. 230–235).
- [Ryan et al., 2008] Ryan, W. J., Duchowski, A. T., & Birchfield, S. T. (2008). Limbus/pupil switching for wearable eye tracking under variable lighting conditions. In *Proceedings of the 2008 symposium on Eye tracking research & applications*, ETRA '08 (pp. 61–64). New York, NY, USA: ACM.
- [Sagaspe et al., 2008] Sagaspe, P., Taillard, J., Åkerstedt, T., Bayon, V., Espié, S., Chaumet, G., NewAuthor7, & Philip, P. (2008). Extended driving impairs nocturnal driving performances. *PLoS ONE*, 3(10).
- [Sandberg, 2011] Sandberg, D. (2011). *Detecting Driver Sleepiness*. PhD thesis, Department of Applied Mechanics, Chalmers University of Technology.
- [Sandberg et al., 2011] Sandberg, D., Akerstedt, T., Anund, A., Kecklund, G., & Wahde, M. (2011). Detecting driver sleepiness using optimized nonlinear combinations of sleepiness indicators. *Intelligent Transportation Systems, IEEE Transactions on*, 12(1), 97–108.
- [Sandberg & Wahde, 2008] Sandberg, D. & Wahde, M. (2008). Particle swarm optimization of feedforward neural networks for the detection of drowsy driving. In *Neural Networks, 2008. IJCNN 2008. (IEEE World Congress on Computational Intelligence). IEEE International Joint Conference on* (pp. 788–793).

- [Schmidt et al., 2009] Schmidt, E. A., Schrauf, M., Simon, M., Fritzsche, M., Buchner, A., & Kincses, W. E. (2009). Drivers' misjudgement of vigilance state during prolonged monotonous daytime driving. *Accident Analysis & Prevention*, 41(5), 1087–1093.
- [SeeingMachines, 2011] SeeingMachines (2011). Seeingmachines. <http://www.seeingmachines.com> Last view, July 2011.
- [Senaratne et al., 2007] Senaratne, R., Hardy, D., Vanderaa, B., & Halgamuge, S. (2007). Driver fatigue detection by fusing multiple cues. In *Proceedings of the 4th international symposium on Neural Networks: Part II—Advances in Neural Networks*, ISNN '07 (pp. 801–809). Berlin, Heidelberg: Springer-Verlag.
- [Sirohey et al., 2002] Sirohey, S. A., Rosenfeld, A., & Duric, Z. (2002). A method of detecting and tracking irises and eyelids in video. *Pattern Recognition*, 35(6), 1389–1401.
- [SmartEye, 2011] SmartEye (2011). Smart eye. <http://www.smarteye.se> Last view, July 2011.
- [Stein, 1995] Stein, A. C. (1995). Detecting fatigued drivers with vehicle simulators. *Drivers Impairment, Driver Fatigue and Driving Simulation*, (pp. 133–150).
- [Stutts et al., 2003] Stutts, J., Wilkins, J. W., Osberg, J. S., & Vaughn, B. V. (2003). Driver risk factors for sleep-related crashes. *Accident Analysis and Prevention*, 35(3), 321–331.
- [Sub-Group, 2006] Sub-Group, S. (2006). Canadian operational definition of driver fatigue. STRID Sub-Group on Fatigue.
- [Suzuki et al., 2006a] Suzuki, M., Yamamoto, N., Yamamoto, O., Nakano, T., & Yamamoto, S. (2006a). Measurement of driver's consciousness by image processing -a method for presuming driver's drowsiness by eye-blinks coping with individual differences -. In *Systems, Man and Cybernetics, 2006. SMC '06. IEEE International Conference on*, volume 4 (pp. 2891–2896).
- [Suzuki et al., 2006b] Suzuki, M., Yamamoto, N., Yamamoto, O., Nakano, T., & Yamamoto, S. (2006b). Measurement of driver's consciousness by image processing -a method for presuming driver's drowsiness by eye-blinks coping with individual differences. *Systems, Man and Cybernetics*, 4, 2891–2896.
- [Sweeney et al., 1995] Sweeney, M., Ellingstad, V., Mayer, D., Eastwood, D., Weinstein, E., & Loeb, B. (1995). The need for sleep: Discriminating between fatigue-related and nonfatigue-related truck accidents. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 5(1), 1122–1126.
- [Takei & Furukawa, 2005] Takei, Y. & Furukawa, Y. (2005). Estimate of driver's fatigue through steering motion. In *Systems, Man and Cybernetics, 2005 IEEE International Conference on*, volume 2 (pp. 1765–1770 Vol. 2).
- [Thiffault & Bergeron, 2003] Thiffault, P. & Bergeron, J. (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis and Prevention*, 35(3), 381–391.
- [Tian et al., 2000] Tian, Y., Kanade, T., & Cohn, J. F. (2000). Dual-state parametric eye tracking. In *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition 2000*, FG '00 (pp. 110–). Washington, DC, USA: IEEE Computer Society.

- [Torkkola et al., 2004] Torkkola, K., Massey, N., & Wood, C. (2004). Driver inattention detection through intelligent analysis of readilyavailable sensors. In *Intelligent Transportation Systems, 2004. Proceedings. The 7th International IEEE Conference on* (pp. 326–331).
- [UN-ECE, 2007] UN-ECE (2007). *Statistics Of Road Traffic Accidents In Europe And North America*, volume LI. United Nations, Economic Commission For Europe.
- [van Winsum et al., 2000] van Winsum, W., Brookhuis, K. A., & de Waard, D. (2000). A comparison of different ways to approximate time-to-line crossing (tlc) during car driving. *Accident Analysis & Prevention*, 32(1), 47–56.
- [Verwey & Zaidel, 2000] Verwey, W. B. & Zaidel, D. M. (2000). Predicting drowsiness accidents from personal attributes, eye blinks, and ongoing driving behaviour. *Personality and individual differences*, 28(1), 123–142.
- [Viola & Jones, 2001] Viola, P. & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proc. CVPR*, 1, 511–518.
- [Viola & Jones, 2004] Viola, P. & Jones, M. J. (2004). Robust real-time face detection. *Int. J. Comput. Vision*, 57(2), 137–154.
- [Volvo, 2011] Volvo (2011). Volvo car corporation. <http://www.media.volvocars.com>
Last view, July 2011.
- [Wakita et al., 2005] Wakita, T., Ozawa, K., Miyajima, C., Igarashi, K., Itou, K., Takeda, K., & Itakura, F. (2005). Driver identification using driving behavior signals. In *Intelligent Transportation Systems, 2005. Proceedings. 2005 IEEE* (pp. 396–401).
- [Wang et al., 2005] Wang, P., Green, M. B., Ji, Q., & Wayman, J. (2005). Automatic eye detection and its validation. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Workshops - Volume 03* (pp. 164–). Washington, DC, USA: IEEE Computer Society.
- [Wright et al., 2007] Wright, N., Stone, B., Horberry, T., & Reed, N. (2007). A review of in-vehicle sleepiness detection devices. *TRL Limited, Published Project Report PPR157*.
- [Xie et al., 1998] Xie, X., Sudhakar, R., & Zhuang, H. (1998). A cascaded scheme for eye tracking and head movement compensation. *IEEE Transactions on Systems, Man, and Cybernetics, Part A*, 28(4), 487–490.
- [Yang et al., 1998] Yang, J., Stiefelhagen, R., Meier, U., & Waibel, A. (1998). Real-time face and facial feature tracking and applications.
- [Yoo & Chung, 2005] Yoo, D. H. & Chung, M. J. (2005). A novel non-intrusive eye gaze estimation using cross-ratio under large head motion. *Comput. Vis. Image Underst.*, 98, 25–51.
- [Young et al., 1997] Young, T., Blustein, J., Finn, L., & Palta, M. (1997). Sleep-disordered breathing and motor vehicle accidents in a population-based sample of employed adults. *Sleep*, 20(8), 608–613.
- [Yuille et al., 1989a] Yuille, A., Cohen, D., & Hallinan, P. (1989a). Feature extraction from faces using deformable templates. In *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR '89., IEEE Computer Society Conference on* (pp. 104–109).

- [Yuille et al., 1989b] Yuille, A. L., Cohen, D. S., & Hallinan, P. W. (1989b). Feature extraction from faces using deformable templates. In *Computer Vision and Pattern Recognition, 1989. Proceedings CVPR '89., IEEE Computer Society Conference on* (pp. 104–109).
- [Zhong et al., 2007] Zhong, Y.-J., Du, L.-P., Zhang, K., & Sun, X.-H. (2007). Localized energy study for analyzing driver fatigue state based on wavelet analysis. In *Wavelet Analysis and Pattern Recognition, 2007. ICWAPR '07. International Conference on*, volume 4 (pp. 1843–1846).
- [Zhu et al., 2002a] Zhu, Z., Fujimura, K., & Ji, Q. (2002a). Real-time eye detection and tracking under various light conditions. In *Proceedings of the 2002 symposium on Eye tracking research & applications, ETRA '02* (pp. 139–144). New York, NY, USA: ACM.
- [Zhu & Ji, 2005] Zhu, Z. & Ji, Q. (2005). Robust real-time eye detection and tracking under variable lighting conditions and various face orientations. *Comput. Vis. Image Underst.*, 98, 124–154.
- [Zhu et al., 2002b] Zhu, Z., Ji, Q., & Fujimura, K. (2002b). Combining kalman filtering and mean shift for real time eye tracking under active ir illumination. In *Proceedings of the 16 th International Conference on Pattern Recognition (ICPR'02) Volume 4 - Volume 4*, ICPR '02 (pp. 40318–). Washington, DC, USA: IEEE Computer Society.