

Escuela de Ingeniería y Arquitectura



Universidad
Zaragoza

Ingeniería de Telecomunicación
Proyecto Fin de Carrera

*Classification of pedestrians,
vehicles and animals using
automotive radar*

Department of Signal Processing
Department of Communications Engineering
Tampere University of Technology
Tampere, Finland

Autor: Sergio Cebollada Gonzalo
Director: Pavlo Molchanov
Ponente: Santiago Cruz Llanas

Zaragoza, Septiembre de 2013

Agradecimientos

A Pavlo Molchanov, Alexey Vinel, Yevgeni Koucheryavy y Santiago Cruz por su colaboración en el desarrollo de este proyecto.

A mis compañeros de carrera por los buenos momentos vividos a lo largo de estos años.

A mi familia, en especial mis padres, y a mis amigos, por su apoyo y su confianza en los momentos en los que más falta hacían.

Resumen

El interés creciente en los ámbitos de seguridad y prevención de accidentes en carretera ha propiciado en los últimos años una gran evolución en el desarrollo de sistemas de detección automática de obstáculos buscando evitar las colisiones.

En concreto, en países como Finlandia se ha realizado especial hincapié, debido al gran número de accidentes causados por animales (principalmente renos) que se cruzan en las carreteras, causando grandes daños, físicos y económicos. Es por eso que la investigación centrada en evitar colisiones es de suma importancia allí. La Universidad Tecnológica de Tampere ha comenzado a trabajar en uno de estos sistemas, centrando el trabajo en los departamentos de procesado de señal e ingeniería de comunicaciones.

El objetivo de este proyecto es desarrollar un sistema que sea capaz de detectar objetos y, además, clasificarlos distinguiendo tres grupos: peatones, vehículos y animales. Para ello se parte desde cero, lo cual implica que el trabajo abarcará desde la toma inicial de muestras de los posibles objetivos, hasta el desarrollo de todas las etapas posteriores, incluyendo una primera aproximación al futuro clasificador.

Como primer paso, se analizarán varios sistemas de detección (radar, videocámaras, etc.) y se elegirá uno de ellos de forma razonada. Se tomarán muestras de los diferentes objetivos con dicho sistema, y serán analizadas. Para ello se elegirá el método de filtrado y análisis más apropiado. Una vez se tenga una representación gráfica de las diferentes muestras, el segundo paso será la implementación de un decisor y un posterior clasificador, que sea capaz de discriminar entre las tres posibles categorías que se estudian: peatones, vehículos y animales (renos). De nuevo se considerarán varias opciones y se elegirá aquella que pueda proporcionar mejores resultados. Terminados estos pasos, cualquier señal procedente del detector debería ser automáticamente filtrada y analizada, para ser capaz rápidamente de decidir si hay un obstáculo o no, y clasificarlo en alguno de los tres grupos.

El objetivo de este proyecto es que, más adelante, pueda ser aplicado a un sistema capaz de generar una respuesta mecánica en el coche después de haber detectado un obstáculo en la carretera, para evitar posibles colisiones.

Índice General

Índice General.....	vii
Lista de Figuras	ix
Lista de Tablas	xi
Lista de Acrónimos y Abreviaciones	xiii
1. Introducción	1
1.1. Motivación	1
1.2. Descripción del problema	1
1.3. Objetivos del trabajo	3
1.4. Materiales utilizados.....	3
1.5. Contexto.....	3
1.6. Estructura.....	4
2. Estado del arte	5
2.1. Sistemas basados en videocámaras.....	5
2.2. Sistemas radar.....	6
2.2.1. Radares micro-Doppler.....	7
2.3. Conclusiones	10
3. Métodos y Materiales	11
3.1. Sensor radar	11
3.1.1. Especificaciones del radar	12
3.1.2. Recopilación de muestras.....	13
3.2. Unidad de procesamiento de señal.....	13
3.2.1. Edición de las muestras	13
3.2.2. Análisis	14
3.2.3. Filtrado.....	15
3.2.4. Decisión	17
3.2.5. Clasificación	18
4. Resultados	21
4.1. Análisis y Filtrado	21
4.1.1. Vehículos	21
4.1.2. Peatones	24
4.1.3. Renos	27
4.2. Decisión.....	28
4.2.1. No hay objetivo detectado	28

4.2.2. Objetivo detectado.....	29
4.2.3. Distribución de potencia	30
4.3. Clasificación	32
5. Conclusiones	33
5.1. Líneas futuras.....	34
Bibliografía	35
Anexos	37
A. Características de los sistemas basados en videocámaras	39
B. Aspectos matemáticos de la etapa de clasificación	43
C. Scripts en MATLAB.....	47

Lista de Figuras

Figura 1: Configuraciones del sensor para detector peatones y objetos.....	2
Figura 2: Línea temporal de las medidas de protección de peatones.....	5
Figura 3: Visión esquemática del concepto de radar armónico y transpondedores.....	7
Figura 4: Dibujo del típico “walk cycle” humano.....	8
Figura 5: Diagrama espacio-velocidad simulado para el cuerpo (azul) y para las piernas (verde y rojo) de un peatón andando.....	8
Figura 6: Modelo de puntos de reflexión para un peatón en movimiento.....	9
Figura 7: Modelo de puntos de reflexión para un vehículo con movimiento longitudinal.....	9
Figura 8: Modelo de puntos de reflexión para un vehículo con movimiento lateral.....	9
Figura 9: Representación esquemática del sistema.....	11
Figura 10: Ejemplo de espectrograma.....	14
Figura 11: Ejemplo de espectrograma de ruido.....	16
Figura 12: Densidad espectral de potencia media de ruido.....	17
Figura 13: Representación esquemática del método SVM.....	19
Figura 14: Espectrograma de un coche en movimiento alejándose de nosotros (antes de filtrar).....	21
Figura 15: Espectrograma de un coche moviéndose linealmente (antes de filtrar).....	22
Figura 16: Espectrograma de un coche en movimiento alejándose de nosotros (tras sustracción espectral con Matlab®).....	23
Figura 17: Espectrograma de un coche moviéndose linealmente (tras sustracción espectral con Matlab®).....	23
Figura 18: Espectrograma de una persona yendo hacia nosotros (antes de filtrar).....	24
Figura 19: Espectrograma de una persona alejándose y luego acercándose (antes de filtrar).....	25
Figura 20: Espectrograma de una persona moviéndose hacia nosotros (tras sustracción espectral con Matlab®).....	26
Figura 21: Espectrograma de una persona alejándose y luego acercándose (tras sustracción espectral con Matlab®).....	26
Figura 22: Espectrograma de un reno (antes de filtrar).....	27
Figura 23: Espectrograma de un reno (tras sustracción espectral con Matlab®).....	27
Figura 24: Espectrograma de varias muestras sin objetivo y sus potencias.....	28
Figura 25: Espectrograma de varias muestras con diferentes objetivos y sus potencias.....	29
Figura 26: Histograma de la distribución de potencia para ambos casos.....	30
Figura 27: Evolución de las probabilidades de detección, falsa alarma y pérdida.....	31
Figura 28: Certificación estéreo entre luz visible e imágenes térmicas.....	40
Figura 29: Integración del sistema.....	41
Figura 30: Aplicación del método HOG para identificación de renos.....	42
Figura 31: Ejemplo de margen de separación entre clases.....	45
Figura 32: Diferentes tipos de márgenes de separación.....	45

Lista de Tablas

Tabla 1: Comparación entre diferentes modalidades de sensor para protección de peatones...	2
Tabla 2: Especificaciones del radar.....	12
Tabla 3: Base de datos de muestras.....	13
Tabla 4: Resultados de probabilidad para diferentes umbrales de decisión.....	31
Tabla 5: Porcentaje de precisión del clasificador para los diferentes objetivos.....	32

Lista de Acrónimos y Abreviaciones

VRU	Vulnerable Road User
FOV	Field Of View
IR	Infra Red
PMD	Photonic Mixing Device
IV	Intelligent Vehicle
ITS	Intelligent Transportation System
EU	European Union
CCD	Charge-Coupled Device
SVM	Support Vector Machine
DVC	Deer-Vehicle Crash
HOG	Histogram of Orientated Gradient
SARH	Semi-Active Radar Homing
US	United States
CW	Continuous-Wave
RF	Radio Frequency
EIRP	Equivalent Isotropically Radiated Power
IF	Intermediate Frequency
STFT	Short-Time Fourier Transform
SNR	Signal-to-Noise Ratio
AGC	Automatic Gain Control

1. Introducción

1.1. Motivación

Un informe de la Organización Mundial de la Salud describe los accidentes de tráfico como una de las principales causas de muerte y lesiones en todo el mundo, alcanzando una cifra estimada de 1,2 millones de muertes y 50 millones de lesiones. En algunos países, una gran mayoría de los muertos no son los ocupantes de los vehículos, sino los demás usuarios de la carretera (VRUs, Vulnerable Road Users).

Además, en algunos países como Finlandia, muchos de los accidentes son causados por animales que se cruzan en la carretera, como por ejemplo renos o alces. De acuerdo con uno de los últimos informes del Instituto de Información de Seguros, hay aproximadamente 500.000 colisiones con renos cada año. Eso se traduce en más de 100 muertes y miles de heridos. El coste para la industria de las aseguradoras es de aproximadamente \$2000 por accidente. El reciente aumento de la población de renos ha aumentado el número de estos accidentes.

La protección de los usuarios de la vía es una tarea continua en los nuevos sistemas de seguridad implantados, con el objetivo de reducir el número de víctimas en accidentes de carretera (1). Cada conductor es responsable de su propia seguridad y la del resto de pasajeros. Todos tenemos la capacidad de distinguir entre vehículos, peatones y animales que se cruzan en la carretera, y sabemos cómo reaccionar (la reacción puede ser distinta en función de los distintos objetivos), pero a veces no es suficiente. El cansancio, las distracciones y la baja velocidad de reacción pueden hacernos tomar una decisión incorrecta o tardía. De hecho, en el 90% de los accidentes la causa está relacionada con un error humano. Es por eso que en los últimos años, muchos fabricantes de automóviles están trabajando duro para implantar en sus coches sistemas que puedan ayudarnos con esta tarea, y puedan tomar la decisión correcta por ellos mismos.

1.2. Descripción del problema

El objetivo de este proyecto es desarrollar un sistema que sea capaz de procesar la señal obtenida por un sensor y ser capaz de decidir si hay un objetivo interponiéndose en la vía, así como clasificarlo en cualquiera de los tres grupos.

Dos sistemas serán considerados principalmente. Por un lado, sistemas ópticos, incluyendo videocámaras, y por otro, sistemas radar como el que vemos en la Fig. 1. Hay varios estudios y pruebas realizadas sobre detección de objetos con ambos sistemas. Aun así, la idea de construir un sistema como éste, incluyendo un clasificador y trabajando con diferentes tipos de objetivos al mismo tiempo, está todavía bastante poco explotada. Y es ahí donde reside el gran interés en trabajar en ello.

En general, es conocido que los sistemas basados en imágenes son altamente influenciados por las condiciones ambientales, como la lluvia, nieve o niebla. Además,

tienen problemas para realizar detecciones correctas durante la noche, debido a la oscuridad. Por otro lado, tienen mejor resolución angular que los sistemas radar.

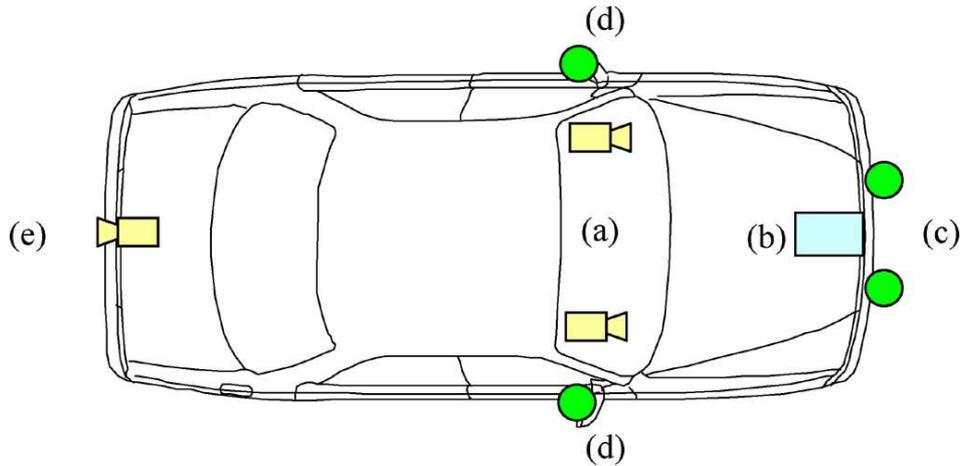


Fig. 1: Configuraciones del sensor para detectar peatones y objetos (2). (a) Línea base estereo para objetos distantes. (b) RADAR para medidas de distancia para un campo de visión (FOV) estrecho. (c) Línea base estrecha para FOV ancho, para objetos cercanos. (d) Omnicámaras para monitorizar puntos ciegos. (e) Cámara inversa monocular.

Pero, como punto de partida, nuestro trabajo será estudiar en profundidad ambos sistemas, comparar sus ventajas y desventajas y prestar atención a sus progresos y la calidad de sus resultados para este tipo de aplicaciones. Es importante conocer en qué punto están los dos sistemas, para aprovecharnos a la hora de decidir cuál es el más apropiado para nosotros, y saber cómo continuar las investigaciones previas y modificarlas para completar nuestro proyecto.

Sensor type	Field of view	Angular resolution	Detection range	Range resolution	Illumination	Hardware cost	Algorithmic complexity
Rectilinear camera	Med.	Med./High	Low/Med.	Med.	Passive reflective, needs ambient light	Low	High
Omni camera	Large	Low/Med.	Low	Low	Passive reflective, needs ambient light	Med.	High
Near IR	Med.	Med./High	Med.	Med.	Active, works in dark	Low	High
Thermal IR	Med.	Low/Med.	Low/Med	Low	Emissive, works in dark	High	Med.
PMD sensor	Med.	Low	Med.	Low/Med.	Modulated light source	Med.	Med.
RADAR	Small	Low	High	High	Active, works in dark, rain, fog.	Med.	Low
LASER scanner	Large	Med.	Med.	High	Active, works in dark	High	Low

Tabla 1: Comparación entre diferentes modalidades de sensor para protección de peatones

1.3. Objetivos del trabajo

Para una mejor organización, se han establecido diferentes objetivos para estructurar el trabajo a llevar a cabo, descritos a continuación:

1. Análisis de los diferentes sistemas usados en seguridad vial (cámaras, radar, etc.)
2. Estudio del sistema que será utilizado finalmente para tomar las muestras (radar de onda continua).
3. Toma de muestras de peatones, vehículos y renos.
4. Edición de audios y vídeos correspondientes a las muestras tomadas, para la extracción de las señales buscadas.
5. Tratamiento de la señal obtenida por el radar en cada caso, de forma separada.
6. Desarrollo de un algoritmo que sea capaz de decidir si la señal detectada corresponde a un obstáculo o no.
7. Desarrollo de un primer algoritmo para clasificar dichas muestras obtenidas.
8. Evaluación de dicho clasificador, en términos de probabilidad de clasificación correcta y, en la medida de lo posible, perfeccionamiento de dicho clasificador.

1.4. Materiales utilizados

En este apartado se enumeran los diferentes materiales y herramientas utilizados a lo largo de todo el proyecto. En los siguientes capítulos se tratará con más detalle su utilización:

- Radar de onda continua. Modelo IPS-146 de InnoSent, a 24 GHz.
- Videocámara de 1.3 mpx.
- Grabador de audio con tasa de muestreo 44100 Hz, 16 bit ADC.
- Programa de mezclado de audio y vídeo.
- Programa de editado de vídeo Virtualdub[®].
- VLC Media Player[®] para la extracción de audio de los vídeos.
- Matlab[®] para la construcción de los algoritmos de análisis, filtrado, decisión y clasificación.

1.5. Contexto

Este trabajo se ha desarrollado a lo largo de 7 meses en la Universidad Tecnológica de Tampere (Finlandia), gracias al programa de intercambio de estudiantes ERASMUS. Se enmarca en el área de investigación, llevada a cabo entre los departamentos de Tratamiento de Señal y de Ingeniería de Comunicaciones.

1.6. Estructura

A continuación se describe brevemente el contenido de los capítulos restantes que conforman esta memoria:

- El capítulo 2 presenta una visión del estado del arte en el que se encuentran las diferentes tecnologías relativas a la detección de objetos que se cruzan en la vía.
- El capítulo 3 muestra las principales herramientas que se utilizan en este trabajo, así como la metodología seguida para su desarrollo.
- El capítulo 4 muestra los resultados obtenidos en cada una de las fases del trabajo.
- Por último, el capítulo 5 expone las conclusiones extraídas de este proyecto, así como futuras líneas de trabajo.

2. Estado del arte

Parte del trabajo desarrollado en este PFC consiste en el análisis de trabajos similares al que se va a desarrollar, con el objetivo de valorar el sistema a utilizar más adelante. Es por eso, que en este apartado se muestran algunos artículos que explican el estado del arte en el que se encuentran las principales tecnologías aplicadas en casos parecidos.

2.1. Sistemas basados en videocámaras

Mejorar la comodidad y seguridad del conductor y los ocupantes de un vehículo es la principal motivación en el desarrollo y la innovación asociada a los llamados “vehículos inteligentes” (IVs, Intelligent Vehicles) y los “sistemas de transporte inteligentes” (ITSs, Intelligent Transportation Systems).

Ha habido numerosos estudios y programas basados en el seguimiento y la detección, muchos de ellos centrados en los peatones. La seguridad de los peatones puede ser mejorada en varios pasos, como se muestra en la Fig. 2. Estas mejoras pueden ser complementadas con sistemas que detecten a los peatones y puedan prevenir los accidentes, bien alertando al conductor, bien realizando maniobras de frenado de forma automática. En los casos donde el accidente no pueda ser prevenido, incorporar al vehículo dispositivos para mitigar la colisión puede reducir el impacto causado en el peatón. Las mejoras en la infraestructura para reducir los accidentes relacionados con peatones pueden ser divididas en tres categorías: control de velocidad, separación vehículo-peatón y medidas para aumentar la visibilidad de los peatones. Por supuesto, todas ellas aplicables también a otros vehículos, animales, ciclistas, y cualquier otro objeto que pueda aparecer en la vía.

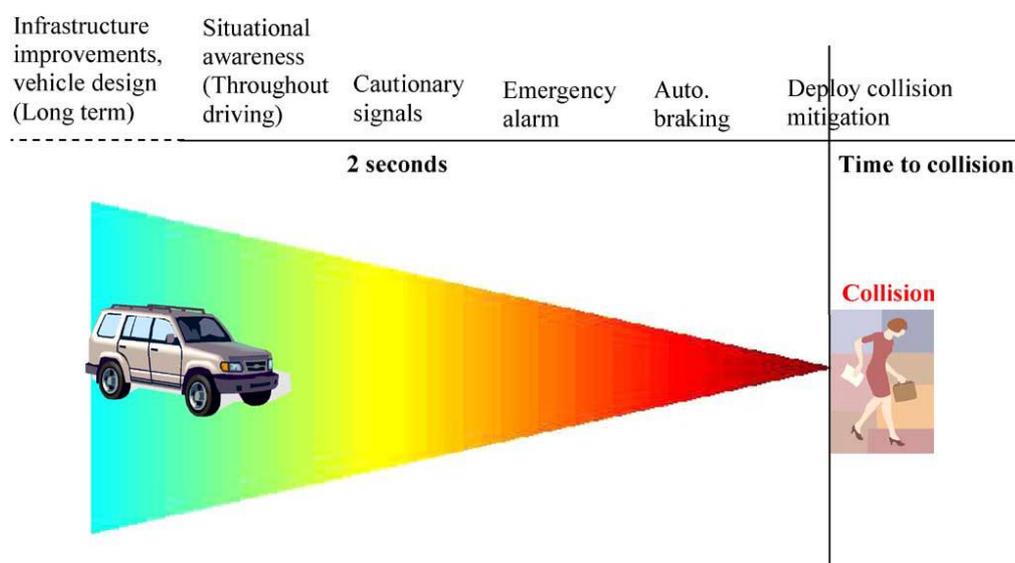


Fig. 2: Línea temporal de las medidas de protección de peatones (2). El tiempo de reacción para una distancia dada decrece con la velocidad.

El programa Pedsmart (3) tiene el objetivo de aplicar la tecnología ITS para mejorar la seguridad de los peatones. Han desarrollado varios dispositivos que detectan a los peatones que cruzan o están esperando, así como los motoristas. También han desarrollado un software llamado “Herramienta de Análisis de choques con Peatones y Ciclistas” para analizar las interacciones entre peatones, ciclistas y vehículos a motor.

La Unión Europea ha conducido varios proyectos en colaboración con diversas industrias e institutos de investigación para el desarrollo de diferentes sistemas en general, y sistemas de seguridad de peatones en particular. El proyecto PReVENT (4) trabaja en el desarrollo de tecnologías de seguridad que ayudan a los conductores a prevenir o mitigar los efectos de un accidente, utilizando análisis de los alrededores, así como del estado del conductor basado en radares. En particular, el subproyecto COMPOSE se centra en la detección de peatones, ciclistas y otros vehículos usando datos provenientes de sensores y usando frenado automático o semiautomático. El proyecto PROTECTOR y su sucesor SAVE-U se centraron particularmente en la reducción de accidentes involucrando VRUs (5). El proyecto europeo PUVAME propone una solución para prevenir colisiones entre VRUs y autobuses. Todos estos sistemas utilizan cámaras “off-board” que observan las intersecciones y las paradas de autobús para rastrear el movimiento de los buses así como de los VRUs.

Los sensores de imagen pueden capturar una perspectiva de alta definición de la escena, pero la extracción de la información implica una gran cantidad de procesado. Además, la visibilidad desde el vehículo es muchas veces limitada. Es muy frecuente el caso en el que es difícil o imposible observar el objeto peligroso desde el propio vehículo.

En el Anexo A se detalla más información acerca de las características de los sistemas basados en videocámaras.

2.2. Sistemas radar

Los radares automotor están empezando a formar parte del equipamiento estándar de los coches de gama alta, y se espera que pronto lo sean en los de gama media-baja. Aunque estos radares a menudo no son capaces de distinguir con facilidad a los VRUs, tales como peatones, ciclistas y motociclistas de otros usuarios u obstáculos de la vía, como coches o señales de tráfico.

Uno de los sistemas que se estudia consiste en un radar armónico, y un pequeño transpondedor que lleva el VRU (6). El radar armónico transmite una señal que consiste en dos componentes frecuenciales distintas (f_1 y f_2). En el transpondedor, las dos frecuencias son mezcladas y un producto armónico es transmitido de vuelta al radar. Para cumplir con la banda designada para radares automotor, f_2 es bastante cercana a f_1 , así el producto armónico de tercer orden, $2f_1 - f_2$, generado en el transpondedor, se sitúa cerca de f_1 y f_2 . El sistema radar procesa las reflexiones de objetivos convencionales usando una de las componentes (f_1 o f_2) directamente. La respuesta del

transpondedor es procesada sintetizando la frecuencia de dicho transpondedor basándose en las componentes f_1 y f_2 . De esta forma, el radar puede tener dos principales salidas, donde una de ellas está relacionada con las reflexiones convencionales y la otra está relacionada con un VRU llevando un transpondedor.

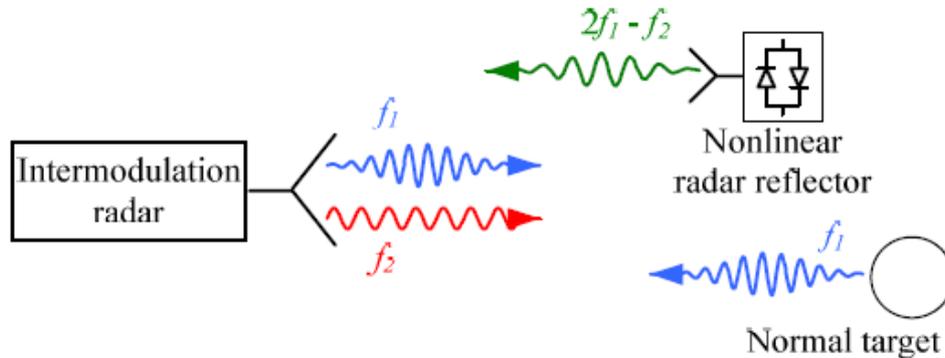


Fig. 3: Visión esquemática del concepto de radar armónico y transpondedores

El radar transmite una señal consistente en dos frecuencias distintas en la banda de 76-81 GHz.

Aunque es posible desarrollar un transpondedor de bajo coste y suficientemente pequeño como para poder ser integrado en la ropa, para poder ser utilizado en este tipo de aplicaciones, puesto que no necesita batería, es básicamente un elemento mezclador unido a una antena, y puede ser mejorado para obtener buenos resultados, se decide abandonar esta idea porque se considera que proveer de un transpondedor a todos los VRUs no es una idea viable.

2.2.1. Radares micro-Doppler

Este tipo de radares trabajan constantemente, enviando una señal de frecuencia estable y procesando el eco reflejado. Así maximizan la potencia total en el objetivo. Usan Doppler para medir la velocidad angular del objetivo, lo que les hace inmunes a la interferencia de grandes objetos estáticos, y les permite incrementar el número de muestras para aumentar el rango de detección sin incrementar la potencia transmitida.

Dentro de las posibles aplicaciones de estos radares, se encuentra el ámbito militar. Por ejemplo, son usados para guiar misiles aire-aire buscadores de blancos con radares semiactivos (SARH, Semi-Active Radar Homing) como el U.S. AIM-7 Sparrow, y otros misiles estándar. También están incluidos en sistemas de vigilancia, usados en diversas tácticas de vigilancia en el campo de batalla, reconocimiento de infantería o vigilancia de fronteras (7).

Publicaciones recientes muestran que un simple sensor radar es capaz de grabar perfiles micro-Doppler de humanos caminantes (8).

El movimiento de los peatones se caracteriza por su no-uniformidad. Incluso cuando el centro de gravedad se mueve con una velocidad constante, todas las partes del cuerpo realizan un proceso de aceleración y deceleración técnicamente complicado,

pero periódico. La periodicidad de este movimiento viene dada por el llamado “walk cycle”, que se define con dos pasos. En este ciclo, un pie está siempre en el suelo para un movimiento de “no avance”. La Fig. 4 muestra un dibujo de este ciclo para visualizar el proceso periódico del tronco, así como de brazos y piernas.

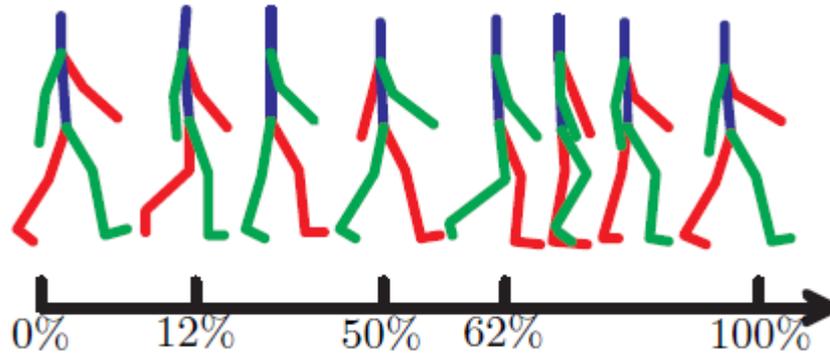


Fig. 4: Dibujo del típico “walk cycle” humano

Esta descripción implica una aproximación de las piernas a simples péndulos, lo cual es suficiente para esta aplicación. Las piernas izquierda y derecha tienen una diferencia de fase de 180° , lo cual se representa con un “menos” a la hora de calcular la velocidad de cada parte, como comprobaremos más adelante en el apartado de resultados.

Después de las consideraciones teóricas, la misma disposición, esta vez para una persona con movimiento radial fue grabada por un sensor radar a 24 GHz. Para esto, dicho sensor se modifica para medidas de alta resolución. Para detectar las señales esperadas, se estableció una resolución espacial de $r = 0.08$ m y una resolución de velocidad de $v = 0.09$ m/s.

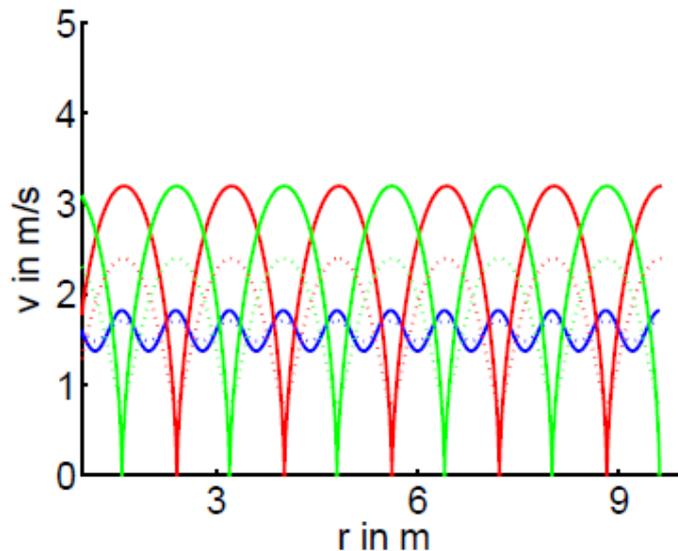


Fig. 5: Diagrama espacio-velocidad simulado para el cuerpo (azul) y par a las piernas (verde y rojo) de un peatón andando

Esto ayuda a comprender que algunos de los resultados que se presentan más adelante, relacionados con algunos de estos parámetros, y aclarar la diferencia de fase de las piernas.

Pero, trabajando con este tipo de radares, también se pueden tomar muestras de vehículos. Tal y como se explica en (9), uno de los principales objetivos es diferenciar entre vehículos en movimiento y peatones, en términos de extracción de características y clasificación. El espectro Doppler es la base del reconocimiento de objetivos integrado en un sensor radar automotor. También la velocidad lateral es usada para clasificar peatones y vehículos en cualquier escenario urbano de forma correcta. Los vehículos moviéndose de forma longitudinal tienen la misma velocidad en todos los puntos de reflexión. Sin embargo, en caso de movimientos laterales, se pueden medir diferentes velocidades provenientes de diferentes puntos de reflexión, debido a la dirección del movimiento del vehículo. Esto hace peligrosos y difíciles de clasificar a los vehículos, porque esta característica es bastante similar a la observada en el caso de detección de peatones.

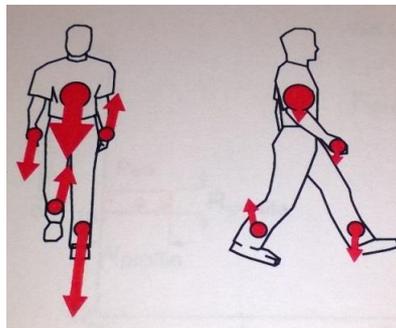


Fig. 6: Modelo de puntos de reflexión para un peatón en movimiento

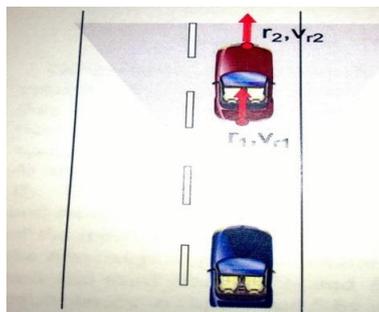


Fig. 7: Modelo de puntos de reflexión para un vehículo con movimiento longitudinal

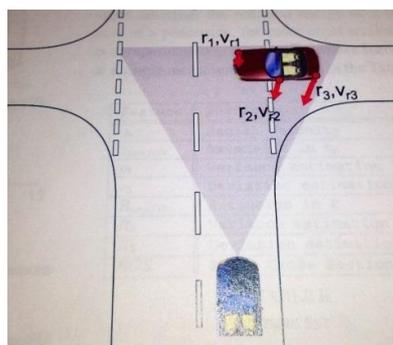


Fig. 8: Modelo de puntos de reflexión para un vehículo con movimiento lateral

La asignación de un objeto medido a cualquiera de las clases es desarrollada por un algoritmo de decisión subjetivo basado en las características extraídas. Una vez más, los parámetros del clasificador son generados en una fase de entrenamiento usando una serie de muestras de entrenamiento. Estas muestras consisten en vectores de características seleccionados aleatoriamente dentro de las medidas realizadas por el radar y correspondientes a las diferentes clases. Los clasificadores más típicos son los basados en máquinas de vectores de soporte (SVMs).

En este estudio, para la detección de peatones, si se comparan solo con vehículos con movimiento longitudinal, se llega a obtener hasta un 99% de precisión. Sin embargo, comparándolos con vehículos con movimiento lateral, solo el 79,9% de los peatones son clasificados correctamente, lo cual no es un buen resultado.

Y, además de todo esto, no hay muchos estudios documentados acerca de la utilización de estos sistemas para la detección clasificación de renos.

2.3. Conclusiones

Aunque, tal y como se explicaba anteriormente, se pueden obtener buenos resultados utilizando sistemas de imagen, su complejidad, y el hecho de ser altamente influenciados por la oscuridad y las condiciones meteorológicas, nos hacen tomar la decisión de decantarnos por el uso de un sistema radar micro-Doppler para la realización de este proyecto.

Además, ha habido un gran progreso en el desarrollo de radares de 24 GHz, y son bastante menos caros en comparación con la tecnología que opera a 77 GHz. Estos radares son simples de fabricar y operar con ellos. Y, por esto, son baratos de fabricar y mantener.

Es por eso que algunas de las compañías (por ejemplo Nissan, Mercedes –ver sistema PRE-SAFE- etc.) están utilizando este tipo de sistemas radar en lugar de los ópticos. Aplicaciones como el aviso de invasión de carril contrario, ayuda de aparcamiento, control de velocidad o Stop&Go ya han sido introducidas en el mercado.

3. Métodos y Materiales

Como se explica anteriormente, en este proyecto se decide trabajar con un sistema basado en un sensor radar, en lugar de uno basado en videocámaras.

Un sistema de reconocimiento de peatones basado en radar tiene dos componentes principales, un sensor radar y una unidad de procesamiento de señal, que consiste en un procesador de los datos recibidos por el radar (análisis y filtrado) combinado con un algoritmo de detección y clasificación (10), tal y como vemos en la Fig. 9.

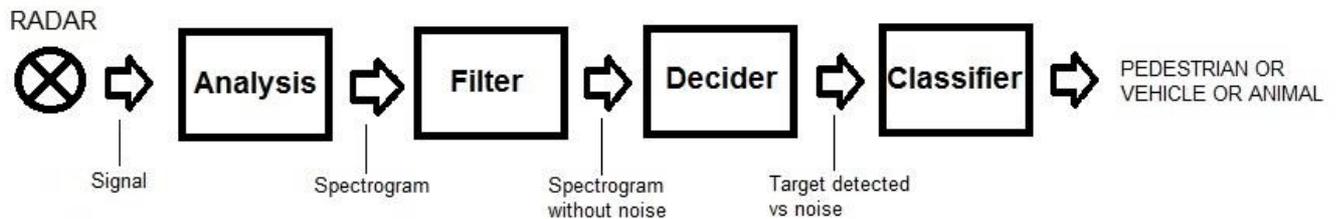


Fig. 9: Representación esquemática del sistema

3.1. Sensor radar

Las medidas se realizan con un radar de onda continua (CW, Continuous-Wave radar). Este tipo de radares están continuamente trabajando, enviando una señal de frecuencia estable y procesando el eco reflejado. De esta manera maximizan la potencia total en el objetivo. Usan Doppler para medir la velocidad del objetivo, lo que les hace inmunes frente a las interferencias provenientes de grandes objetos estáticos, y les permite incrementar el número de muestras para ampliar el rango de detección sin incrementar la potencia transmitida.

Para este tipo de aplicaciones, las dos bandas frecuenciales principales son 24 y 77 GHz. Se decide trabajar en la banda de 24 GHz porque este tipo de radares son mucho más atractivos, principalmente debido a su menor coste.

En nuestro caso, se utiliza el modelo IPS-146 de InnoSenT. Un radar de onda continua, a 24 GHz, con un preamplificador de radiofrecuencia para operar en el modo más bajo de ruido, y una antena separada para transmisión y recepción.

3.1.1. Especificaciones del radar

Parameter	Symbol	Min.	Typ.	Max.	Units	Comment
Oscillator						
transmit frequencies	$f_{IPS-146}$	24.000	24.125	24.250	GHz	delivery condition
	$f_{IPS-146_UK}$	24.150	24.200	24.250	GHz	
	$f_{IPS-146_F}$	24.075	24.125	24.175	GHz	
temperature drift (frequency)	Δf		-1		MHz/°C	
output power (EIRP)	P_{out}		18		dBm	@ 25°C
Receiver						
I/Q balance	amplitude			6	dB	
	phase	60	90	120	°	
IF-output	voltage offset		$V_{cc}/2$		V	
IF-amplifier	bandwidth		50 - 10		kHz	
	gain		50		dB	
Antenna pattern (compare with antenna plot on page 3)						
full beam width @ -3dB	horizontal		30		°	azimuth
	vertical		32		°	elevation
side-lobe suppression	horizontal		20		dB	azimuth
	vertical		20		dB	elevation
Power supply						
supply voltage	V_{cc}	4.75	5.0	5.25	V	
supply current	I_{cc}		60	80	mA	IF-amp included
Environment						
operating temperature	T_{OP}	-20		+60	°C	
storage temperature	T_{STG}	-40		+85	°C	
Mechanical Outlines						
outline dimensions	height		9.0		mm	
	length		60.0			
	width		37.0			

Tabla 2: Especificaciones del radar

3.1.2. Recopilación de muestras

La primera tarea consiste en tomar las muestras de los diferentes objetivos. El equipo necesario para tomarlas incluye el radar antes mencionado, un grabador de audio que graba la señal recibida por el radar y la convierte en audio, una videocámara para capturar vídeo, y un programa para mezclar ambos, de manera que el resultado final sea una grabación simultánea de audio y video, lo cual facilitará el posterior análisis de cada muestra por separado.

Para los peatones y los vehículos nos acercamos al aparcamiento de la universidad. Se tomaron tantas muestras distintas como fue posible, para estudiar diferentes situaciones (coches moviéndose hacia nosotros, alejándose de nosotros, con movimiento lateral, peatones hacia nosotros, alejándose de nosotros, etc.). El mismo método fue empleado para tomar las muestras de los renos, pero, en este caso, tuvimos que ir a un zoo cercano a la ciudad, para poder encontrarlos. Agradecimiento especial al personal del Ahtari Zoo por su ayuda en la toma de muestras.

Al final de la sesión de toma de muestras, se dispone de varias grabaciones de cada uno de los objetivos, que serán editadas y analizadas para generar una base de datos de muestras de corta duración, listas para ser utilizadas en las siguientes etapas.

3.2. Unidad de procesamiento de señal

3.2.1. Edición de las muestras

La primera fase consiste en la edición de los videos capturados. Primero se realizan cortes correspondientes a capturas de los diferentes objetivos utilizando un programa de edición de vídeo (Virtualdub[®]), y posteriormente se extrae de cada uno de ellos la señal de audio utilizando el programa VLC Media Player[®]. Al terminar esta fase, se dispone de las siguientes muestras de audio.

Objetivo	Nº de muestras
Vehículo	13
Peatón	8
Reno	9
Sin obstáculo	10

Tabla 3: Base de datos de muestras

Se descartan varias muestras en las que aparecen varios objetivos, para evitar aumentar la complejidad de las sucesivas etapas.

3.2.2. Análisis

Una vez extraída la señal de audio de cada una de las muestras obtenidas, comienza la fase de análisis.

La mejor forma de obtener una buena descripción de esta señal es representando su espectrograma. El espectrograma es una representación del espectro de una señal frente a un periodo de tiempo. El formato más común es una gráfica con dos dimensiones: el eje horizontal representa el tiempo, el vertical la frecuencia. Una tercera dimensión indica la amplitud de una frecuencia en concreto para cada instante de tiempo, en función de la intensidad del color de cada punto de la imagen. Los espectrogramas son calculados habitualmente utilizando la transformada de Fourier de tiempo reducido (STFT). Tras varias pruebas, se observa que la mejor representación se obtiene utilizando la transformada de Fourier de 1024 puntos, con un solapamiento de 10 muestras.

Por otro lado, podemos asociar fácilmente la frecuencia a la velocidad en este caso, utilizando la relación frecuencia-velocidad asociada al fenómeno Doppler:

$$v = f \cdot \frac{\lambda}{2}$$

donde λ es la longitud de onda del radar.

Por lo tanto, a partir de este punto, hablaremos de los espectrogramas como representaciones de velocidad respecto al tiempo.

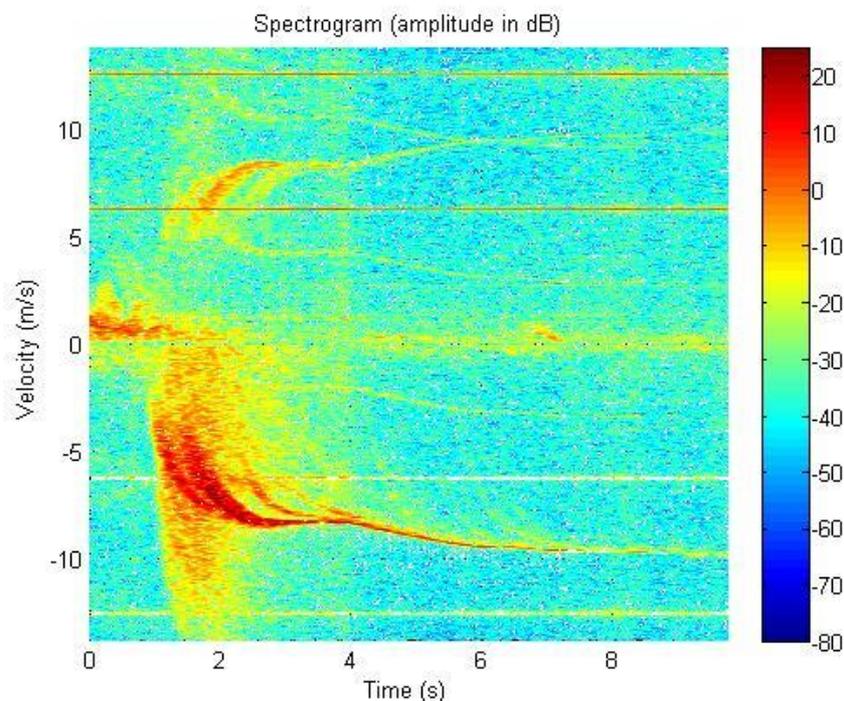


Fig. 10: Ejemplo de espectrograma

Con estas representaciones podemos empezar a apreciar las diferencias entre los distintos objetivos. Como explicaremos más adelante, hay grandes variaciones en las características de los objetivos estudiados, y algunas de ellas son fácilmente identificables simplemente ojeando el espectrograma.

3.2.3. Filtrado

Observando los espectrogramas, rápidamente nos damos cuenta de que están altamente afectados por el ruido. Además del ruido típico que puede ser producido por cualquier objeto a nuestro alrededor, o por el medio ambiente, se aprecia otro tipo de ruido, introducido por el propio radar. En la gráfica se aprecia al ver cuatro líneas, correspondientes a las frecuencias 1 KHz y 2 KHz (y sus réplicas a -1 KHz y -2 KHz).

Puesto que se trataba de una tarea complicada eliminar esas líneas sin perder información útil de los objetivos, se opta por utilizar el método de sustracción espectral.

Sustracción espectral: Es un método simple y efectivo de reducción de ruido (11). En este método, la idea es realizar una estimación del espectro medio de la señal de ruido, y extraerlo de la señal original, de modo que la relación señal a ruido (SNR) es claramente mejorada.

La señal ruidosa $y(m)$ es la suma de la señal deseada $x(m)$ y el ruido $n(m)$:

$$y(m) = x(m) + n(m)$$

En el dominio frecuencial, se denota:

$$Y(j\omega) = X(j\omega) + N(j\omega) \Rightarrow X(j\omega) = Y(j\omega) - N(j\omega)$$

Así pues, la idea es analizar algunas de las muestras tomadas en las que no aparece ningún objetivo de la misma forma que se ha hecho antes, y usarlas para estimar el espectro de la señal de ruido.

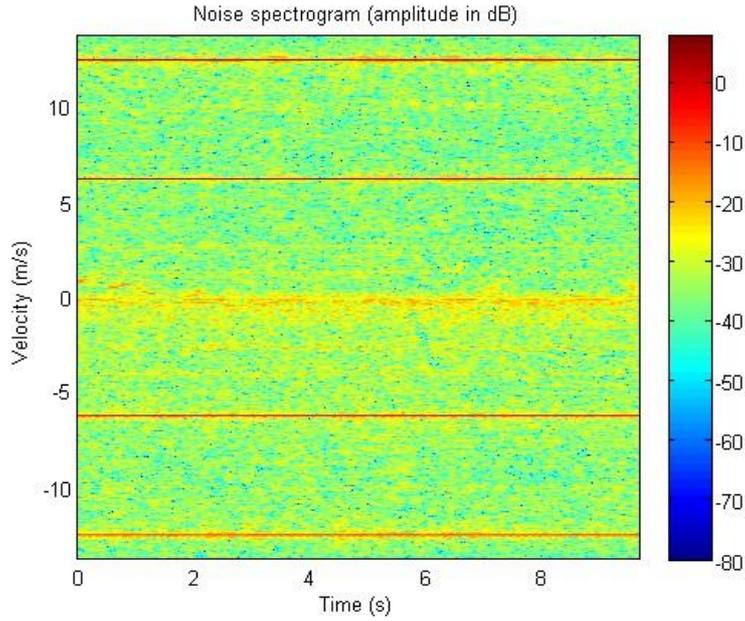


Fig. 11: Ejemplo de espectrograma de ruido

Después de esto, el espectro de la señal de ruido es utilizado para extraerlo de todas las muestras que queremos analizar. Pero este proceso no es tan simple, puesto que muchos de los radares tienen el llamado Control Automático de Ganancia (AGC). Debido a esto, la potencia del ruido cambia a lo largo del tiempo (si el objetivo está lejos, entonces el radar incrementa la potencia, y por lo tanto, el ruido).

La ganancia del ruido es estimada como la longitud normalizada de la proyección de la señal de ruido en el espectro medio de ruido:

$$G(t) = \sum_{f=1}^M \frac{\gamma(f)S(f, t)}{\|\gamma\|^2}$$

donde $\|\gamma\| = \sqrt{\gamma(1)^2 + \gamma(2)^2 + \dots + \gamma(M)^2}$ es la norma euclídea del patrón de ruido. Ahora el ruido puede ser eliminado así:

$$X(f) = Y(f) - G(f)N(f)$$

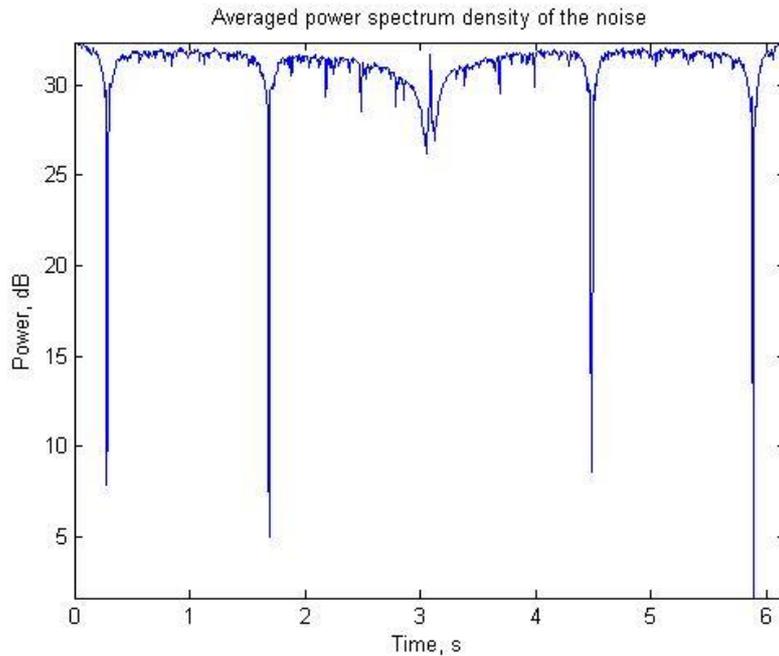


Fig. 12: Densidad espectral de potencia media de ruido

3.2.4. Decisión

Una vez que las señales de audio extraídas de las muestras tomadas con el sensor radar han sido analizadas y filtradas, están listas para los pasos finales. Primero, hay que decidir si una señal corresponde a un objetivo detectado o es simplemente ruido. Este paso se realiza como test de hipótesis (12), intentando obtener una alta probabilidad de detección, y una baja probabilidad de falsa alarma.

Puesto que se trata de un ejercicio basado en teoría de decisión estadística, necesitamos trabajar con una descripción estadística de la función de densidad de probabilidad (pdf) de las señales con las que queremos trabajar.

Con este objetivo, se desarrolla el decisor. Se necesita la función de densidad de probabilidad para los dos casos: objetivo más ruido, o solo ruido. Por lo tanto, se construyen dos matrices diferentes, una de ellas con espectrogramas de varias muestras de los diferentes objetivos, y la otra con espectrogramas de varias muestras donde no se ha capturado ninguno. Se calculan los valores de potencia para cada instante en ambas matrices, y se representa la distribución de potencia de ambas para observar el rango de potencia en el que se encuentran. Aparte de esto, se representan sus funciones de distribución de probabilidad, o histograma. Observando la diferencia entre ambas distribuciones, el paso final consiste en establecer un umbral. Un valor que sirva como frontera entre ambas hipótesis. Más adelante se muestran ejemplos donde se observa que muestras analizadas cuyos valores son mayores que dicho umbral serán consideradas objetivos, mientras que las muestras que tengan valores por debajo del umbral, se considerarán ausencia de objetivos.

Antes de decidir el valor de dicho umbral, se deben considerar los siguientes conceptos relacionados con el método de hipótesis:

- Probabilidad de detección, P_D : La probabilidad de que un objetivo sea declarado, cuando existe tal objetivo.
- Probabilidad de falsa alarma, P_{FA} : La probabilidad de que un objetivo sea declarado, cuando no hay un objetivo presente.
- Probabilidad de pérdida, P_M : La probabilidad de que un objetivo no sea declarado, cuando hay un objetivo presente.

Basándonos en estas probabilidades, hay diferentes opciones para establecer el umbral. Para determinadas aplicaciones, es interesante tener una alta probabilidad de detección, lo que implica que cada vez que se declara un objetivo, es porque existe la certeza casi absoluta de que dicho objetivo está presente. Pero en nuestro caso la decisión no es tan fácil. Dado que dicha decisión puede implicar una reacción mecánica posterior para evitar la posible colisión, no podemos descuidar la probabilidad de falsa alarma, ni la de pérdida. Un sistema que no avisa de la presencia de un objetivo cuando está, simplemente por su baja probabilidad, se convierte en inútil para su propósito. Pero también lo es un sistema que avisa de la presencia de un objetivo cuando no lo hay, puesto que también puede resultar peligroso. Por lo tanto, es lógico pensar que la decisión final razonable será un punto intermedio, el cual será analizado en la siguiente sección.

Una vez se decide el valor umbral, cada nueva muestra será analizada, se calculará la potencia media asociada su espectrograma y, en función del valor obtenido, se decidirá si corresponde a un objetivo o se trata simplemente de ruido.

3.2.5. Clasificación

Después de filtrar y analizar la señal medida por el sensor radar, y decidir si corresponde a un objetivo o no, solo queda decidir de qué objetivo se trata.

Hay muchos tipos de clasificadores, basados en diferentes métodos. Algunos de los más utilizados son el clasificador Naïve Bayes (13) (14), basado en el teorema de Bayes (15) y el clasificador basado en máquinas de vectores de soporte (SVM) (16), que es el que se usa en este proyecto. Se elige éste porque se trata de un clasificador con el que se pueden llegar a obtener buenos resultados y, como se explica en apartados anteriores y en el Anexo A, es el más comúnmente utilizado en aplicaciones de este tipo.

Como muchos sistemas de clasificación, las SVM trabajan sobre lo que se denomina vector de características, y proporcionan a su salida la clase asociada a dicho vector (en nuestro caso peatón, vehículo o reno). El vector de características ha de ser descriptivo de la muestra introducida, permitiendo clasificar la señal. En nuestro caso, como todas las secciones anteriores han ilustrado el interés de las representaciones tiempo-frecuencia para esta tarea, usaremos como vector de características el espectrograma y coeficientes Cepstrum (17) (18). Los coeficientes Cepstrum se calculan

como la antitransformada del logaritmo del módulo de la transformada de Fourier de la señal. Es típica su utilización en aquellos problemas de clasificación en los que la variación temporal de las características espectrales es relevante, como en reconocimiento de voz, clasificación sobre señales sísmicas, etc.

Las SVM requieren una etapa de entrenamiento, para la que se utilizan vectores de características correctamente clasificados. En esta etapa se determinan una serie de parámetros que permitirán clasificar nuevos vectores.

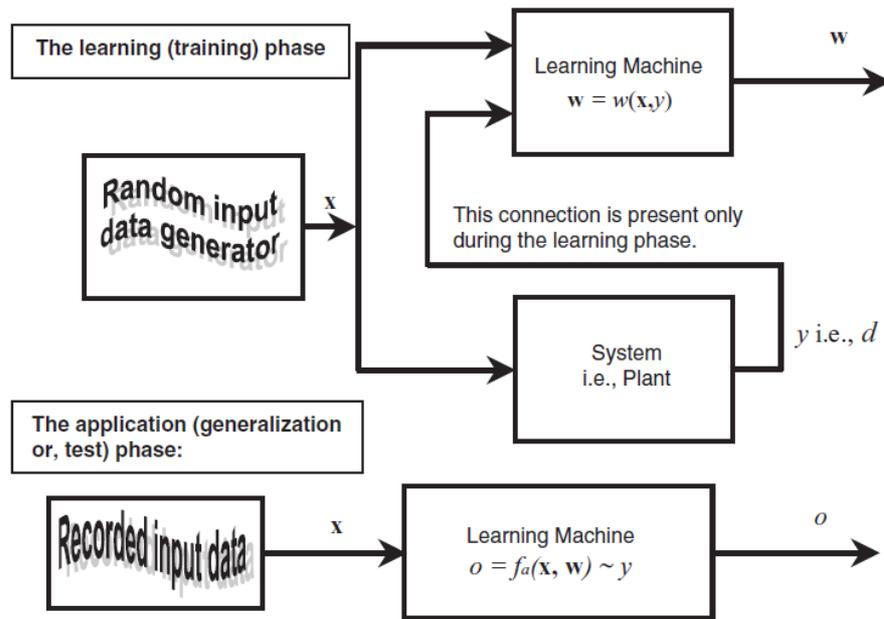


Fig. 13: Representación esquemática del método SVM

De nuevo, y puesto que estamos simulando un sistema de seguridad y el objetivo principal es evitar colisiones, buscamos un resultado lo más preciso posible.

Algunos detalles adicionales, tanto sobre el Cepstrum como sobre las SVM, se describen en el Anexo B.

4. Resultados

En este apartado se presentan los resultados obtenidos en las diferentes etapas del proyecto, tras haber seguido los pasos indicados en los capítulos anteriores para su realización.

4.1. Análisis y Filtrado

En esta primera parte se observan los espectrogramas correspondientes a algunas de las muestras utilizadas a lo largo del proyecto, correspondientes a las tres clases: vehículos, peatones y renos. Al lado de cada una de ellas aparece la imagen real del objetivo.

4.1.1. Vehículos

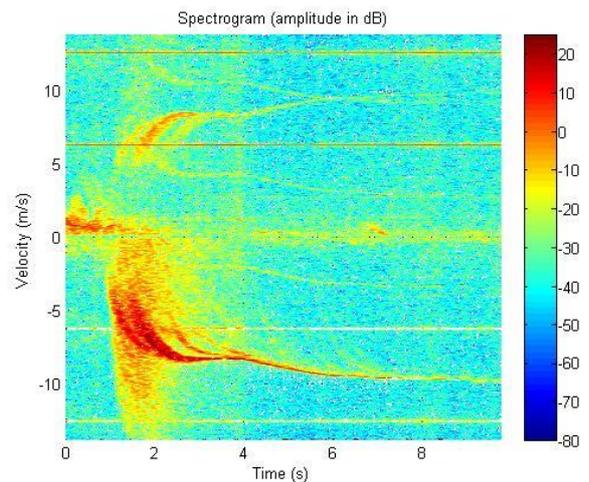


Fig. 14: Espectrograma de un coche en movimiento alejándose de nosotros (antes de filtrar)

En este primer espectrograma se puede apreciar la forma de un coche que aparece detrás de nosotros, y empieza a acelerar mientras se aleja, hasta que desaparece. La velocidad negativa representa el hecho de que el coche se aleja. La velocidad negativa en esta situación está relacionada con las frecuencias negativas medidas por el sensor radar, debidas al movimiento del coche.

Se observa también las líneas de ruido introducidas por el propio radar. Calculando la relación entre velocidad y frecuencia, se deduce que las frecuencias son 1 y 2 KHz.

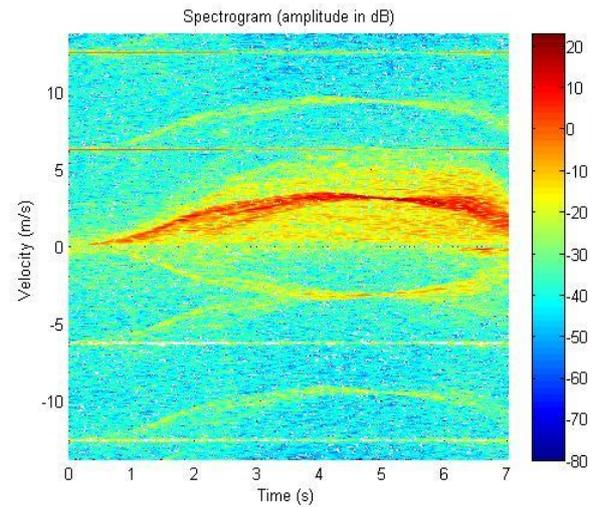


Fig. 15: Espectrograma de un coche moviéndose linealmente (antes de filtrar)

Aquí tenemos un caso diferente. Aunque se trata también de un coche, podemos observar que la forma del espectrograma es diferente. Esto se debe a que el movimiento del coche también lo es. Estaba parado a nuestra izquierda, y empezó a moverse hacia la derecha, mientras aceleraba. La velocidad aumenta porque al principio el coche está acercándose al radar (desde la izquierda hasta el centro), después de unos segundos, empieza a alejarse (desde el centro hasta la derecha), así que la velocidad empieza a decrecer hasta que desaparece de nuestro alcance.

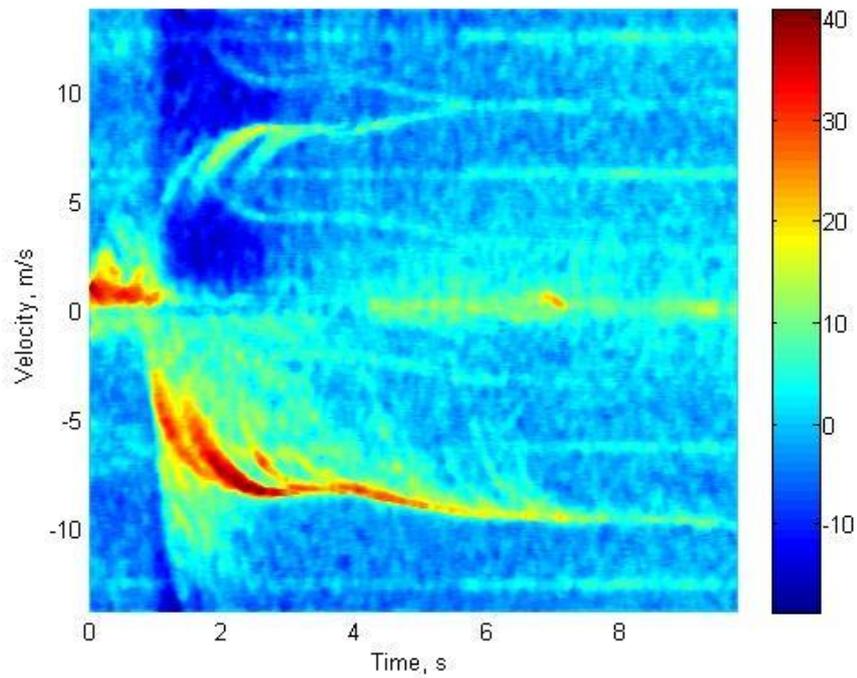


Fig. 16: Espectrograma de un coche en movimiento alejándose de nosotros (tras sustracción espectral con Matlab®)

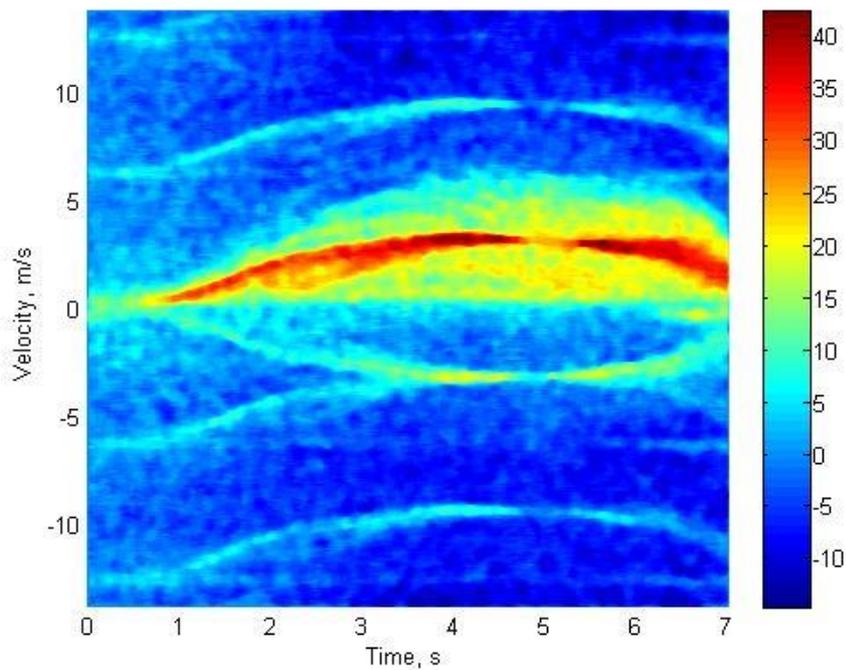


Fig. 17: Espectrograma de un coche moviéndose linealmente (tras sustracción espectral con Matlab®)

Como el espectro de ruido usado para la sustracción espectral se construyó con diferentes muestras, tomadas de coches, peatones y renos, no se obtiene un filtro perfecto para cada muestra. Además, se le añade un filtro de mediana como último paso, para que el resultado sea más limpio.

4.1.2. Peatones

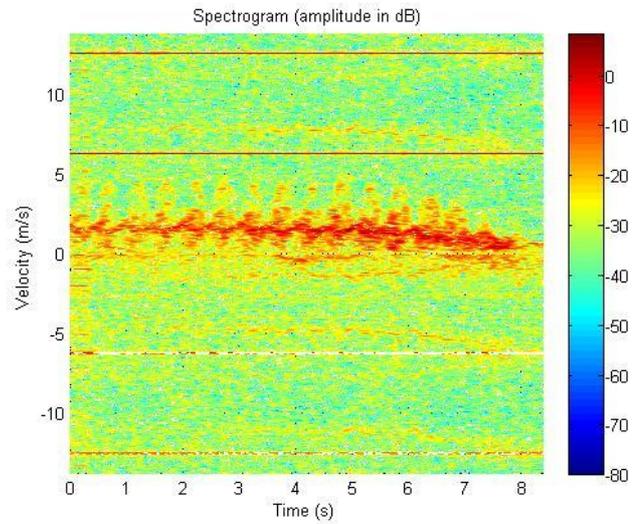


Fig. 18: Espectrograma de una persona moviéndose hacia nosotros (antes de filtrar)

Ahora tenemos un caso distinto. La velocidad cambia mucho más rápidamente que en los casos anteriores, y está relacionado con la explicación de la Fig. 9. Puesto que cada paso tenemos tres partes principales, todas ellas tienen su velocidad. Si prestamos atención a la gráfica, podemos observar una línea gruesa roja, que representa el tronco del peatón y su velocidad media. Aparte de esa línea, y no tan fáciles de apreciar, tenemos otros dos movimientos diferentes que representan las piernas. Mientras uno de ellos incrementa su velocidad, el otro la reduce.

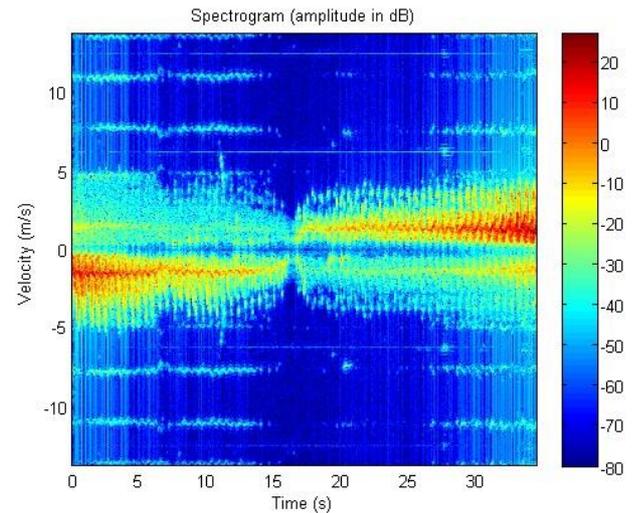


Fig. 19: Espectrograma de una persona alejándose y luego acercándose (antes de filtrar)

Aunque, como se puede observar, el espectrograma original en este caso es mucho más limpio que en los anteriores (esta muestra fue tomada en condiciones diferentes, un día diferente, con el objetivo de compararlas), podemos observar similitudes. Primero, la forma del espectrograma es casi la misma, puesto que se aprecian las diferentes velocidades asociadas al tronco y las piernas. En este caso además, podemos observar la diferencia entre detectar una persona que se aleja o que se acerca. Al principio, el peatón se está alejando, por lo que la velocidad se considera negativa (como se observaba para el primer coche). Unos segundos después, el objetivo vuelve hacia nosotros, y la velocidad pasa a ser positiva.

En este caso también hay que eliminar las líneas de ruido.

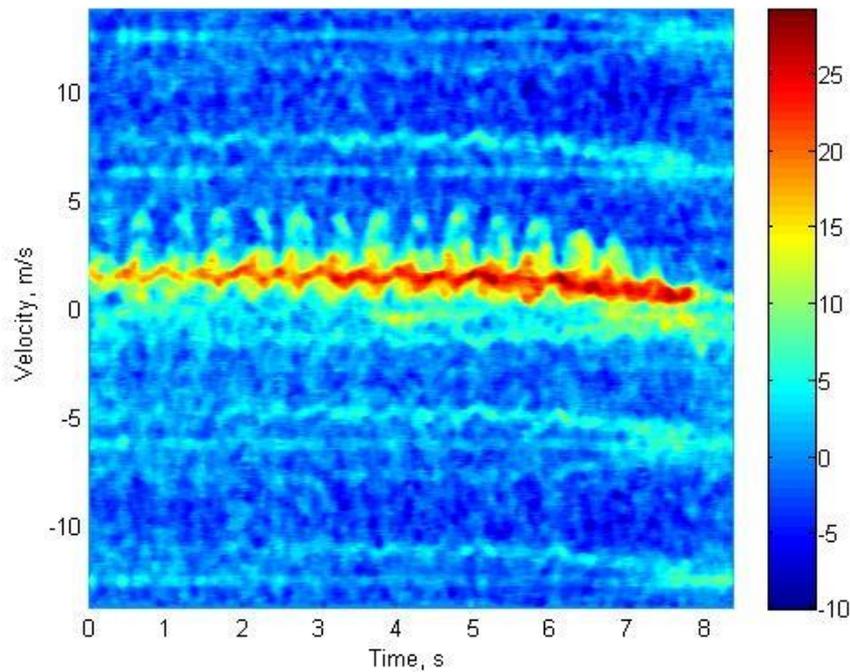


Fig. 20: Espectrograma de una persona moviéndose hacia nosotros (tras sustracción espectral con Matlab®)

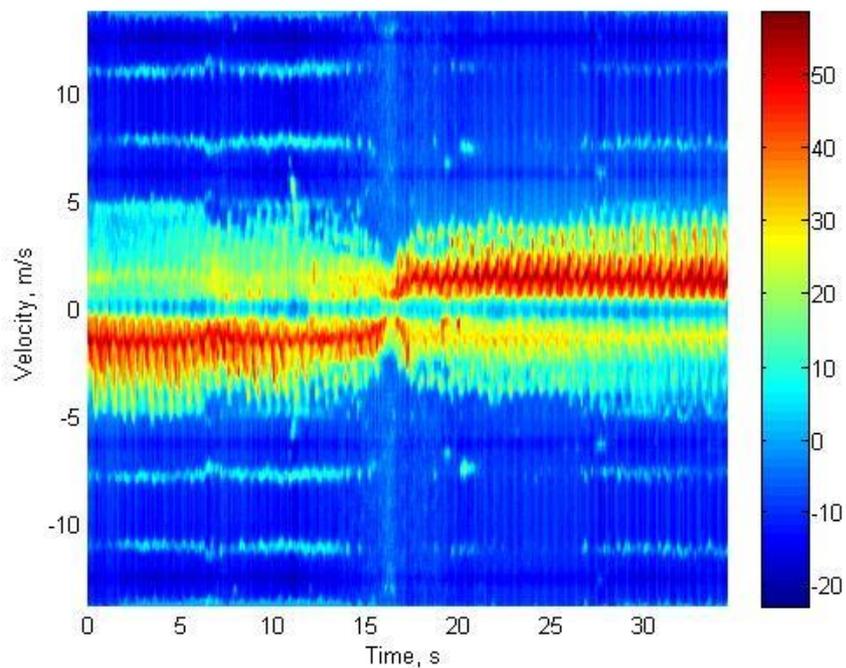


Fig. 21: Espectrograma de una persona alejándose y luego acercándose (tras sustracción espectral con Matlab®)

En esta situación tenemos una gran diferencia entre ambos casos. Esto se debe al hecho de que la segunda muestra fue tomada un día diferente, por lo que el ruido es distinto y su espectrograma ya era bastante limpio antes de filtrar. Aun así, podemos observar que las líneas de ruido han sido eliminadas.

4.1.3. Renos

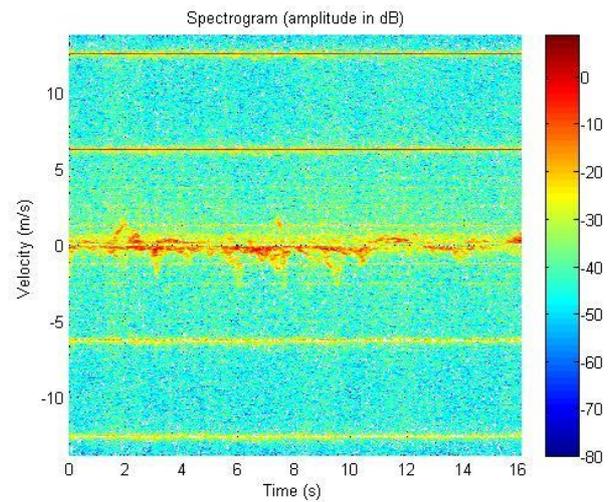


Fig. 22: Espectrograma de un reno (antes de filtrar)

Las muestras de renos fueron tomadas otro día, como se ha explicado al detallar el proceso. Es por eso que la potencia de las señales y la del ruido que las afecta son distintas.

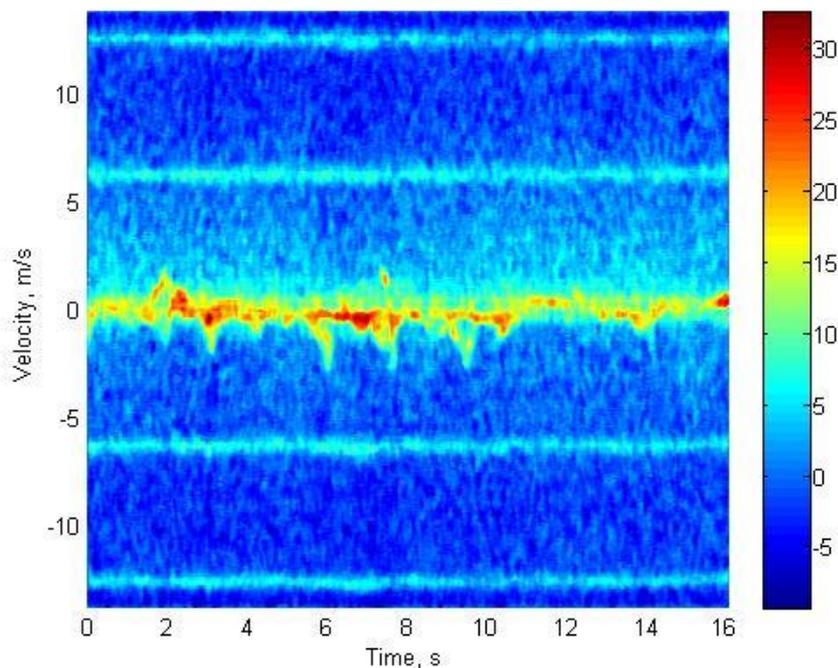


Fig. 23: Espectrograma de un reno (tras sustracción espectral con Matlab®)

Aquí se observa que las líneas de ruido afectan mucho más a estas muestras. Esto se debe a la diferencia de potencia, y al hecho de que el espectro medio de ruido usado para la sustracción espectral está construido con muestras correspondientes a todos los objetivos. Esto explica por qué las líneas parecen no ser eliminadas usando este método.

4.2. Decisión

En este apartado se muestran las dos posibles opciones a considerar en esta etapa: presencia o ausencia de obstáculo. Se comparan y se visualiza de forma gráfica la elección del umbral de decisión y su repercusión.

4.2.1. No hay objetivo detectado

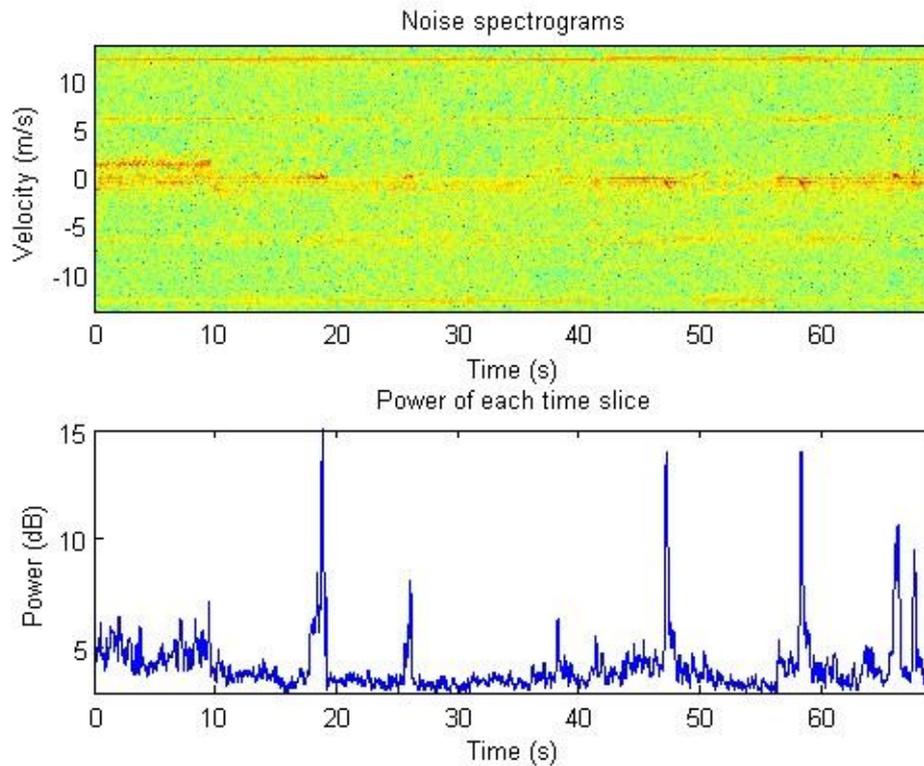


Fig. 24: Espectrograma de varias muestras sin objetivo y sus potencias

En este caso se muestra el espectrograma filtrado de diferentes muestras seguidas en las que no hay ningún objetivo.

En la segunda gráfica se observa la potencia de cada punto del espectrograma mostrado en la primera gráfica. Se puede apreciar que, salvo en algunos puntos (muy cortos en tiempo), la amplitud es realmente baja, como se esperaba.

4.2.2. Objetivo detectado

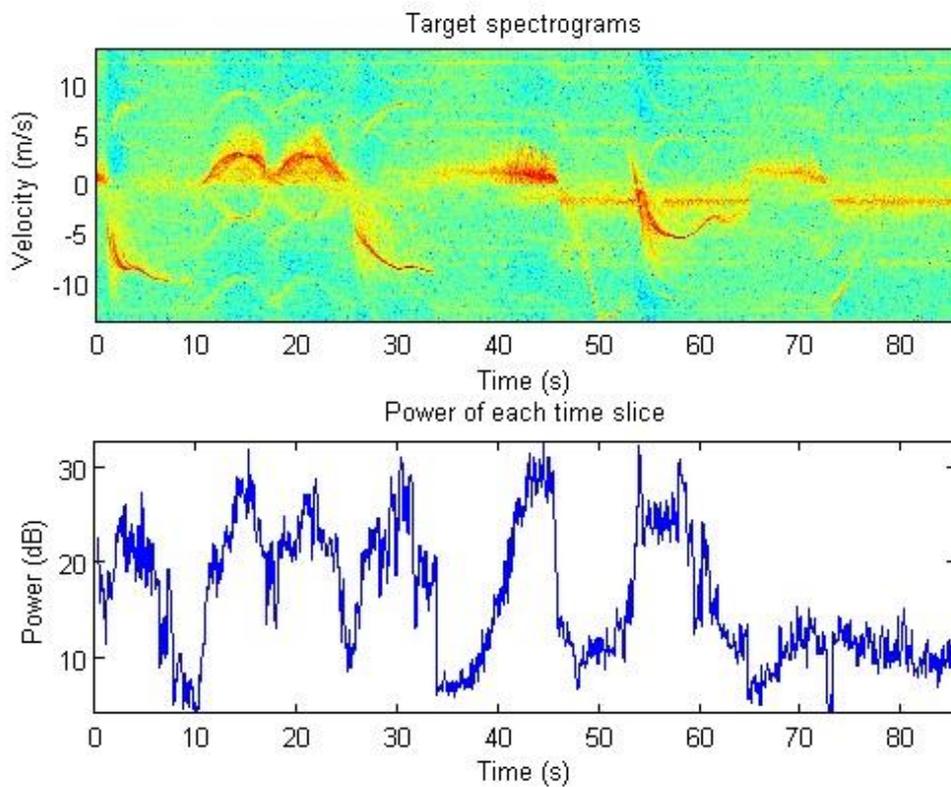


Fig. 25: Espectrograma de varias muestras con diferentes objetivos y sus potencias

Para este segundo caso mostramos el espectrograma filtrado de diferentes muestras seguidas de objetivos detectados.

Ahora se observa en la segunda gráfica que la amplitud es mucho mayor que en el caso anterior. Hay también algunos puntos de potencia baja, correspondientes todos ellos a los comienzos o finales de las muestras, donde el objetivo aún no ha aparecido o ya se ha ido. Se debe tener especial cuidado con esos puntos, porque pueden producir fallos en la clasificación, tal y como se explica más adelante.

4.2.3. Distribución de potencia

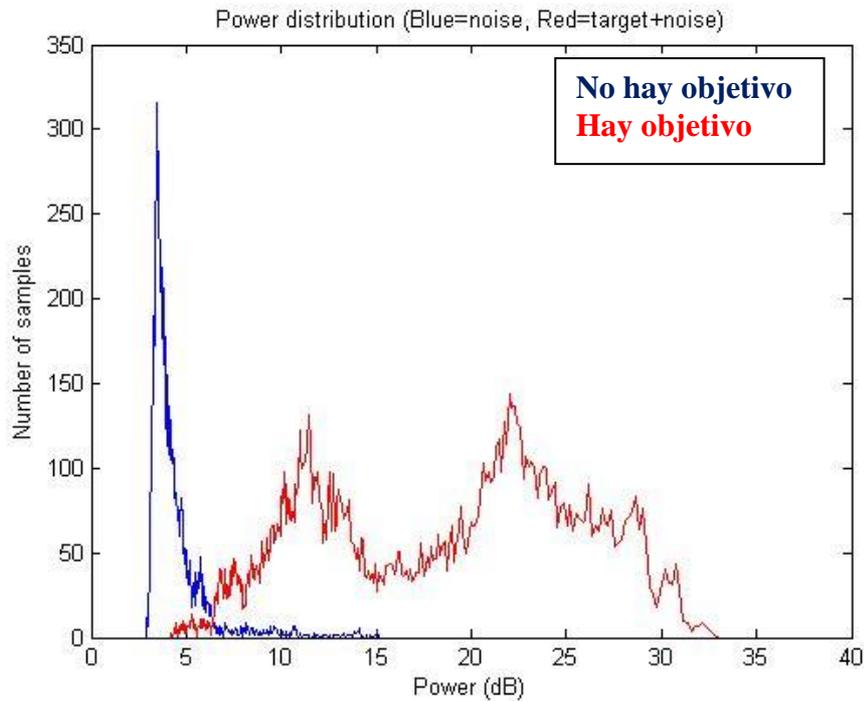


Fig. 26: Histograma de la distribución de potencia para ambos casos.

La representación final en este paso es la distribución de potencia de ambos casos, sin y con objetivo detectado. Para el primer caso, sin objetivo, las muestras están distribuidas entre 3 y 15 dB, pero principalmente entre 3 y 5 dB, como era de esperar. Esto es debido a los puntos que tienen mayor amplitud en breves periodos de tiempo. Para el segundo caso, se observa que las muestras están distribuidas principalmente entre 7 y 33 dB, pero también aparece alguna entre 4 y 7 dB, debido a los segmentos de los espectrogramas que corresponden a los principios y finales de las muestras medidas, donde la amplitud es más baja, como ya se ha explicado antes.

Para decidir el mejor umbral, hay que valorar los resultados obtenidos en cuanto a probabilidad de detección, probabilidad de falsa alarma y probabilidad de pérdida. Si se desea una probabilidad de detección cercana al 100%, se debería establecer un umbral ligeramente inferior a 5 dB. Entonces, cada posible objetivo sería declarado, pero obtendríamos una alta probabilidad de falsa alarma, lo cual podría resultar peligroso. La decisión opuesta sería establecer un umbral ligeramente superior a 15 dB, pero entonces se ignorarían muchos casos en los que sí que habría un objetivo presente, resultando en un sistema bastante ineficaz.

Por lo tanto, la decisión final supondrá establecer un punto intermedio como umbral, que debería ser cercano a los 6 dB. Para ello se realizan varias pruebas con distintos umbrales y se observan los valores obtenidos para las diferentes probabilidades, antes de tomar una decisión.

Valor de umbral (dB)	Probabilidad de detección	Probabilidad de falsa alarma	Probabilidad de pérdida
5	0,9917	0,1216	0,0083
5,2	0,9891	0,1020	0,0109
5,4	0,9852	0,0883	0,0148
5,6	0,9835	0,0766	0,0165
5,8	0,9803	0,0594	0,0197
6	0,9776	0,0495	0,0224
6,2	0,9752	0,0414	0,0248
6,4	0,9728	0,0352	0,0272
6,6	0,9663	0,0323	0,0337
6,8	0,9569	0,0303	0,0431
7	0,9486	0,0294	0,0514

Tabla 4: Resultados de probabilidad para diferentes umbrales de decisión

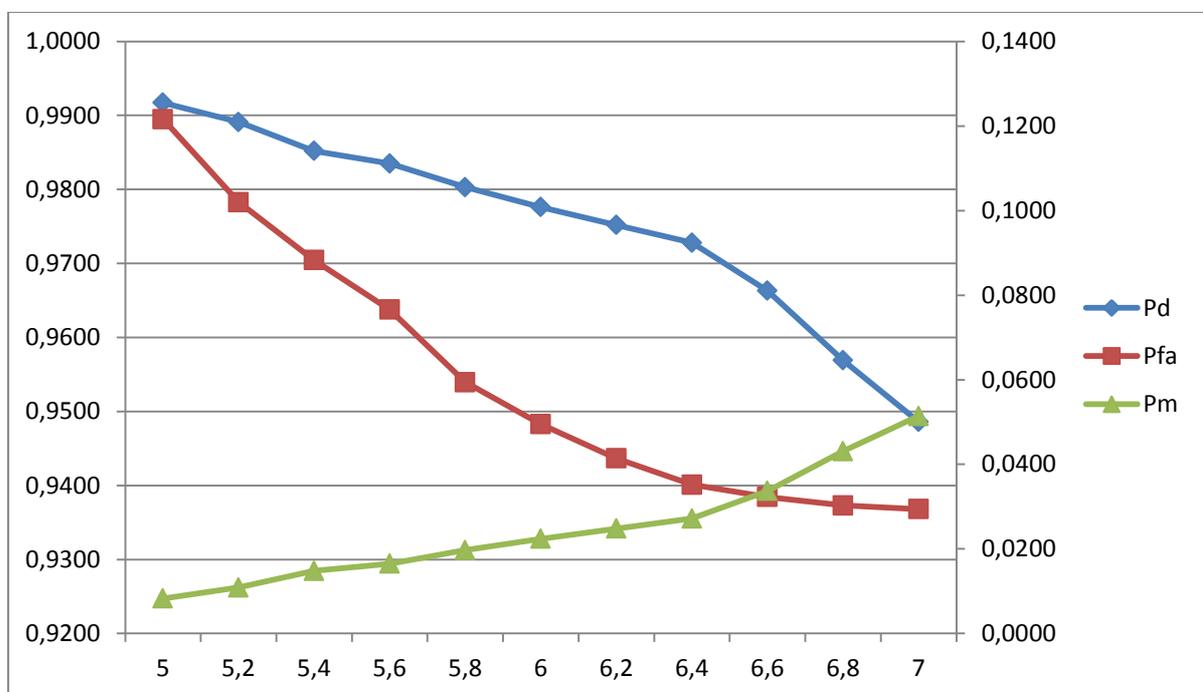


Fig. 27: Evolución de las probabilidades de detección, falsa alarma y pérdida

Se observa que, tal y como se esperaba, las probabilidades de detección y de falsa alarma decrecen conforme el valor del umbral aumenta, mientras que la probabilidad de pérdida disminuye. Prestando atención a los valores, se propone establecer el umbral en 6.4 dB, puesto que a partir de ese valor la probabilidad de detección empieza a decrecer de forma bastante considerable, y en ese punto obtenemos unos valores de probabilidad de pérdida y probabilidad de falsa alarma bastante aceptables.

4.3. Clasificación

Como se detalla anteriormente, para la realización del clasificador, se opta por la idea de trabajar con máquinas de vectores de soporte.

En este apartado, sin embargo, los resultados son bastante menos prometedores que en el anterior, puesto que se alternan casos con porcentajes de acierto bastante altos (en torno al 99%) con casos de porcentaje de acierto muy bajo (llegando incluso al 0% en casos puntuales). Tras realizar la prueba con varias muestras correspondientes a las tres categorías (peatones, coches, renos), estos son los resultados obtenidos:

Obstáculo	Mejor resultado	Peor resultado	Precisión media
Peatón	86,62%	14,36%	46,58%
Coche	89,75%	48,93%	72,77%
Reno	99,32%	0,00%	76,99%

Tabla 5: Porcentaje de precisión del clasificador para los diferentes objetivos

Por lo tanto, pese a que el porcentaje de precisión medio no se aleja mucho de los observados en estudios previos documentados (nunca superiores a un 79%, y nunca trabajando a la vez con tres clases de objetivos distintos) todavía no se puede llegar a considerar aceptable para ser implantado en un sistema de seguridad en tiempo real. Aumentar el número de muestras de entrenamiento, para que exista un mejor equilibrio con la cantidad de características extraídas, podría ayudar a que estos resultados mejoraran.

Por otro lado, la calidad de las muestras analizadas y enviadas al clasificador puede ser una de las causas de estos resultados. Mejorar las fases de análisis y filtrado podría también suponer un incremento de las prestaciones de nuestro clasificador.

Se deja, por tanto, la puerta abierta a seguir evolucionando este clasificador, así como las etapas previas, con el objetivo de mejorar estos resultados. Tarea que, por falta de tiempo, no se ha podido realizar en este PFC.

5. Conclusiones

Comparando los diferentes estudios presentados en las secciones anteriores, podemos admitir que los resultados obtenidos en este PFC son prometedores.

El principal objetivo se ha cumplido. Se ha desarrollado un sistema de detección y clasificación completo partiendo de cero. Desde la fase de recopilación de muestras de los diferentes objetivos, hasta el desarrollo de los algoritmos que realizaban las siguientes etapas.

Como bien ha sido expuesto, no hay demasiados resultados previos documentados acerca de sistemas de este tipo, por lo que las posibilidades de comparación no son muy grandes.

La metodología empleada para la toma de muestras y su posterior edición y extracción de señales se considera acertada. Los espectrogramas son una clara representación de los diferentes objetivos, y todos ellos son fácilmente identificables tras una simple observación.

Los resultados obtenidos tras la fase de filtrado también son positivos. Se eliminan casi por completo los tonos de ruido introducidos por el propio radar, y se consigue una representación de los espectrogramas mucho más limpia que permite apreciar con más claridad las primeras características observadas de los diferentes objetivos.

En cuanto al decisor, el hecho de dejar abierta la posibilidad a seleccionar de forma manual el valor umbral permite que el sistema pueda ser utilizado para diferentes situaciones, puesto que, en función de la probabilidad que se quiera maximizar (o minimizar), se puede desplazar dicho valor tanto como se desee. Sin duda, uno de los puntos fuertes de este sistema es la adaptabilidad de esta etapa.

El principal problema llega con el clasificador, 71% no es un resultado demasiado fiable para un sistema de seguridad. Aunque se trata de una posibilidad que se contemplaba desde la propuesta inicial antes de comenzar el desarrollo de este PFC. El hecho de que se trate de un sistema realizado por completo desde su fase inicial, sumado al hecho de que no haya investigaciones ni pruebas documentadas del uso de un clasificador como este para distinguir entre más de dos clases (e incluso estos no alcanzan resultados superiores a un 79%) hace que, más que considerarlo un fracaso, se piense en este resultado como punto de partida para futuras investigaciones y pruebas, que permitan mejorarlo y convertirlo en aplicable a futuros sistemas que sean capaces de trabajar en tiempo real.

5.1. Líneas futuras

Una vez este PFC ha sido terminado, se quieren proponer algunas líneas de trabajo futuras. Algunos pasos o acciones que se pueden realizar con el objetivo, tanto de mejorar los resultados obtenidos, como de ampliar las posibilidades que ofrece este campo, así como proponer algunas preguntas o casos adicionales.

En primer lugar, tal y como se explica en apartados anteriores, la tarea principal debe consistir en mejorar los resultados que ofrece el clasificador. Si el objetivo es que el trabajo realizado en este PFC sea implantado en un sistema de seguridad, la precisión, tanto de las etapas de decisión como de clasificación, debería ser lo más cercano posible al 100%.

Varias son las opciones que se proponen. En primer lugar, podría ser de utilidad aumentar la base de datos de muestras. Cuantas más muestras de entrenamiento se introduzcan en el clasificador, mejor debería funcionar. Así mismo, un posible perfeccionamiento de las etapas de análisis y filtrado puede producir que las muestras que llegan al clasificador estén mejor definidas y esto ayude a una mejor clasificación.

En cuanto al clasificador en sí, una mejora en la extracción de características ayudaría, sin duda, a una mejora en los resultados. Cuanto más compacto sea el vector de características, es decir, cuanta más información se pueda ofrecer en dicho vector, utilizando menos muestras, más fácil resultará para el clasificador la separación entre las clases, y mejor será el resultado de la clasificación final. Este tipo de clasificadores tienen problemas a la hora de procesar muestras que proporcionan información redundante, y esto puede afectar a los resultados ofrecidos.

Por otro lado, además de los resultados observados en este PFC, sería una buena idea contemplar la posibilidad de añadir nuevos objetivos. En países como Finlandia, Canadá o EEUU hay muchos problemas con renos cruzando carreteras, tal y como se presenta en la introducción del PFC. Pero esta no es una situación real en muchos otros países, por lo que este objetivo puede ser adaptable a la región donde se quiera trabajar. Perros, gatos, vacas, liebres o ciervos son también intrusos comunes en muchas de las carreteras de algunos países, entre ellos España. Pero también ciclistas o motoristas quedan fuera de este estudio y podría ser interesante incluirlos.

Otros problemas que no han sido contemplados y podrían serlo en el futuro se pueden dividir principalmente en dos grupos: trabajar con objetivos parados, que pueden interferir en la detección de nuestros objetivos ya estudiados, y trabajar con grupos, lo cual sería básicamente una evolución de lo analizado en este PFC. Grupos de personas caminando juntas, grupos de animales, o situaciones que involucren a varios vehículos, como pueden ser los atascos, o un simple adelantamiento. Todas ellas son situaciones de las que ahora mismo no se conoce el resultado que se obtendría analizándolas con este sistema, pero podría ser interesante trabajar en ellas.

Bibliografía

1. **Heuer M, Al-Hamadi A, Meinecke M-M, Mende R.** *Requirements on Automotive Radar Systems for Enhanced Pedestrian Protection.* Warsaw, Poland : 19th International Radar Symposium.
2. **Gandhi T, Trivedi M M.** *Pedestrian Protection Systems: Issues, Survey, and Challenges.* s.l. : IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, 2007. VOL. 8, NO. 3.
3. <http://www.walkinginfo.org/pedsmart>. [En línea]
4. http://prevent.ertico.webhouse.net/en/prevent_subprojects/vulnerable_road_users_collision_mitigation/apalaci/. [En línea]
5. http://www.gavrila.net/Computer_Vision/Smart_Vehicles/smart_vehicles.html. [En línea]
6. **Saebboe J, Viikari V, Varpula V, Seppä H, Cheng S, Al-Nuaimi M, Hallbjörner P, Rydberg A.** *Harmonic Automotive Radar for VRU Classification.* Stavanger (Norway), Finland, Uppsala (Sweden) : s.n.
7. **Bilik I, Tabrikian J, Cohen A.** *GMM-Based Target Classification for Ground Surveillance Doppler Radar.* s.l. : Ben-Gurion University of the Negev, Israel (<http://ieeexplore.ieee.org/>).
8. **Fölster F, Rohling H, Ritter H.** *Observation of a walking pedestrian with a 24 GHz automotive radar.* Hamburg : Hamburg University of Technology (http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-14694/Final_Papers/GM0014-F.pdf [En línea]).
9. **Heuel S, Rohling H.** *Pedestrian Classification in Automotive Radar Systems.* Hamburg, Germany : Hamburg University of Technology (<http://ieeexplore.ieee.org/>).
10. **Bartsch A, Fitzek F, Raschofer R H.** *Pedestrian recognition using automotive radar sensors.* Munich, Germany : Advances in Radio Science, 2012.
11. http://dea.brunel.ac.uk/cmstp/home_esfandiar/spectral%20subtraction.htm. [En línea]
12. **Richards M A.** *Fundamentals of Radar Signal Processing.* s.l. : McGraw-Hill, 2005.
13. **Rish I.** *An empirical study of the naive Bayes classifier.* s.l. : T.J. Watson Research Center.
14. **Murphy K P.** *Naive Bayes classifiers.* 2006 (<http://www.ic.unicamp.br/~rocha/teaching/2011s1/mc906/aulas/naive-bayes.pdf> [En línea]).
15. **Bohling G.** *Applications of Bayes' Theorem.* s.l. : Kansas Geological Survey, 2005.
16. **Hsu C-H, Chang C-C, Lin C-J.** *A Practical Guide to Support Vector Classification.* Taiwan : National Taiwan University, 2010 (<http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf> [En línea]).
17. **Childers D G, Skinner D P, Kemerait R C.** *The Cepstrum: A Guide to Processing.* s.l. : Proceedings of the IEEE, 1977.
18. **D'haes W, Rodet X.** *Discrete Cepstrum Coefficients as Perceptual Features.* Paris (France), Brussels (Belgium) : IRCAM and University of Antwerpen.

19. **Fardi B, Dousa J, Wanielik G, Elias B, Barke A.** *Obstacle detection and pedestrian recognition using a 3D PMD camera.* s.l. : Proc. IEEE Intell. Veh. Symp., 2006. págs. 225-230.
20. **Fang Y, Yamada K, Ninomiya Y, Horn B K P, Masaki I.** *A In thermal IR images, pedestrians that are warmer than the images.* s.l. : IEEE Trans. Veh. TechnolIEEE Trans. Veh. Technol., 2004. págs. 1679–1697. vol. 53, no. 6.
21. Raytheon Company, Thermal-Eye TSC. [En línea] <https://lgpssignal.com/pdf/THERMAL-Eye%20TSC.pdf>.
22. **Zhou D, Dillion M, Kwon E.** *Tracking-Based Deer Vehicle Collision Detection Using Thermal Imaging.* s.l. : IEEE International Conference on Robotics and Biomimetics, 2009.
23. **Zhou D, Wang J.** *Identification of Deer in Thermal Images to Avoid Deer-Vehicle Crashes.* University of Minnesota : International Conference on Electronics and Optoelectronics, 2011.

Anexos

A. Características de los sistemas basados en videocámaras

Una de las principales cuestiones es que los sensores de imagen dan una proyección 2-D de un escenario 3-D, perdiendo la información de la profundidad durante el proceso, puesto que la profundidad es una característica fundamental a la hora de intentar evitar colisiones. Aunque el uso de binoculares estéreos pueden recuperar dicha información, se necesita la difícil, y muchas veces ambigua, correspondencia del problema de relacionar las características entre las imágenes. Para intentar resolver este problema, un nuevo sistema con sensor 3-D usando un dispositivo mezclador fotónico es propuesto en (19). Este sistema contiene una cámara 64x16 CCD con iluminación LED emitiendo una señal IR modulada. El sensor calcula la distancia del objeto midiendo la diferencia de fase entre las señales transmitida y reflejada. Así, las imágenes de distancia y amplitud son creadas, y son usadas en el siguiente paso para identificar a los peatones.

Otra limitación explicada al principio es el trabajo en condiciones de oscuridad. Para superar estas y otras limitaciones, se puede usar una combinación de múltiples sensores que proporcionen información complementaria. Por ejemplo, la capacidad de los sensores que trabajan con luz visible diurna podría ser combinada con la capacidad de los sensores infrarrojos (IR), para obtener información 3-D a una resolución mayor.

En caso de los sensores de imagen, la forma y la apariencia de los peatones puede ser usada para separarlos del fondo. Para ello, se extraen sus características de las imágenes, y un clasificador entrenado es usado para separar a los peatones de otros objetos y del paisaje de fondo. Algunas de las características usadas son la simetría, los histogramas o los llamados contornos activos. En el caso de imágenes IR térmicas, los peatones que son más calientes que el fondo forman puntos que son usados para la detección, pero este sistema suele fallar en entornos en los que haya altas temperaturas. En (20), características basadas en el histograma, la inercia y los contrastes son usadas para distinguir peatones.

El movimiento es también una señal importante para detectar peatones. En caso de cámaras estacionarias, se sustrae el fondo para separar a los objetos en movimiento de dicho fondo estático. Aunque, en el caso de plataformas en movimiento, el movimiento propio de la cámara puede producir la sensación de un falso movimiento del propio fondo. Para el movimiento lateral de los peatones, es factible separar el movimiento de los peatones de este falso movimiento del fondo. Sin embargo, para peatones con movimiento longitudinal, el movimiento de la imagen es paralelo a este movimiento del fondo y, por tanto, difícil de separar. Además, las características de periodicidad del movimiento del cuerpo humano son útiles en la detección de peatones y su separación de otros objetos, estacionarios o en movimiento. La distribución espacial del movimiento representada por las características del momento, la distribución espectral de potencia del movimiento, las características simétricas de las piernas y los patrones de los pasos son algunas de las señales usadas para detectar y verificar a los candidatos.



Fig. 28: Certificación estéreo entre luz visible e imágenes térmicas (2). (a) Imagen en color de la entrada, (b) imagen IR térmica de la entrada, (c) superposición desalineada de las imágenes en color e IR antes de la certificación, y (d) superposición desalineada de las imágenes en color e IR después de la certificación

Después de la detección del objetivo, el siguiente paso es la clasificación. Varios tipos de clasificadores han sido utilizados para distinguir peatones de otros objetos. La entrada al clasificador es un vector de valores o características extraídas de dichos objetos, y la salida es la decisión, explicando si hay un objeto detectado o no. Los clasificadores son entrenados normalmente usando un número determinado de ejemplos positivos y negativos para determinar la frontera de decisión entre ellos. Después del entrenamiento, el clasificador procesa las muestras desconocidas y decide la presencia o ausencia del objetivo basándose en qué lado de la frontera de decisión se encuentra el vector. Los clasificadores más populares usados para la detección de peatones son los siguientes: máquinas de vectores de soporte (SVMs, Support Vector Machines), que será explicado más adelante, varios tipos de redes neurales, y clasificadores de aprendizaje estadístico, como AdaBoost.

A.1. Renos

Otro ejemplo de VRU son los animales. Accidentes de coche debidos a colisiones con renos (DVCs, Deer-Vehicle Crashes) son comunes en las carreteras secundarias de Europa y Norteamérica. Se han estudiado, y se estudian, varias opciones para evitar estos accidentes, pero pocas han tenido éxito hasta ahora.

Son muchos los métodos estudiados y probados para la detección y disuasión de reno, como reflectores o señales con flashes para la carretera, o bien cámaras térmicas implantadas en los vehículos (21). Los investigadores han intentado usar las cámaras térmicas para detectar la presencia de los renos. En este sistema, no sólo se detectan renos en un rango muy grande, sino que también se pueden rastrear para indicar si se pueden tratar de una amenaza inmediata para el vehículo.

En (22), el método del histograma de gradiente orientado (HOG) y las máquinas de vectores de soporte (SVM) se usan para identificar a los renos. En este algoritmo, basado en el cálculo del HOG en imágenes térmicas de miles de muestras (muestras térmicas de renos), el método SVM se usa en un primer momento para obtener el patrón del reno con el objetivo de describirlo. Una vez obtenido el llamado descriptor, éste se usa para compararlo con el HOG de la imagen actual. El resultado de la comparación explicará la existencia o no de renos en la imagen actual. Para aumentar la precisión, se realiza un segundo entrenamiento, donde las imágenes con falsos positivos son usadas como muestras de entrenamiento. Los resultados de las pruebas realizadas, tanto en laboratorio como en campo, muestran una precisión de hasta un 85%. Consiguiendo este objetivo, los renos pueden ser identificados, y una señal de alerta puede ser enviada al conductor. Así, la probabilidad de una colisión vehículo-reno puede ser reducida. Este método ha sido validado para la identificación de objetos como peatones, combinándolo con otros métodos de tratamiento de imagen como HOG y SVM. Hay también una gran cantidad de trabajos realizados usando este método para detectar otro tipo de objetos. Pero no hay investigaciones realizadas sobre el uso del HOG para detectar renos en imágenes térmicas para aplicaciones de seguridad en la carretera anteriores a (22).

El dispositivo de obtención de imágenes consiste en un sistema de grabado y procesado de imagen, conectado a una cámara térmica. La dirección de visión de la cámara es dirigida por dos motores con control de movimiento para obtener dos grados de libertad de movimiento, por ejemplo rotación lateral y vertical.

La cámara infrarroja usada en este trabajo es la ThermalEye™ TSCss-FF camera (23). Las imágenes capturadas pueden ser en color o monocromáticas, siendo este segundo tipo el usado en esta aplicación.

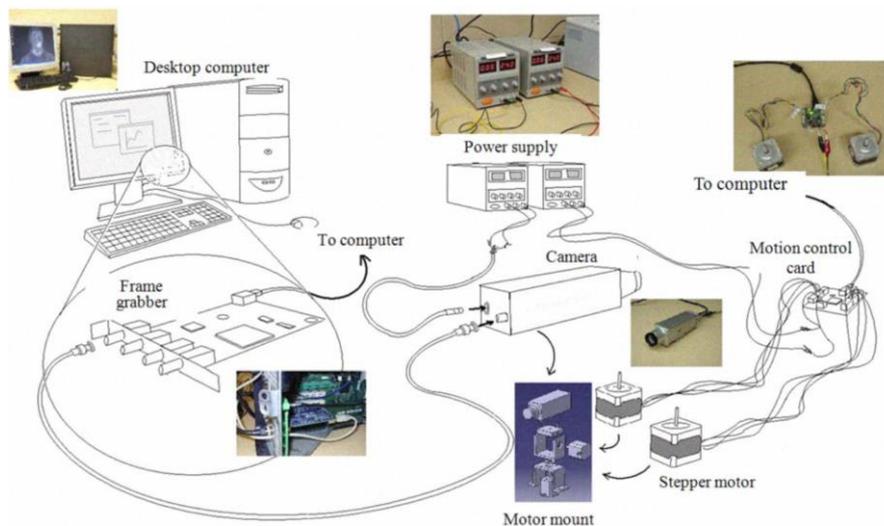


Fig. 29: Integración del sistema

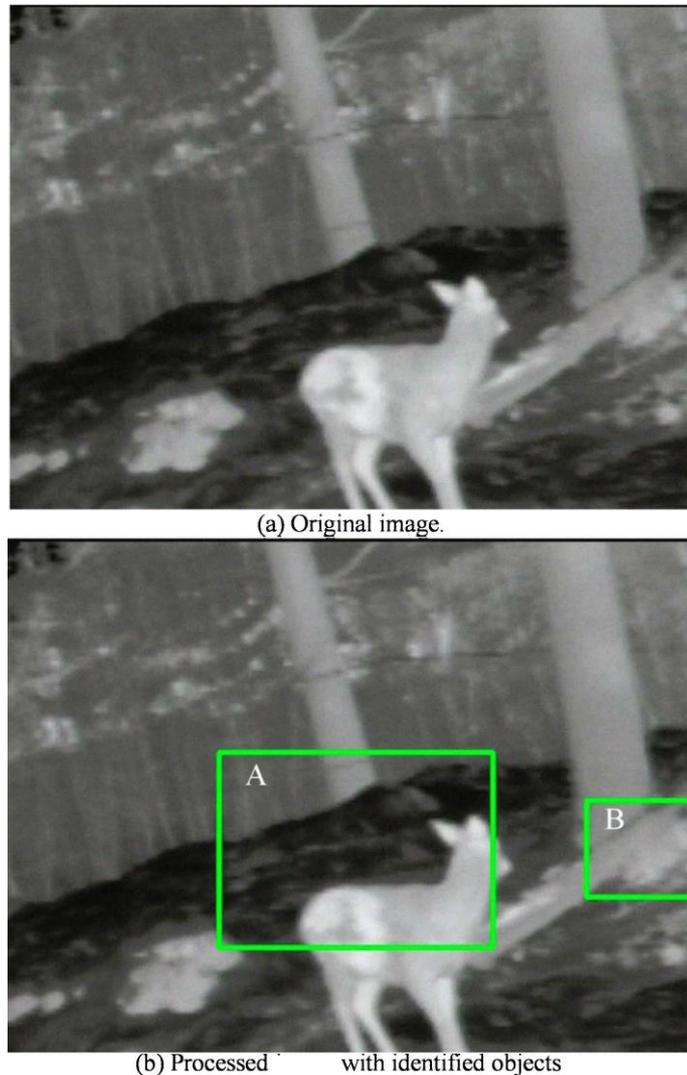


Fig. 30: Aplicación del método HOG para identificación de renos

Los métodos HOG y SVM podrían alcanzar hasta un 30% de precisión. Esto implica que otros muchos objetos podrían ser identificados como renos, lo cual hace ver que se trata de una precisión poco o nada aceptable.

Otro método, llamado “segundo entrenamiento”, se usa para mejorar esta precisión. Tras el segundo entrenamiento, la parte sin objetos (renos), como la mostrada como B en la Fig. 30(b), es cortada para formar una nueva imagen. Investigaciones posteriores muestran que la precisión es mejorada considerablemente hasta alcanzar un 85%. Por el contrario, el tiempo de procesado de cada imagen usando este método es de 1.5 segundos.

Así, el método no es apropiado para ser usado en el rastreo e identificación de renos en tiempo real. En el futuro, se espera que este tiempo de procesado pueda ser reducido para poder usarlo en este tipo de aplicaciones.

B. Aspectos matemáticos de la etapa de clasificación

En este anexo se detallan algunos conceptos matemáticos relacionados con la última etapa de las realizadas en este PFC y que no se detallan en el capítulo correspondiente de la memoria.

B.1. Cepstrum

Dos son los conceptos principales a tratar. Por un lado, el Cepstrum. Como ya se ha explicado, es un método de extracción de características que, en este caso, se utilizan para formar el vector que se introduce al clasificador en sí, que será el encargado de realizar la última fase del proceso. El Cepstrum de una señal es el resultado de aplicar una transformada y antitransformada de Fourier, realizando ciertas operaciones matemáticas en el proceso (logaritmos, módulos, cuadrados, etc.).

La fórmula que lo define es la siguiente:

$$\text{Power Cepstrum} = |F^{-1}\{\log(|F\{f(t)\}|^2)\}|^2$$

No se trata pues, de una operación lineal, pues se realiza el logaritmo, pero el resultado final contiene información relacionada con el espectro inicial de la señal. El Cepstrum de una señal muestra información sobre la tasa de cambio de las diferentes bandas del espectro.

B.2. SVM

Como se detalla en la memoria, para la fase de clasificación se utiliza un clasificador basado en máquinas de vectores de soporte (SVM). Se trata de un clasificador utilizado con éxito en bastantes aplicaciones (principalmente en el campo de reconocimiento de patrones), y cuyo objetivo es maximizar la separación entre clases, para favorecer la clasificación. Utiliza unas muestras, llamadas de entrenamiento, para una primera etapa de aprendizaje, y luego analiza las muestras, llamadas de test, para llegar al resultado final.

En lo que sigue, para realizar la explicación de las SVM, se detallará el caso binario, por su simplicidad. En el trabajo que nos ocupa en este PFC se trabaja con una SVM que ha de distinguir entre tres clases.

Una SVM toma una decisión sobre la clase del vector de características X atendiendo al signo de $D(x)$ en la siguiente expresión:

$$D(x) = \sum_{\forall x_i \in S} \alpha_i \lambda_i K(x_i, x) + \alpha_0$$

donde S representa al conjunto de entrenamiento y $\lambda_i = \pm 1$ la clase del vector de entrenamiento x_i .

La función bidimensional $K(x_i, x)$ se conoce, en el contexto de las SVM, como función Kernel, y su elección entre diversas opciones típicas (gaussiana, polinómica, etc.) tiene repercusión en el tipo de fronteras que se esperan entre las clases. En nuestro caso $K(x_i, x)$ es el producto escalar entre x_i y x , lo que se conoce como Kernel lineal. Los parámetros $\alpha_i \geq 0$ son el resultado de la solución de un proceso de optimización en la etapa de entrenamiento. El problema de optimización es:

$$\min_{\alpha} (\alpha^T \Lambda K \Lambda \alpha + C \sum_j \varepsilon_j)$$

con restricciones $\lambda_j \cdot D(x_j) \geq 1 - \varepsilon_j \quad \forall x_j$ en el conjunto de entrenamiento. Λ es una matriz diagonal con las etiquetas λ_j , y matriz K almacena todos los valores de la función Kernel, $K(x_i, x_j)$, entre pares de vectores del conjunto de entrenamiento. Las variables auxiliares ε_j hacen que sea posible cierto solapamiento entre clases, controlado por el peso de penalización, $C > 0$. Para $C = \infty$ no se permitiría solapamiento entre clases. Ocurre que los valores α_i que se obtienen como resultado del proceso de optimización son típicamente cero. Sólo son distintos de cero para un subconjunto de vectores del conjunto de entrenamiento, que se conocen como vectores soporte. Los vectores soporte permiten establecer la frontera entre clases independientemente del resto de vectores, y son los únicos que intervienen en la toma de decisión. Esta propiedad de haber aprendido automáticamente en el proceso de entrenamiento lo compleja que será la decisión en la etapa de test, es uno de los puntos fuertes de los SVM frente a otros tipos de clasificadores.

En la etapa de entrenamiento de una SVM para un problema de clasificación binaria, se dispondría de unos pares $(x_1, y_1) \dots (x_n, y_n)$, donde x es el vector de características del objeto, e y representaría la clase a la que dicho objeto corresponde, codificada con los valores $+1$ ó -1 . La idea es minimizar esta fórmula:

$$\min D(\alpha) = \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j \Phi(x_i) \cdot \Phi(x_j) - \sum_i y_i \alpha_i$$

Otra de las ventajas que tiene este clasificador, es que el margen de separación entre clases no es siempre una línea recta, si no que en función del modelo aplicado, puede generar regiones de separación más complejas, dando lugar a mejores resultados que otros clasificadores.

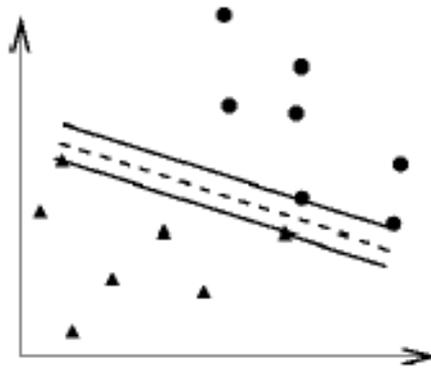


Fig. 31: Ejemplo de margen de separación entre clases

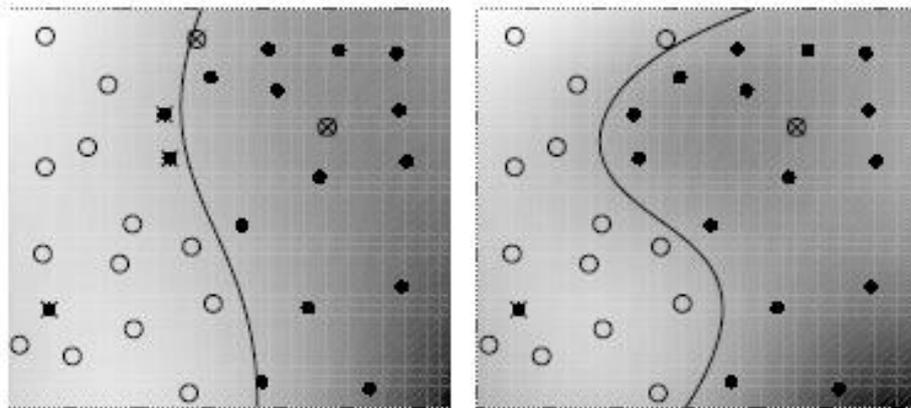


Fig. 32: Diferentes tipos de márgenes de separación

C. Scripts en MATLAB

C.1. spec_filter.m

```
% -----  
% spec_filter.m  
%  
% Master Thesis 'Classification of pedestrians, vehicles and animals  
% using automotive radar'  
% Tampere University of Technology  
% Author: Sergio Cebollada Gonzalo  
% 06/02/2013  
% -----  
% This script loads an audio FILE, draws its spectrogram and filters  
% it using the method called Spectral Substraction.  
  
function spec_filter(FILE)  
  
resfactor = 10;  
fftlength = 1024;  
over = 1000;  
  
V = 44100/resfactor/fftlength * ([-floor(fftlength/2) : 1 :  
floor(fftlength/2)-1]) * 3*10^8/(24*10^9)/2;  
% V = 44100/resfactor/fftlength * ([-floor(fftlength) : 1 :  
floor(fftlength)-1]) * 3*10^8/(24*10^9)/2 * 3.6;  
  
% 1: Extracting the FILE  
[Y,FS,NBITS] = wavread(FILE);  
  
Y = Y(:,1)+1i*Y(:,2);  
Yres = resample(Y,1,resfactor);  
Yres = Yres(1:round(length(Yres)));  
Tmax = (length(Yres)) / (44100/resfactor);  
  
% 2: Calculating the Spectrogram  
S = spectrogram(Yres,fftlength,over); % spectrogram of the signal  
S = db(S);  
S(S<-80) = -80;  
X = fftshift(S,1);  
  
% 3: Spectral Substraction Method  
N = averages; % plot of the noise estimations  
norm_N = sqrt(sum(N.^2,1));  
alpha = (N'*X)./(norm_N^2); % noise gain estimation  
Y = X - N*alpha; % spectral subtraction  
  
% 4: Median Filter  
Y2 = MEDFILT2(Y,[20 15]);  
  
% 5: Drawing the spectrogram already filtered  
T = linspace(0,Tmax,size(S,2));  
l = size(X,2);  
surf(T,V,Y2,'EdgeColor','none');  
axis([min(T) max(T) min(V) max(V)]);  
view(0,90);
```

```
xlabel('Time (s)');
ylabel('Velocity (m/s)');
% ylabel('Velocity, Km/h');
title('Spectrogram (amplitude in dB)');
colorbar

threshold = 6.4; % set this value manually
detection2(Y, threshold)
end
```

C.2. averages.m

```
% -----  
% averages.m  
%  
% Master Thesis 'Classification of pedestrians, vehicles and animals  
using automotive radar'  
% Tampere University of Technology  
% Author: Sergio Cebollada Gonzalo  
% 07/02/2013  
% -----  
% This script uses several noise FILES, and builds the Averaged power  
% spectrum density of the noise based on their spectrograms.  
  
function N = averages  
  
% 1: Building the matrix of noise spectrograms  
N = zeros(1024,10);  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,1) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise2.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,2) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise3.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,3) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise4.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,4) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise5.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,5) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise6.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,6) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise7.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,7) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise8.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,8) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise9.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,9) = av;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise10.wav';  
[av, Tmax] = average_noise(FILE);  
N(:,10) = av;
```

```

N = mean(N,2);

% 2: Drawing the Averaged power spectrum density of the noise
T = linspace(0,Tmax,size(N,1));
plot(T, N) % mirar la escala
axis tight
title('Averaged power spectrum density of the noise');
xlabel('Time, s');
ylabel('Power, dB');
end

```

C.3. average_noise.m

```

% -----
% average_noise.m
%
% Master Thesis 'Classification of pedestrians, vehicles and animals
using automotive radar'
% Tampere University of Technology
% Author: Sergio Cebollada Gonzalo
% 07/02/2013
% -----
% This script loads an audio FILE, draws its spectrogram and
calculates
% its power spectrum in db

function [av, Tmax] = average_noise(FILE)
resfactor = 10;
fftlength = 1024;
over = 1000;

% 1: Extracting the FILE
[Y,FS,NBITS] = wavread(FILE);

Y=Y(:,1)+1i*Y(:,2);

Yres = resample(Y,1,resfactor);

Yres = Yres(1:round(length(Yres)));
Tmax = (length(Yres)) / (44100/resfactor);

% 2: Calculating the Spectrogram
S = spectrogram(Yres,fftlength,over);
S = db(S);
S(S<-80) = -80;

X = fftshift(S,1);

% 3: Calculating the power spectrum in db
av = mean(X,2);
av = db(av);
end

```

C.4. detection2.m

```
% -----  
% detection2.m  
%  
% Master Thesis 'Classification of pedestrians, vehicles and animals  
using automotive radar'  
% Tampere University of Technology  
% Author: Sergio Cebollada Gonzalo  
% 26/02/2013  
% -----  
% This script decides whether the file analyzed (X) corresponds to a  
% target or not, according to the threshold (th) set.  
  
function detection2(X, th)  
  
% 1: Noise matrix  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = Tmax;  
S = det;  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise2.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise3.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise4.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise6.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise7.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise8.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
vídeos\noise9.wav';  
[det, Tmax, V] = detection_b(FILE);  
T1 = T1 + Tmax;  
S = [S det];
```

```

FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\noise10.wav';
[det, Tmax, V] = detection_b(FILE);
T1 = T1 + Tmax;
S = [S det];
Sb = sum(10.^(S./10),1)/size(S,1);
Sb = 10.*log10(Sb);

% 2: Target+noise matrix
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\car1.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = Tmax2;
S2 = det2;
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\car3.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\car5.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\car7.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\twopeople.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\car9.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\car11.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\person1.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y
vídeos\person2.wav';
[det2, Tmax2, V2] = detection_b(FILE);
T2 = T2 + Tmax2;
S2 = [S2 det2];
S2b = sum(10.^(S2./10),1)/size(S2,1);
S2b = 10.*log10(S2b);

% 3: histograms
figure
subplot(2,1,1)
T = linspace(0,T1,size(S,2));

```

```

surf(T,V,S,'EdgeColor','none');
axis([min(T) max(T) min(V) max(V)]);
title('Noise spectrograms');
xlabel('Time (s)');
ylabel('Velocity (m/s)');
subplot(2,1,2)
plot(T,Sb)
axis tight
title('Power of each time slice');
xlabel('Time (s)');
ylabel('Power (dB)');
figure
subplot(2,1,1)
T = linspace(0,T2,size(S2,2));
surf(T,V2,S2,'EdgeColor','none');
axis([min(T) max(T) min(V2) max(V2)]);
title('Target spectrograms');
xlabel('Time (s)');
ylabel('Velocity (m/s)');
subplot(2,1,2)
plot(T,S2b)
axis tight
title('Power of each time slice');
xlabel('Time (s)');
ylabel('Power (dB)');
figure
B = logspace(-4, 1.6, 2000);
P = histc(Sb, B);
P2 = histc(S2b, B);
Pb = P / sum(P);           % Normalize
P2b = P2 / sum(P2);
plot(B, P)                 % Plot noise
hold on
plot(B, P2, 'r')          % Plot signal+noise
xlabel('Power (dB)');
ylabel('Number of samples');
title('Power distribution (Blue=noise, Red=target+noise)');

% 4: Calculating the probabilities
Pd = sum(P2(B > th)) / sum(P2)
Pfa = sum(P(B > th)) / sum(P)
Pm = sum(P2(B < th)) / sum(P2)

% 5: FILE to be analyzed
det1b = sum(10.^(X./10),1)/size(X,1);
det1b = sum(det1b,2)/size(det1b,2);
det1b = 10.*log10(det1b);
if (det1b > th)
    disp('There is a target');
    disp(' ');
    classification(X)
else
    disp('There is no target');
    disp(' ');
end
end
end

```

C.5. classification.m

```
% -----  
% classification.m  
%  
% Master Thesis 'Classification of pedestrians, vehicles and animals  
using automotive radar'  
% Tampere University of Technology  
% Author: Sergio Cebollada Gonzalo  
% 16/04/2013  
% -----  
% This script builds a training vector of features extracted from  
different  
% samples and loads a FILE, which will be used as a test vector, to  
% classify it.  
  
function classification(FILE2)  
  
% 1: Training  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
videos\car1.wav';  
[det, Yres] = detection_c(FILE);  
f = features(det, Yres);  
trainX = f;  
trainY = 2*ones(size(det,1),1); % 1=human, 2=car, 3=reindeer  
  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
videos\car3.wav';  
[det, Yres] = detection_c(FILE);  
f = features(det, Yres);  
trainX = [trainX; f];  
trainY = [trainY; 2*ones(size(det,1),1)]; % 1=human, 2=car, 3=reindeer  
  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
videos\person1.wav';  
[det, Yres] = detection_c(FILE);  
f = features(det, Yres);  
trainX = [trainX; f];  
trainY = [trainY; ones(size(det,1),1)]; % 1=human, 2=car, 3=reindeer  
  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
videos\person2.wav';  
[det, Yres] = detection_c(FILE);  
f = features(det, Yres);  
trainX = [trainX; f];  
trainY = [trainY; ones(size(det,1),1)]; % 1=human, 2=car, 3=reindeer  
  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
videos\deer1.wav';  
[det, Yres] = detection_c(FILE);  
f = features(det, Yres);  
trainX = [trainX; f];  
trainY = [trainY; 3*ones(size(det,1),1)]; % 1=human, 2=car, 3=reindeer  
  
FILE = 'C:\Users\Sergio\Documents\CPS\Proyecto\Audios y  
videos\deer2.wav';  
[det, Yres] = detection_c(FILE);
```

```

f = features(det, Yres);
trainX = [trainX; f];
trainY = [trainY; 3*ones(size(det,1),1)]; % 1=human, 2=car, 3=reindeer

% 2: Testing
[det2, Yres2] = detection_c(FILE2);
l = lab(FILE2);
f2 = features(det2, Yres2);
testX = [f2; f2; f2; f2; f2; f2];
testY = [1*ones(size(det,1),1); 1*ones(size(det,1),1);
1*ones(size(det,1),1); 1*ones(size(det,1),1); 1*ones(size(det,1),1);
1*ones(size(det,1),1)];

[tsACC, cv_acc, tspY, Opt_sigma, Opt_c, model] =
SVM_CV_estimate(trainX,trainY,testX,testY);
x = mean(tspY);
result(x)
end

```

C.6. features.m

```

% -----
% features.m
%
% Master Thesis 'Classification of pedestrians, vehicles and animals
using automotive radar'
% Tampere University of Technology
% Author: Sergio Cebollada Gonzalo
% 17/04/2013
% -----
% This script extracts the features of the file (S) loaded, and builds
% the features vector.

function f = features(S, Yres)

f = (abs(ifft((S(:,1)./10)))).^2;
f = sum(f,1)/size(f,1); % feature 1=mean
f2 = (abs(ifft(log10((abs(fft(Yres))))).^2))).^2; % feature 2=cepstrum

for i = 2:size(S,1)
    f_aux = (abs(ifft((S(:,i)./10)))).^2;
    f_aux = sum(f_aux,1)/size(f_aux,1);
    f = [f; f_aux];
end
f = [f; f2];
end

```