



Universidad de Valladolid

ESCUELA DE INGENIERÍAS INDUSTRIALES

INGENIERÍA DE SISTEMAS Y AUTOMÁTICA

TESIS DOCTORAL:

**INSPECCIÓN VISUAL AUTOMÁTICA DE
SUPERFICIES CONTINUAS CARACTERIZANDO
ANOMALÍAS LOCALES EN EL DOMINIO
ESPACIO-FRECUENCIAL**

Presentada por Roberto Medina Aparicio para optar
al grado de doctor por la Universidad de Valladolid

Dirigida por:

Prof. Jaime Gómez García-Bermejo

Prof. Eduardo Zalama Casanova

RESUMEN

Esta tesis propone una metodología general para la inspección visual automática de superficies continuas que abarca las etapas de adquisición de imágenes, su procesamiento y la utilización de los resultados obtenidos.

El objetivo último es la determinación de qué zonas de las superficies bajo análisis son defectuosas por alejarse de la homogeneidad esperada y cuál es el tipo de defecto presente. Para ello, se propone un nuevo enfoque para el procesamiento de las imágenes que se basa en la caracterización de anomalías en el dominio espacio-frecuencial, explotando las posibilidades que ofrece el filtro de Gabor. Se ha definido una metodología para el diseño de bancos de filtros de Gabor que analiza una zona del espacio de frecuencias y orientaciones. Se propone que la información extraída por estos filtros sean las características que se evalúen en la detección y clasificación de defectos.

Este enfoque general ha sido particularizado a la resolución de tres problemas reales de reconocida trascendencia: la inspección de la superficie de bobinas de chapa de acero laminado, del pavimento de carreteras y del revestimiento de túneles de hormigón.

Para la inspección de bobinas de chapa de acero laminado se propone una arquitectura jerárquica capaz de manipular las imágenes generadas por un conjunto de cámaras lineales. La detección y clasificación de 6 tipos de defectos es realizada mediante la combinación de tres clasificadores que utilizan características extraídas de imágenes filtradas por filtros de Gabor. Todo el proceso transcurre en tiempo real, a la velocidad de corte de las bobinas.

La inspección del pavimento de carreteras se realiza mediante un vehículo equipado con el sistema de adquisición de imágenes LRIS. Las imágenes provistas por este sistema son acompañadas con la información sincronizada que proporcionan dos perfilómetros láser, una cámara de color para visualizar el entorno de la zona inspeccionada y un GPS para georreferenciar toda la información. Las grietas, uno de los deterioros más habituales y representativos del estado de la carretera, son detectadas y clasificadas en una única etapa combinando clasificadores binarios que utilizan características de imágenes filtradas por filtros de Gabor mediante una adaptación del algoritmo AdaBoost.

Por último, para la inspección de la superficie del revestimiento de túneles de hormigón, se ha definido una arquitectura que permite la adquisición del perímetro completo de túneles de distintos diámetros, con el objetivo de detectar grietas y fisuras cuya orientación no es conocida a priori. Para ello se propone un nuevo filtro que hemos denominado filtro de Gabor invariante a la rotación, así como una variante para el cálculo de la textura de anisotropía condicional. La combinación de ambas técnicas permite la detección precisa de los defectos buscados.

La dos primeras soluciones propuestas en esta tesis se encuentran ya operativas a nivel industrial, mientras que la tercera se encuentra aún en fase de prototipo.

AGRADECIMIENTOS

Esta tesis contiene los avances realizados en los últimos 8 años en la inspección de superficies continuas durante los cuales mis compañeros de trabajo de la Fundación CARTIF y mis directores de tesis han contribuido significativamente, y mi familia y amigos me han proporcionado el apoyo necesario para seguir adelante.

En primer lugar, me gustaría agradecer a mis directores de tesis, Dr. Jaime Gómez y Dr. Eduardo Zalama, su esfuerzo y dedicación en la elaboración de esta tesis, así como la confianza depositada durante todos estos últimos años. Sus conocimientos y orientaciones han sido esenciales para llevar esta tesis a buen puerto.

También quiero agradecer el apoyo recibido de mis compañeros de trabajo de la división de Robótica y Visión Artificial de la Fundación CARTIF, sin los cuales nunca hubiera sido posible realizar esta investigación.

Los sistemas de inspección implementados en esta tesis han sido fruto de la colaboración con tres empresas, Gonvarri, Corporación Llorente y Dragados, a cuya dirección y personal quiero agradecer la ayuda prestada, especialmente en la identificación y descripción del contenido de las imágenes.

Finalmente quiero dar las gracias a mis familiares y amigos por el ánimo recibido y la paciencia mostrada. Especialmente debo agradecer el empeño que mis padres han puesto en que llegue hasta aquí y la paciencia y comprensión de mi pareja, que siempre estuvo a mi lado durante la preparación de esta tesis, dándome fuerzas para seguir adelante.

CONTENIDO

1	Introducción	1
1.1	Marco de la investigación	1
1.2	Objetivos y alcance de la tesis.....	5
1.3	Principales aportaciones.....	6
1.4	Producción científica e industrial.....	7
1.4.1	Publicaciones.....	7
1.4.2	Producción industrial.....	10
1.5	Organización de la tesis	11
2	Inspección visual automática de superficies continuas.....	13
2.1	Descripción de la problemática	13
2.2	Características comunes para los sistemas de inspección visual de superficies continuas.....	15
2.3	Inspección de bobinas de chapa laminada	17
2.3.1	Complejidad de un sistema automático para la inspección de chapa laminada	19
2.3.2	Elementos clave de la estructura hardware de un sistema de inspección superficial de chapa laminada.....	20
2.3.3	Procesamiento de imágenes	22
2.4	Inspección de la superficie de carreteras.....	28
2.4.1	Principales deterioros en pavimentos bituminosos.....	29
2.4.2	Sistemas de adquisición de imágenes de la superficie de la carretera	30
2.4.3	Procesamiento de imágenes	38
2.5	Inspección del revestimiento de túneles de hormigón.....	41
2.5.1	Defectos en túneles de hormigón.....	42
2.5.2	Caracterización de las fisuras.....	46
2.5.3	Sistemas comerciales para la inspección de túneles	46
2.5.4	Procesamiento de imágenes	48
2.6	Conclusiones.....	51

3	Arquitectura para la inspección de superficies continuas.....	53
3.1	Etapas para la inspección visual de superficies continuas	53
3.2	Adquisición	54
3.2.1	Iluminación	55
3.2.2	Configuración	56
3.2.3	Sincronización	57
3.2.4	Almacenamiento	58
3.3	Procesamiento	59
3.3.1	Preprocesamiento	59
3.3.2	Detección.....	60
3.3.3	Clasificación	60
3.4	Utilización de los resultados	61
3.4.1	Visualización de la información obtenida	61
3.4.2	Generación de informes	62
3.4.3	Toma de decisiones	62
3.5	Conclusiones.....	62
4	Técnicas de detección y clasificación de defectos	65
4.1	Introducción	65
4.2	Evaluación del desempeño de la detección y clasificación de defectos superficiales	66
4.2.1	Clasificadores binarios.....	67
4.2.2	Clasificadores no binarios	70
4.3	Cálculo de umbrales para clasificadores binarios	71
4.4	Combinación de clasificadores binarios mediante el algoritmo de entrenamiento AdaBoost	72
4.5	Técnicas para la clasificación de tipos de defectos.....	74
4.5.1	Redes neuronales	75
4.5.2	Clasificador de K vecinos más cercanos	82
4.5.3	Clasificador bayesiano. Clasificador naïve Bayes	86
4.6	Combinación de clasificadores de tipos de defectos	88
4.7	Conclusiones.....	90

5	Análisis espacio-frecuencial de las imágenes mediante filtros de Gabor	93
5.1	Motivación	93
5.2	El filtro de Gabor	104
5.3	Diseño de bancos de filtros de Gabor	105
5.4	Análisis de las imágenes filtradas.....	113
5.5	Filtro de Gabor invariante a la rotación.....	115
5.6	La textura de anisotropía condicional en imágenes filtradas por filtros de Gabor	118
5.7	Conclusiones.....	123
6	Sistemas de inspección implementados	125
6.1	Introducción	125
6.2	Inspección superficial de bobinas de chapa laminada	125
6.2.1	Arquitectura jerárquica.....	126
6.2.2	Adquisición de imágenes	127
6.2.3	Transmisión de la información y sincronización.....	129
6.3	Inspección del estado de la superficie de la carretera	132
6.3.1	Equipos para la captura de datos	132
6.3.2	Configuración y sincronización de la adquisición de datos.....	134
6.3.3	Visualización de la información obtenida	136
6.4	Inspección de grietas y fisuras en túneles	137
6.4.1	Adaptación del sistema de inspección al tipo de túnel y a la resolución de la inspección	138
6.4.2	Descripción del prototipo desarrollado	144
6.5	Conclusiones.....	151
7	Resultados	153
7.1	Detección y clasificación de defectos superficiales en chapa laminada....	154
7.1.1	Técnicas de procesamiento de imágenes utilizadas.....	154
7.1.2	Resultados obtenidos	162
7.1.3	Conclusiones	170

7.2	Detección y clasificación de grietas en la superficie de la carretera	172
7.2.1	Técnicas de procesamiento de imagen utilizadas	172
7.2.2	Resultados obtenidos	178
7.2.3	Conclusiones	190
7.3	Detección y clasificación de grietas en la superficie de túneles	194
7.3.1	Filtro de Gabor invariante a la rotación	196
7.3.2	Textura de anisotropía condicional	200
7.3.3	Textura de anisotropía condicional de la imagen filtrada por un filtro de Gabor invariante a la rotación.....	205
7.3.4	Conclusiones	208
8	Conclusiones.....	211
	Bibliografía	215

LISTA DE FIGURAS

Figura 1. Superficies continuas objeto de estudio en esta tesis. (a) Bobina de acero laminado; (b) Porción de carretera en la que se aprecia una grieta en carretera; (c) Túnel de hormigón	14
Figura 2. Imágenes de algunos de los defectos que pueden aparecer en la superficie de una bobina de chapa laminada. (a) Marcas de asentamiento; (b) Moho blanco; (c) Impresiones; (d) Helechos; (e) Viruela; (f) Grumos de cinc	20
Figura 3. Descripción de los deterioros en firmes según el "Catálogo de deterioros en firmes" publicado por el MOPU en Abril de 1989, [MOP89]	29
Figura 4. Imágenes de algunos de los defectos que pueden aparecer en la superficie de la carretera. (a) Grieta longitudinal; (b) Fisuras finas; (c) Grieta transversal; (d) Piel de cocodrilo; (e) Cuarteo de malla gorda; (f) Fisura sellada	30
Figura 5. Sistemas de ayuda para la recogida de datos de defectos en la carretera. A la izquierda VIZIROAD. A la derecha DESY.....	31
Figura 6. Sistema de inspección de carreteras Roadrecon de PASCO. Fuente Highway & Airport Research Lab, http://pavement.hanyang.ac.kr/research/pms.html#	32
Figura 7. Sistema de inspección de carreteras ARAN de Fugro Roadware. [Fug15]	33
Figura 8. Sistema de inspección de carreteras PicCrack. Fuente: http://www.go-explore-trans.org/david-lee/2/	35
Figura 9. Laser Road Imaging System (LRIS), [INO15]	36
Figura 10. Fisuras en la superficie de un túnel	42
Figura 11. Desprendimiento en una parte de una dovela.....	43
Figura 12. Armadura vista en dovela de hormigón.....	43
Figura 13. Eflorescencia en el interior de un túnel	44
Figura 14. Diversos grados de humedades en el interior de un túnel.....	45
Figura 15. Sistema de escaneado de túneles LTSS de Pavemetrics.....	47
Figura 16. Sistema de detección de fisuras en túneles tCrack de terra vermessungen.....	48

Figura 17. Etapas para la inspección visual automática	54
Figura 18. Ejemplo de curva ROC. Esta curva corresponde a uno de los clasificadores utilizados en la sección 7.2 para la detección grietas longitudinales en la superficie de la carretera. Se han marcado los puntos de la recta correspondientes a los umbrales obtenidos al utilizar los métodos de cálculo de umbrales descritos en la sección 4.3.....	69
Figura 19. Modelo completo de una neurona artificial	77
Figura 20. Modelo de red multicapa	78
Figura 21. Algoritmo backpropagation	80
Figura 22. Dos imágenes de 512x512 píxeles con una resolución de 1mm de grietas en carretera y su segmentación mediante el método de Otsu. En la fila superior una grieta longitudinal. En la final inferior una grieta transversal.	94
Figura 23. A la izquierda, dos grietas en la superficie de un túnel de hormigón con una apertura aproximada de 1mm; a su derecha, el resultado de la segmentación de la imagen utilizando el método Otsu. Imágenes de 1024x918 píxeles adquiridas con una resolución de 0,2mm/píxel.....	95
Figura 24. A la izquierda, la grieta longitudinal de la Figura 22 segmentada eligiendo el umbral manualmente; a la derecha la misma imagen tras la operación morfológica de cierre	96
Figura 25. Dos imágenes filtradas por el Filtro Prewitt. La columna de la izquierda es una grieta de carretera, la columna de la derecha es una grieta en un túnel. La fila superior es la imagen original preprocesada, la fila central es el resultado de aplicar el filtro de Prewitt, la fila inferior es la imagen filtrada segmentada por el método Otsu	97
Figura 26. Dos imágenes filtradas por el Filtro Sobel. La columna de la izquierda es una grieta de carretera, la columna de la derecha es una grieta en un túnel. La fila superior es la imagen original preprocesada, la fila central es el resultado de aplicar el filtro de Sobel, la fila inferior es la imagen filtrada segmentada por el método Otsu...	98
Figura 27. Dos imágenes de grietas en carretera y en túneles a los que se les han aplicado los filtros de Canny y Roberts.....	99
Figura 28. Dos imágenes filtradas por el filtro top-hat con un elemento estructural circular de radio 7. La columna de la izquierda es una grieta de carretera, la columna de la derecha es una grieta en un túnel. La fila superior es la imagen original preprocesada, la fila central es el resultado de aplicar el filtro top-hat, la fila inferior es la imagen filtrada segmentada por el método Otsu	100

Figura 29. La elipse es la isocurva de valor de medio pico de la salida de un filtro de Gabor en el espacio de la frecuencia. Para cumplir el teorema de Nyquist-Shannon esta isocurva no debe superar el valor de frecuencia de 0,5	107
Figura 30. (a) Solapamiento radial. (b) Solapamiento tangencial.....	109
Figura 31. Representación de las curvas de nivel de medio pico. Las curvas de nivel de medio pico son tangentes entre sí tanto en la dirección radial como en la tangencial.	110
Figura 32. Esquema de la metodología definida para el cálculo de los parámetros de cada filtro de Gabor de un banco de filtros.....	113
Figura 33. Probabilidad de que una línea de la imagen sea defecto dependiendo del nivel de gris medio de la imagen. En rojo la ecuación propuesta por Nguyen et al. en [Ngu09], en azul la ecuación propuesta en este trabajo	121
Figura 34. Arquitectura jerárquica del sistema de inspección visual de chapa laminada.....	126
Figura 35. Módulos de la cara superior del sistema de inspección de chapa laminada. Cada módulo dispone de cámara e iluminación. Las cubiertas para aislar el sistema de la iluminación ambiente han sido retiradas.	128
Figura 36. Diagrama de flujo del proceso de adquisición y procesamiento de imágenes.	130
Figura 37. Puesto de control del sistema de inspección	131
Figura 38. Capturas de pantalla del software de visualización de resultados	132
Figura 39. Vehículo utilizado para la inspección de carreteras. (a) Vista trasera (b) Vista frontal	133
Figura 40. Captura de pantalla del software de adquisición de datos.....	135
Figura 41. Interfaz para mostrar la información. (a) Interfaz para mostrar la información sin procesar de todos los dispositivos (b) Interfaz para mostrar el mapa de defectos y la información extraída de las imágenes.....	137
Figura 42. Secciones tipo de túnel ferroviario. En rojo las zonas de trabajo donde es posible situar las cámaras (a) Túnel del AVE Chamartín-Atocha; (b) Línea 1 del metro de Sevilla; (c) Túnel de interconexión eléctrica de Inelfe	139
Figura 43. Diseño del prototipo de inspección de túneles ferroviarios.....	144
Figura 44. Plataforma de inspección en los túneles tipo descritos.....	145

Figura 45. Prototipo de la plataforma de inspección de túneles.....	145
Figura 46. Ruedas para el apoyo de la plataforma. (a) Rueda sobre vía de tren con sistema de guiado. (b) Rueda con el codificador incremental acoplado.....	146
Figura 47. Sistema cámara-iluminación. (a) Cámara e2v Aviiva SM2 CL 1010 con óptica Tamron AF 70-300mm F/4-5.6 Di LD Macro 1:2. (b) Iluminación halógena de 100W con reflector AR111 para concentrar la luz con un ángulo de 8°	147
Figura 48. Esquema del software de adquisición y control.....	147
Figura 49. Captura de pantalla del software de interfaz de usuario	149
Figura 50. Captura de pantalla del módulo de adquisición de imágenes.....	150
Figura 51. Captura de pantalla del módulo de visualización de datos. (a) Datos brutos (b) Mapa de defectos.....	150
Figura 52. Proceso de segmentación	156
Figura 53. Imágenes de los defectos y sus correspondientes imágenes filtradas por uno de los filtros seleccionados. (a) soldadura, (b) moho blanco, (c) marcas de asentamiento, (d) calvas, (e) grumos de cinc, (f) impresiones, (g), (h), (i) son posibles defectos que no son defectos. El Filtro 1 de la Tabla 8 se aplica a (d), (c), (i). El Filtro 2 a (e), (h). El Filtro 3 a (b), (g). El Filtro 4 a (a). El Filtro 5 a (c)	160
Figura 54. Muestras de los defectos objeto de estudio adquiridos por el sistema durante su funcionamiento	163
Figura 55. Dos soldaduras (a) y (b); y dos marcas de transporte (c) y (d). Puede verse la gran variabilidad en la apariencia visual de defectos del mismo tipo.....	164
Figura 56. (a) imagen adquirida con el LRIS sin defectos. (b) Imagen preprocesada. (c) Imagen adquirida con el LRIS con presencia de grietas. (d) Imagen preprocesada	174
Figura 57. Dos filtros diferentes en el dominio espacial y frecuencial. El rango de valores de la parte real e imaginaria es de -1 a 1, el del módulo de 0 a 1 y la fase de -180° a 180°. (a) Filtro del banco 4. La frecuencia central es 0,010 píxeles ⁻¹ y la orientación es 45°. (b) Filtro del banco 3. La frecuencia central es 0,013 píxeles ⁻¹ y la orientación es 0°.	177
Figura 58. Imágenes de grietas transversales utilizadas en el entrenamiento	179
Figura 59. Imágenes de grietas longitudinales utilizadas en el entrenamiento ...	179

Figura 60. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 1 para grietas transversales.....	184
Figura 61. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 2 para grietas transversales.....	184
Figura 62. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 3 para grietas transversales.....	185
Figura 63. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 1 para grietas longitudinales.....	185
Figura 64. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 2 para grietas longitudinales.....	186
Figura 65. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 3 para grietas longitudinales.....	186
Figura 66. De izquierda a derecha, imagen preprocesada y el valor del módulo de la imagen filtrada por los cuatro mejores filtros de Gabor seleccionados por cada clasificador fuerte. Debajo de cada imagen, la inecuación que debe cumplir cada clasificador; las condiciones que son satisfechas están escritas en verde, mientras que las que no están en rojo. (a) Grietas transversales. (b) Grietas longitudinales.....	191
Figura 67. (a) y (c) son imágenes preprocesadas de grietas en túneles. (b) y (d) son la segmentación manual de estas grietas.....	195
Figura 68. Evolución del error ponderado frente al número de generaciones al aplicar el algoritmo genético Evolución Diferencial para establecer los parámetros del filtro de Gabor invariante a la rotación para la detección de grietas en túneles.....	198
Figura 69. Ejemplo de aplicación del filtro de Gabor invariante a la rotación. (a) y (d) son las imágenes originales normalizadas; (d) y (e) son las imágenes filtradas; (c) y (f) las imágenes segmentadas	199
Figura 70. Ejemplo de aplicación del método original de textura de anisotropía condicional. (a) y (d) son las imágenes originales normalizadas; (d) y (e) son las imágenes filtradas; (c) y (f) las imágenes segmentadas	201
Figura 71. Ejemplo de aplicación de variante del método de textura de anisotropía condicional en el que solo se utiliza la media. (a) y (d) son las imágenes originales normalizadas; (d) y (e) es la imagen filtrada; (c) y (f) las imágenes segmentadas...	203
Figura 72. Ejemplo de aplicación de variante del método de textura de anisotropía condicional en el que se utiliza la función sigmoide para calcular la probabilidad. (a) y	

(d) son las imágenes originales normalizadas; (d) y (e) es la imagen filtrada; (c) y (f) las imágenes segmentadas..... 205

Figura 73. Evolución del error ponderado frente al número de generaciones al aplicar el algoritmo genético Evolución Diferencial para establecer los parámetros del filtro de Gabor invariante a la rotación utilizado para filtrar las imágenes de entrada al algoritmo CTA para la detección de grietas en túneles 207

Figura 74. Ejemplo de aplicación de la textura de anisotropía condicional de la imagen filtrada de un filtro de Gabor invariante a la rotación. (a) y (d) son las imágenes originales normalizadas; (d) y (e) es la imagen filtrada; (c) y (f) las imágenes segmentadas..... 208

LISTA DE TABLAS

Tabla 1. Comparativa de sistemas de inspección de carreteras.....	37
Tabla 2. Matriz de confusión. Relación entre los clasificadores y la clasificación realizada por un experto	68
Tabla 3. Matriz de confusión de un clasificador no binario.....	71
Tabla 4. Máxima y mínima distancia de trabajo (DT) para cada uno de los tres tipos de túneles propuestos. En un entorno de 25cm al radio del túnel (R_T) tampoco será posible situar las cámaras	140
Tabla 5. Cinco cámaras lineales tipo con diferentes resoluciones	140
Tabla 6. Resolución y número de cámaras para los distintos tipos de túneles estudiados	143
Tabla 7. Parámetros del banco de filtros de Gabor para la clasificación de defectos en chapa laminada	158
Tabla 8. Parámetros de los filtros de Gabor utilizados.....	159
Tabla 9. Matriz de confusión del clasificador red neuronal (ANN).....	166
Tabla 10. Matriz de confusión del clasificador de vecinos cercanos (KNN)	167
Tabla 11. Matriz de confusión del clasificador Bayesiano	167
Tabla 12. Matriz de confusión de la combinación de los tres clasificadores empleados	168
Tabla 13. Resultados de sensibilidad (Sen.) y especificidad (Esp.) de las muestras utilizadas para la red neuronal (ANN), el método de los vecinos cercanos (KNN), el clasificador Bayesiano y la combinación de los tres, aplicado a cada uno de los defectos estudiados	168
Tabla 14. Parámetros de los bancos de filtros de Gabor	176
Tabla 15. Muestras utilizadas en el entrenamiento y el test para la detección de grietas transversales y longitudinales	180
Tabla 16. Filtros individuales que han alcanzado los menores errores para los tres métodos utilizados para grietas transversales.....	181
Tabla 17. Filtros individuales que han alcanzado los menores errores para los tres métodos utilizados para grietas longitudinales.....	181

Tabla 18. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas transversales por el método 1	182
Tabla 19. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas transversales por el método 2 y 3.....	182
Tabla 20. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas longitudinales por el método 1	183
Tabla 21. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas longitudinales por el método 2 y 3.....	183
Tabla 22. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del clasificador fuerte seleccionado para detectar grietas transversales	189
Tabla 23. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del clasificador fuerte seleccionado para detectar grietas longitudinales	189
Tabla 24. Comparativa de distintos clasificadores.....	192
Tabla 25. Parámetros del banco de filtros de Gabor invariante a la rotación de la generación inicial utilizada por el algoritmo genético de Evolución Diferencial	196
Tabla 26. Matriz de confusión obtenida en la detección de grietas en túneles mediante la utilización de filtros de Gabor invariantes a la rotación	198
Tabla 27. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante un filtro de Gabor invariante a la rotación	198
Tabla 28. Parámetros de los filtros de Gabor en las generaciones 19 y 141	199
Tabla 29. Matriz de confusión obtenida en la detección de grietas en túneles mediante el algoritmo CTA descrito por Nguyen et al.....	200
Tabla 30. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante el algoritmo CTA de Nguyen et al.....	201
Tabla 31. Matriz de confusión obtenida en la detección de grietas en túneles mediante el algoritmo CTA con un único valor del vector de características y utilizando la media de la imagen en un entorno circular	202

Tabla 32. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante el algoritmo CTA con un único valor del vector de características y utilizando la media de la imagen en un entorno circular.. 202

Tabla 33. Matriz de confusión obtenida en la detección de grietas en túneles mediante el algoritmo CTA cuya función de probabilidad se ajusta a una sigmoide.. 204

Tabla 34. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante el algoritmo CTA cuya función de probabilidad se ajusta a una sigmoide 204

Tabla 35. Matriz de confusión obtenida en la detección de grietas en túneles mediante la aplicación de una modificación del algoritmo CTA a imágenes filtradas por un filtro de Gabor invariante a la rotación 206

Tabla 36. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante la aplicación de una modificación del algoritmo CTA a imágenes filtradas por un filtro de Gabor invariante a la rotación... 206

Tabla 37. Parámetros de los filtros de Gabor en las generaciones 27 y 50..... 207

1 INTRODUCCIÓN

1.1 Marco de la investigación

La inspección visual de superficies continuas se realiza para determinar qué zonas son defectuosas por alejarse de la homogeneidad esperada y cuál es el tipo de defecto presente. Es una necesidad actual tanto en el mundo de la industria como en el de las infraestructuras civiles. Este tipo de superficies se caracteriza por tener un ancho acotado y una longitud indeterminada pero mucho mayor que su anchura, así como una apariencia visual esperada homogénea en toda la superficie, en el sentido de que, cuando se observa cualquier pequeña porción de la superficie, no se distingue a qué zona de la misma pertenece. Existen multitud de productos industriales que se ajustan a esta definición como telas, papel, plásticos, papel de aluminio, parqués, metales, madera, etc. Así mismo, las infraestructuras lineales, como las carreteras, el revestimiento de los túneles, las vías del tren, etc., también se pueden catalogar como superficies continuas homogéneas.

El marco de investigación de esta tesis se centra en la inspección de tres tipos de superficies continuas: la superficie de bobinas de chapa de acero laminado, el pavimento de carreteras y el revestimiento de túneles de hormigón. El desarrollo de sistemas automáticos para su inspección tiene un gran interés en la actualidad, tanto por la gran demanda existente para la inspección de este tipo de superficies, como por la dificultad que supone su inspección, aunque las técnicas descritas en esta tesis también pueden generalizarse a otro tipo de superficies. Estas superficies son inspeccionadas en movimiento y a gran velocidad, generando un volumen de información muy elevado, por lo que la adquisición de las imágenes supone un gran reto en sí misma. La gran variabilidad de la apariencia visual que presenta un mismo tipo de defecto pero a su vez el escaso contraste respecto a la zona libre de defectos constituye una gran dificultad a la hora de procesar automáticamente las imágenes.

En particular, el acero es el material más utilizado en la ingeniería y en la construcción, siendo el acero laminado su uso más habitual. En 2013 la producción de acero fue de 1.582 millones de toneladas, [Wor14]. El acero laminado es habitualmente suministrado en bobinas de varios kilómetros de longitud y un ancho de aproximadamente dos metros. Estas bobinas son procesadas en líneas de corte de alta velocidad por empresas dedicadas expresamente a esta tarea. Los defectos superficiales presentes sobre estas bobinas deberían ser detectados y clasificados con

precisión para que los responsables de calidad de la empresa puedan decidir si el producto es apropiado para un determinado uso o cliente.

La **inspección visual de la superficie de chapa de acero laminado** en la línea de producción presenta grandes retos. Por un lado, la velocidad de producción de la línea puede alcanzar los 200m/min lo que impide a los operarios realizar la inspección del 100% de la producción a la velocidad normal de funcionamiento de la línea. Además, el emplazamiento donde se realiza la inspección es peligroso, con pesados equipos moviéndose dentro y fuera de la instalación durante el proceso normal de trabajo. Por otro lado, el número de tipos de defectos es muy elevado, la apariencia visual difiere en gran medida entre defectos del mismo tipo y el acero se encuentra lubricado, lo que enmascara algunos defectos.

Por su parte, la red de carreteras pavimentadas es una infraestructura lineal extremadamente extensa, alcanzando los 4,4 millones de kilómetros en la Unión Europea (EU27), [Eur12]. El gasto de mantenimiento también es muy elevado, siendo el valor recomendado por la Unión Europea del 2% del valor de la infraestructura, [Lax01], y de 22.654 millones de euros en Estados Unidos en una red de 5 millones de kilómetros, [OEC11]. Para decidir cómo se dedica el abultado presupuesto de mantenimiento y rehabilitación de carreteras se requiere disponer de información precisa sobre el estado de la red vial, para lo cual se utilizan sistemas de gestión del pavimento, [Lep99].

Estos sistemas requieren la recolección y evaluación de grandes volúmenes de datos sobre el estado del pavimento entre los que destaca el deterioro superficial de la carretera. Este deterioro es evaluado mediante la **inspección visual de la carretera**, que habitualmente la llevan a cabo operarios que caminan por el arcén y anotan los defectos que observan. Además de ser una tarea muy peligrosa, los resultados que se obtienen no alcanzan los requerimientos de valoración del estado de la carretera deseados, ya que estos resultados no son ni fiables ni repetibles. Más aún, la velocidad a la que los usuarios inspeccionan la superficie de la carretera no supera los 5km/h.

El catálogo de deterioros de firmes en carreteras, [MOP89], sigue siendo la principal fuente de referencia para la clasificación de defectos superficiales en España. En él se describen los principales defectos que pueden aparecer en pavimentos bituminosos, entre los que destacan las grietas y fisuras.

La inspección visual de la carretera presenta dos grandes retos. Por un lado se debe recoger la información del estado de la carretera sin interrumpir el tráfico y sin poner en peligro a los operarios. Por otro lado, se debe disponer de un registro electrónico de los defectos superficiales evaluados de forma objetiva. Diversos equipos de adquisición de imágenes están irrumpiendo en el mercado en los últimos años, destacando entre

todos ellos el sistema de adquisición de imágenes LRIS ("*Laser Road Imaging System*") de INO, [INO15]. A pesar de ello el procesamiento automático de las imágenes aún no alcanza los niveles de satisfacción deseados, debido, fundamentalmente, a la gran variabilidad de la apariencia visual de la superficie de la carretera y de los defectos que sobre ella pueden aparecer, que se ve afectada tanto por la mezcla bituminosa de la que está compuesta como por el desgaste que sufre por la circulación de vehículos. Además, la superficie de un pavimento bituminoso se encuentra altamente texturizada, sobre su superficie se encuentran las marcas viales, manchas y otros objetos, y se encuentra expuesta a la luz directa del sol, provocando sombras y reflejos no deseados.

Por otro lado, los túneles son obras subterráneas lineales utilizadas principalmente en las infraestructuras para el transporte de personas y mercancías. Este tipo de infraestructura es utilizado por las redes de carreteras, ferrocarriles y metros. Solo en España la Red Nacional de Carreteras cuenta con más de 200km de túneles, los túneles ferroviarios superan los 800km de longitud total y la de red de metro alcanza los 475km. A nivel mundial existen 68 túneles carreteros o ferroviarios de más de 10km de longitud.

La correcta conservación de túneles implica prever actuaciones anticipándose a la aparición de defectos muy graves o irreparables. Los métodos tradicionales de inspección de túneles están basados en los criterios subjetivos del observador (mediante una inspección visual se va rellenando una ficha de inspección) y en equipos que, normalmente, necesitan instalación y contacto con la superficie del túnel (testigos de yeso, fisurómetros, etc.).

La inspección visual del revestimiento de túneles de hormigón presenta también grandes retos que resolver. Por un lado, el tráfico debe ser interrumpido durante la inspección. En túneles de carretera las inspecciones se realizan durante la noche, desviando el tráfico por carreteras secundarias o inhabilitando ciertos carriles, lo que repercute sobre el usuario de la vía aumentando el tiempo empleado en realizar el recorrido y, lo que es más grave, aumentando el riesgo de accidente. En los túneles ferroviarios las inspecciones se realizan cuando la vía no está ocupada, normalmente también por la noche. Sin embargo, la tendencia actual es disminuir el tiempo en el que la vía no está ocupada para aumentar la rentabilidad de la infraestructura. Por otro lado, los criterios de seguridad son cada vez más exigentes, lo que, unido a los efectos de la fatiga y la subjetividad humana, hacen muy difícil que los resultados de la inspección tengan el grado de precisión requerido.

Como se ha puesto de manifiesto, la inspección visual de superficies continuas homogéneas, además de ser una tarea repetitiva y monótona, debe realizarse a gran

velocidad para no limitar la cadencia de producción en las aplicaciones industriales y para reducir los efectos sobre la circulación de vehículos y la seguridad de usuarios y operarios en infraestructuras lineales. Los resultados de la inspección visual llevada a cabo por inspectores humanos, debido a la fatiga que causa este tipo de tareas, empeoran significativamente con el paso del tiempo y, en raras ocasiones, alcanzan las velocidades de inspección deseadas. Además, la inspección humana adolece de objetividad, lo que lleva a que los resultados de la inspección dependan del operario que realiza la inspección, o, incluso, del momento en el que se realiza la misma.

Para poder abordar la inspección visual del 100% de la superficie de forma objetiva y a gran velocidad es necesario introducir **sistemas de inspección visual automático**. La introducción en el mercado de estos sistemas es una tarea que se está llevando a cabo con lentitud debido a que no siempre alcanzan el grado de fiabilidad esperado. Para conseguirlo, se deben mejorar tanto los procesos de adquisición de imágenes como las técnicas de procesamiento de las mismas. La adquisición de imágenes en superficies continuas tiene una serie de peculiaridades comunes, si bien, cada problema concreto requiere ajustar el binomio cámara-iluminación para realzar, en la medida de lo posible, los defectos que se pretenden detectar frente a la superficie libre de ellos. El procesamiento de las imágenes, cuyo objetivo es la detección y clasificación de las anomalías que aparecen en la superficie, también debe abordarse desde el conocimiento de la superficie y el tipo de anomalías que sobre ella pueden presentarse.

Las técnicas de procesamiento de imágenes para la detección y clasificación de defectos en superficies continuas son cada vez numerosas. Las técnicas de detección de defectos suelen basarse en la segmentación de la imagen original, bien utilizando un umbral fijo como en [Zha09], un umbral que depende de la evolución de imágenes anteriores como en [Yaz08] y en [Yan10], o utilizando un umbral distinto para cada columna de la imagen como en [Miy07]. Sin embargo, otros autores segmentan la imagen modificada mediante algún filtro, como en [Zha10], [Yaz08] y [Oli09a] que utilizan la entropía, o en [Mar10] que utilizan la transformada de Hough, o filtros morfológicos como en [Zha14], así como distintos filtros de detección de bordes como en [Mah91], [Riv15] y [Fuj11].

Para la clasificación de defectos se utilizan sistemas inteligentes entre los que destacan las redes neuronales, [Yaz08], [Pen09], [Mas12], el algoritmo de vecinos cercanos, [Dup97], y las máquinas de soporte vectorial, [Suv12], [Liu02]. Otros autores combinan distintos clasificadores, como en [Yun11], donde se combinan los tres clasificadores anteriores mediante el algoritmo de AdaBoost.

Tanto la detección como la clasificación requieren la evaluación de características que distingan unos defectos de otros o la superficie con defecto de la que no lo tiene.

Sin embargo, muchos de los defectos analizados en esta tesis no se pueden identificar con el grado de fiabilidad deseable utilizando las técnicas propuestas en la literatura, que se basan fundamentalmente en la evaluación de la distribución del nivel gris de la imagen. El análisis de la información espacio-frecuencial localizada de la imagen puede aportar información adicional que permita mejorar los resultados de la inspección visual. Dentro de estas técnicas se encuentran los filtros de Gabor, que permiten analizar las imágenes filtradas de forma similar a como lo hace el cerebro de los mamíferos, [Hub65], [Val82].

Se trata de un campo de investigación que no ha sido explotado en profundidad debido a su gran versatilidad y a que se requieren grandes tiempos de procesamiento que hacían inviable su implantación hasta hace pocos años. Estos filtros tienen la capacidad de poder ser utilizados para detectar defectos que ocasionan un cambio en el patrón normal de la imagen, para resaltar ciertas características de las imágenes, para clasificar diferentes tipos de texturas, o para extraer características de las imágenes filtradas para usarlas como entradas a un clasificador. Sin embargo, no existe una metodología clara para su utilización, que debe ser adaptada a cada aplicación.

1.2 Objetivos y alcance de la tesis

Esta tesis tiene por objetivo abordar el proceso completo de la inspección de superficies continuas, analizando el proceso de adquisición de la imagen, su procesamiento y la utilización de los resultados obtenidos.

La inspección de superficies continuas tiene una serie de características comunes a todas ellas y otras que son específicas de cada sistema de inspección. Es por ello que el objetivo final de esta tesis es aplicar y comprobar el funcionamiento de las metodologías y técnicas desarrolladas a tres problemáticas concretas:

- La inspección de bobinas de chapa de acero laminado
- La inspección de la superficie de carreteras de asfalto
- La inspección del recubrimiento de túneles de hormigón

Para lograr este objetivo general se plantean los siguientes objetivos específicos:

- Analizar la problemática de la inspección visual de superficies continuas, profundizando en los tres tipos de superficie objetivo de esta tesis. La investigación comienza analizando las necesidades de la inspección y describiendo los sistemas de adquisición de imágenes más apropiados y su

configuración. Finalmente se analizan en profundidad las técnicas de procesamiento de imágenes utilizadas en aplicaciones similares.

- Describir una arquitectura para la inspección de superficies continuas, estableciendo los elementos comunes a todas ellas.
- Diseñar el proceso de adquisición de datos de las tres aplicaciones propuestas, definiendo el tipo de cámara, el sistema de iluminación y las estrategias para manejar la considerable cantidad de datos generados.
- Establecer la forma de evaluar los resultados del procesamiento de imágenes, lo que permite comparar distintas técnicas de forma objetiva.
- Analizar las técnicas de detección y clasificación de defectos superficiales, así como su aplicabilidad a los tipos de superficie objetivo de esta tesis.
- Extraer características espacio-frecuenciales de las imágenes que proporcionen información adicional a las técnicas de detección y clasificación de defectos. Para ello se propone la utilización de filtros de Gabor que permiten el análisis de anomalías locales en el dominio espacio-frecuencial.
- Establecer una metodología para el diseño de bancos de filtros de Gabor que permitan analizar de forma eficiente una zona del espacio de frecuencias y orientaciones.
- Desarrollo de nuevas aplicaciones para el análisis de las imágenes filtradas por filtros de Gabor, explorando nuevas posibilidades de utilización.
- Utilizar las características de las imágenes filtradas por este tipo de filtros para la clasificación de defectos.
- Implementar los sistemas de inspección visual descritos en entornos reales.
- Analizar los resultados del procesamiento de imágenes comparándolos con trabajos similares siempre que sea posible.

1.3 Principales aportaciones

Las principales aportaciones realizadas en esta tesis hacen referencia tanto a la arquitectura de inspección de superficies continuas como a las técnicas de procesamiento de imágenes implementadas y se resumen en la siguiente lista:

- Definición de la **arquitectura** para la inspección de superficies continuas, describiendo las peculiaridades de los tres sistemas de inspección implementados
- **Implementación** de tres sistemas de inspección visual automático en entornos reales.
- Desarrollo de herramientas para la **visualización** de los datos adquiridos y los resultados del procesamiento, así como la generación de informes de resultados y toma de decisiones.
- Descripción de una metodología para el **diseño de bancos de filtros de Gabor**

- Adaptación del algoritmo de combinación de clasificadores binarios **AdaBoost** para la selección y análisis de imágenes filtradas por filtros de Gabor.
- **Extracción de características** de imágenes filtradas por filtros de Gabor para su utilización en clasificadores de tipos de defectos.
- Desarrollo del **filtro de Gabor invariante a la rotación** para segmentar defectos superficiales sin conocimiento previo de su orientación.
- Aplicación de una **nueva función de probabilidad basada en la función sigmoidea** al método de la textura de anisotropía condicional.
- **Combinación** del filtro de Gabor invariante a la rotación con la textura de anisotropía condicional.

1.4 Producción científica e industrial

Los resultados de la investigación realizada en esta tesis han sido publicados en revistas científicas de relevancia internacional y se han dado a conocer en diversos congresos. Por otro lado, han sido integrados en tres sistemas de inspección visual desarrollados en la Fundación CARTIF durante la ejecución de diversos proyectos de investigación.

1.4.1 Publicaciones

Durante el desarrollo de la presente tesis, fruto de un prolongado esfuerzo de investigación, se han llevado a cabo una serie de publicaciones en revistas científicas y congresos.

Se han realizado tres publicaciones en revistas científicas que se enumeran a continuación:

- Road Crack Detection Using Visual Features Extracted by Gabor Filters. *Computer-Aided Civil and Infrastructure Engineering*, 29(5): 342–358, 2014, [Zal14]

Impact Factor: 4,925, Q1 en las categorías *Computer Science, Interdisciplinary Applications* (puesto 1 de 102), *Construction & Building Technology* (puesto 1 de 59), *Engineering Civil* (puesto 1 de 124) y *Transportation Science & Technology* (puesto 1 de 32).

Artículo en el que se detallan los algoritmos utilizados en esta tesis para la detección de grietas en la superficie de la carretera combinando características extraídas por filtros de Gabor mediante el algoritmo AdaBoost.

- An Effective Texture Mapping Approach for 3D Models Obtained from Laser Scanner Data to Building Documentation. *Computer-Aided Civil and Infrastructure Engineering*, 26(5): 381-392, 2011 [Zal11]

Impact Factor: 3,382, Q1 en las categorías *Computer Science, Interdisciplinary Applications* (puesto 9 de 99), *Construction & Building Technology* (puesto 1 de 56), *Engineering Civil* (puesto 3 de 118) y *Transportation Science & Technology* (puesto 2 de 28).

Artículo en el que se propone una metodología para el mapeado automático y eficiente de texturas capturadas con cámaras de color sobre superficies tridimensionales obtenidas mediante escáneres láser.

- Automated Visual Classification of Frequent Defects in Flat Steel Coils. *The International Journal of Advanced Manufacturing Technology*, 57, 1087-97, 2011, [Med11]

Impact Factor: 1,103, Q2 en las categorías *Automation & Control Systems* (puesto 25 de 58) y *Engineering, Manufacturing* (puesto 12 de 37).

Artículo en el que se presentan los principales resultados obtenidos en la clasificación de defectos superficiales en la superficie de chapa laminada mediante la combinación de tres clasificadores que utilizan descriptores de distribución del nivel de gris, forma y características extraídas por filtros de Gabor.

Por otro lado se han realizado 7 contribuciones en congresos y jornadas de reconocido prestigio nacional e internacional:

- Fusión de Información Visual y Dimensional para la Detección de Grietas y Fisuras en la Superficie de Carretera. *X Jornada Nacional de ASEFMA, Zaragoza, España, Jul 2015, Comunicación 32*, [Gar15]

Jornadas en la que se describen los últimos avances en la detección y clasificación de grietas en carreteras y la utilización de los datos obtenidos por el gestor de la carretera para evaluar el deterioro de la carretera. El artículo fue seleccionado entre los 10 mejores del congreso.

- Enhanced Automatic Detection of Road Surface Cracks by Combining 2D/3D Image Processing Techniques. *IEEE International Conference on Image Processing, ICIP2014, CNIT La Défense, Paris, France, October 27-30, 2014, 778-782*, [Med14]

Conferencia en la que se han presentado los últimos avances en la detección de grietas en la superficie de la carretera combinando métodos de detección 2D/3D. El artículo fue seleccionado en el "Top 10%" del ICIP 2014.

- Detección Rápida de Regiones Uniformes en Chapa Laminada para Control de Calidad, *XVIII Congreso de Máquinas-Herramienta y Tecnologías de Fabricación, San Sebastian, España, Nov. 2010*. [Olm10]
Congreso en el que se han presentado algunos de los avances realizados en la detección de defectos en chapa laminada.
- Automated Visual Inspection of Road Surface Cracks. *Proceedings of the 27th International Symposium on Automation and Robotics in Construction, ISARC'10, Bratislava, Slovakia, Jun 25-27, 2010, 155-64, ISBN:978-80-7399-974-2*, [Med10]
Conferencia en la que se describen los primeros avances realizados en la detección de defectos superficiales en superficie de carreteras mediante la utilización de filtros de Gabor.
- Detección Automática en Línea de Defectos Superficiales en Chapa Laminada Lubricada mediante Técnicas de Visión Artificial, *XVII Congreso de Máquinas-Herramienta y Tecnologías de Fabricación, San Sebastián, España, Nov. 2008*. [Olm08]
Congreso en el que se han publicado algunos de los avances realizados en la detección de defectos en chapa laminada lubricada.
- Surface Defects Detection on Rolled Steel Strips by Gabor Filters. *VISAPP 2008: Proceedings of the Third International Conference on Computer Vision Theory and Applications, Funchal, Madeira, Portugal, January 22-25, 2008 – Volume 1; 479-485, ISBN:978-989-8111-21-0*, [Med08]
Conferencia en la que se presentan los primeros avances realizados en la detección de defectos superficiales en chapa laminada mediante filtros de Gabor.
- Inspección visual en línea de defectos en piezas de chapa embutida. *XVI Congreso de Máquinas-Herramienta y Tecnologías de Fabricación, San Sebastián, España, Oct. 2006*. [Gay06b]
Congreso en el que se presentan técnicas de visión artificial para el procesamiento de imágenes en chapa embutida.
- Detección Visual en Línea de Roturas y Poros en Piezas de Chapa Embutida. *XXVII Jornadas de Automática, Almería, España, Sep. 2006*. [Gay06a]
Jornadas en las que se presentan distintas técnicas de visión artificial para la detección de defectos en chapa embutida.

1.4.2 Producción industrial

Los trabajos de investigación descritos en esta tesis se enmarcan dentro de las líneas de investigación de la División de Robótica y Visión Artificial de la Fundación CARTIF y han sido integrados en los siguientes sistemas automáticos de inspección:

- **Sistema de inspección de chapa de acero laminado:** Este sistema se encuentra integrado en un entorno industrial real, concretamente en dos líneas de corte de chapa laminada en las instalaciones del Grupo GONVARRI en Burgos. El sistema de inspección adquiere y procesa la imagen de las dos caras de la chapa laminada en tiempo real a una velocidad de hasta 200m/min, almacenando la información y produciendo alarmas en caso de detectar defectos graves. El desarrollo y la mejora del sistema de inspección se ha abordado durante la ejecución de los siguientes proyectos de investigación:
 - Monitorización y diagnóstico de líneas de corte de chapa y detección de defectos en chapa laminada mediante visión industrial (ADE, 04/04/BU/0025)
 - Detección defectos en chapa laminada mediante visión industrial (CDTI, FIT-350300-2004-22)
 - Acevis: Combinación de técnicas espectroscópicas y algorítmicas para la detección de defectos superficiales (Ministerio de Ciencia e Innovación , PPT-020000-2008-6)
 - Tratamientos inteligentes en imágenes digitales de bajo contraste de superficies aceitadas con elevada reflectancia especular (Ministerio de Ciencia e Innovación, PID-560410-2009-12)
- **Vehículo de inspección de superficies de carreteras:** Desarrollado en colaboración con Corporación Llorente S.A., permite la toma de datos del estado superficial de la carretera a la velocidad de la vía para su posterior análisis en el laboratorio. El resultado del análisis permite la toma de decisiones para el mantenimiento y la rehabilitación de las carreteras inspeccionadas. Su desarrollo se ha llevado a cabo durante la ejecución de los siguientes proyectos de investigación:
 - ACVA: Auscultación de carreteras mediante visión artificial y tratamiento de imágenes (ADE)
 - MULTIVIAL: Fusión multisensorial dinámica de equipos para la inspección de viales: actuaciones preventivas para la seguridad y la conservación (CDTI, IDI- 20110467)
- **Prototipo para la inspección de fisuras y grietas en hormigón:** Desarrollado en colaboración con Dragados S.A. permite la adquisición de imágenes en túneles ferroviarios y carreteros. El alcance de esta investigación se limita a la adquisición de una pequeña porción del túnel

para su posterior análisis en el laboratorio. Ha sido desarrollado en el marco del siguiente proyecto de investigación:

- INFIT: Inspección y detección de fisuras en túneles ferroviarios (INNPACTO 2012, IPT-2012-0980-370000)

1.5 Organización de la tesis

La presente tesis está organizada en 8 capítulos o secciones. En el **Capítulo 1** se presenta el marco de la investigación, los objetivos, el alcance del trabajo y las principales aportaciones de esta tesis. También se describe la organización del trabajo y las principales publicaciones en revistas científicas y congresos, que han sido fruto de la investigación realizada.

En el **Capítulo 2** se describe la problemática de la inspección visual automática de superficies continuas, para a continuación describir con más detalle tres tipos concretos de inspección en los que se centra esta tesis: la inspección de chapa laminada, de superficies de carretera y del revestimiento de túneles de hormigón. Para cada tipo de superficie se analizan los principales retos que deben ser abordados, tanto desde el punto de vista de la adquisición de imágenes como de su procesamiento, describiendo el estado actual de la tecnología y los principales trabajos de investigación relacionados con la detección y clasificación de los defectos que pueden aparecer sobre estas superficies.

En el **Capítulo 3** se analizan las principales características de la arquitectura necesaria para la inspección de superficies continuas. El proceso de inspección ha sido dividido en tres grandes etapas. En la primera etapa se aborda la adquisición de imágenes. Por ser superficies continuas la captura de imágenes requiere un sistema de adquisición de imágenes continuas, para lo cual es preferible la utilización de cámaras lineales cuyos requisitos son tenidos en cuenta en esta sección, así como la problemática de la iluminación y el almacenamiento. En la segunda etapa se realiza el procesamiento de las imágenes adquiridas. El objetivo del procesamiento es la detección y la clasificación de defectos superficiales. En la última etapa se utilizan los resultados obtenidos, bien mediante la visualización de los resultados o para la toma de decisiones.

En el **Capítulo 4** se resumen las principales técnicas de detección y clasificación de defectos que han sido utilizadas y adaptadas en esta tesis para la inspección de los tres tipos de superficies presentados en Capítulo 2. En primer lugar, se describe como evaluar el desempeño de las técnicas empleadas. A continuación, se analizan algunas técnicas para el cálculo de umbrales en clasificadores binarios y cómo combinarlos

mediante el algoritmo de AdaBoost. Finalmente se presentan tres tipos de clasificadores de tipos de defectos y distintos métodos para su combinación y mejora de resultados.

En el **Capítulo 5** se encuentran las principales aportaciones al procesamiento de imágenes de esta tesis. Las técnicas de procesamiento de imágenes presentadas en el Capítulo 4 requieren la evaluación de una o varias características de la imagen. En este capítulo se describe como extraer información espacio-frecuencial de las imágenes mediante la utilización de filtros de Gabor. Se propone una metodología para el diseño de bancos de filtros de Gabor y distintas técnicas para el análisis de las imágenes filtradas. También se propone una modificación al filtro de Gabor, que se ha denominado filtro de Gabor invariante a la rotación, para la detección de defectos cuya orientación no es conocida a priori. Finalmente se analiza el concepto de textura de anisotropía condicional, proponiendo mejoras en el algoritmo original y combinándolo con el filtro de Gabor invariante a la rotación.

En el **Capítulo 6** se describen los sistemas de inspección implementados en esta tesis. Se han desarrollado tres sistemas. El primero es una aplicación industrial para la detección y clasificación de defectos en chapa laminada en tiempo real que se encuentra instalado en una línea de corte de chapa laminada. El segundo es un sistema de inspección de carreteras embarcado en un vehículo que realiza la toma de datos a la velocidad de circulación de la vía. El tercero de los sistemas implementados es un prototipo para la inspección de grietas y fisuras en túneles. Para cada sistema se detallan los principales retos que se han tenido que abordar para la adquisición y el tratamiento de las imágenes, así como la utilización de los resultados obtenidos.

En el **Capítulo 7** se analizan los principales resultados del procesamiento de imágenes de los tres sistemas de inspección implementados. En este capítulo se describen los objetivos de cada sistema de inspección y se detallan las principales técnicas de procesamiento de imágenes utilizadas en cada caso. Se muestran los principales resultados experimentales obtenidos comparándolos, cuando es posible, con trabajos similares y realizando un análisis crítico de los mismos.

En el **Capítulo 8** se resumen los principales logros alcanzados en esta tesis y se proponen futuras líneas de investigación.

2 INSPECCIÓN VISUAL AUTOMÁTICA DE SUPERFICIES CONTINUAS

2.1 Descripción de la problemática

La inspección es el proceso para determinar si un producto se desvía de unas especificaciones dadas. La inspección visual es uno de los principales procesos de inspección que se encarga de determinar si las características visuales de un producto se desvían de las especificaciones establecidas.

La inspección visual superficial es el proceso de detección, análisis y clasificación de estructuras anómalas en la superficie inspeccionada. Este tipo de inspección se realiza en muchos tipos de productos industriales, como telas, papel, plásticos, papel de aluminio, parqués, metales, madera, etc., pero también puede aplicarse a otros tipos de superficies, como las infraestructuras lineales (carreteras, túneles, vías, etc.)

La principal característica que tienen en común las superficies enumeradas anteriormente es que son homogéneas en toda su superficie, entendiendo esta homogeneidad como que observando cualquier pequeña porción de la superficie no se puede conocer en qué zona de la superficie nos encontramos. Esta homogeneidad en muchas ocasiones es una textura que puede consistir en un patrón que se repite (como el de una tela) o no seguir ningún patrón (como en la madera). La inspección visual de estas superficies tiene como primer objetivo determinar qué zonas de la superficie son defectuosas por alejarse de la homogeneidad esperada. En muchas ocasiones también se busca determinar qué tipo de defectos están presentes en la superficie para poder llevar a cabo medidas correctoras en el proceso o clasificar la calidad del producto.

Otra de las características que tienen en común estas superficies es que tienen un ancho acotado y una longitud indeterminada que es mucho mayor que el ancho, por lo que reciben el nombre de superficies continuas.

La inspección visual de estos productos es una tarea repetitiva y monótona, que cuando es llevada a cabo por personas provoca una fatiga visual en los operarios que hace difícil conseguir unos niveles de fiabilidad adecuados. La máxima velocidad a la que los operarios pueden inspeccionar una superficie continua es bastante inferior a la velocidad a la que se procesa en las líneas de producción industrial, o la velocidad normal de circulación en carreteras o ferrocarriles. Además, aunque la inspección

visual cuando es realizada por operarios se adapta muy bien a nuevas situaciones, tiene un alto grado de subjetividad, lo que provoca que ante una misma situación dos operarios distintos, o el mismo operario en distintas ocasiones, ofrezcan resultados diferentes.

Es por ello por lo que se hace necesario el desarrollo de sistemas de inspección visual automáticos que permitan inspeccionar superficies a altas velocidades y con objetividad. El desarrollo de estos sistemas implica abordar grandes retos tanto en la fase de adquisición de imágenes como en el procesamiento de las mismas. La adquisición de las imágenes debe ser llevada a cabo a gran velocidad y, en muchas ocasiones, en ambientes hostiles. Tanto la adquisición como el procesamiento de las imágenes deben manejar grandes volúmenes de datos que deben ser tratados a gran velocidad. La apariencia visual, tanto de la superficie sin defectos como de los propios defectos, presenta una gran variabilidad pero una baja separabilidad. Es decir, las características de un mismo tipo de superficie o tipo defecto puede tener un amplio rango de valores, pero la diferenciación entre la superficie sin defecto y los defectos, o los defectos entre sí, es difícil de apreciar.

Aunque no es posible desarrollar un sistema de inspección visual automático que pueda abordar la detección y clasificación de todos los defectos que pueden aparecer sobre estas superficies continuas, sí que tienen algunas características comunes y muchas de las técnicas pueden extrapolarse o adaptarse de una aplicación a otra. En esta tesis se aborda la inspección de tres tipos de superficie que se pueden ver en la **Figura 1**: bobinas de acero laminado, pavimento de carreteras y revestimiento de túneles de hormigón.

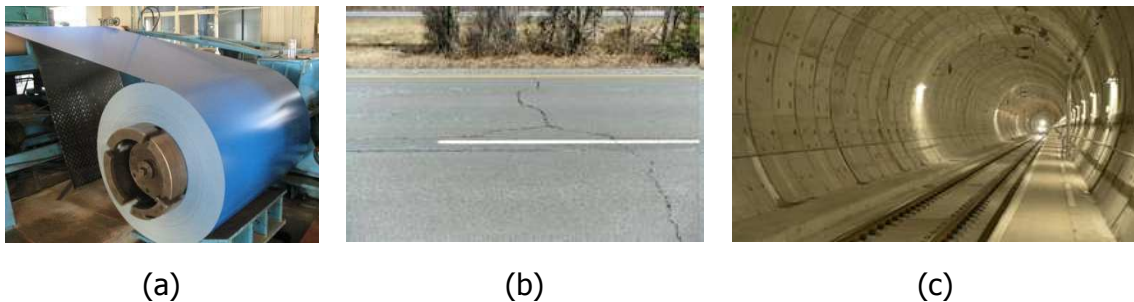


Figura 1. Superficies continuas objeto de estudio en esta tesis. (a) Bobina de acero laminado; (b) Porción de carretera en la que se aprecia una grieta en carretera; (c) Túnel de hormigón

En este capítulo se va a abordar la problemática específica de cada tipo de superficie, analizando el proceso completo de inspección visual automática, desde la adquisición de las imágenes a la obtención y presentación de los resultados, prestando especial interés a las técnicas de procesamiento de imágenes para la detección y

clasificación de algunos de los defectos más comunes que aparecen sobre estas superficies.

2.2 Características comunes para los sistemas de inspección visual de superficies continuas

Aunque la inspección visual de superficies continuas debe abordar aspectos concretos para cada aplicación, existen algunas características comunes que se van a analizar en esta sección, para en posteriores secciones de este mismo capítulo detallar las características específicas de tres soluciones implementadas en esta tesis.

El primer elemento que tienen en común estos sistemas es el objetivo final de la inspección. En un sistema de inspección visual de superficies continuas persigue un doble objetivo: la detección de anomalías en la superficie y su clasificación. Una superficie continua se caracteriza por ser homogénea en toda su superficie, por tanto, la detección de anomalías en la superficie busca medir la calidad de producto o su deterioro y localizar la posición de las anomalías detectadas. La clasificación de estas anomalías, también llamadas defectos o deterioros, permite determinar cuáles son y establecer las actuaciones a realizar. El tipo y las características de estos deterioros son específicos de cada aplicación, y determinan tanto la forma de adquirir las imágenes como los algoritmos de procesamiento. También son específicas de cada aplicación las acciones a realizar con la información extraída, aunque se pueden dividir en dos grandes grupos, las que actúan sobre el proceso de fabricación del producto o las que actúan directamente sobre el producto.

El segundo elemento común que tienen estos sistemas es la arquitectura de adquisición de datos, aspecto que se tratará con mayor profundidad en la sección 3. La adquisición de la imagen de una superficie continua se lleva a cabo de forma más eficiente mediante la utilización de cámaras lineales. La inspección mediante cámaras lineales se denomina inspección en continuo (*web inspection*) y suele realizarse en materiales como la tela, papel, vidrio, bobinas de metal, y, en general, cualquier producto que tenga una longitud indeterminada y un ancho acotado. Por tanto, las cámaras lineales capturan una imagen de una anchura conocida y una longitud ilimitada.

Debido a la alta velocidad de adquisición que suelen requerir estos sistemas, los tiempos de exposición de las cámaras deben ser muy bajos, por lo que se requieren cámaras de alta sensibilidad y fuentes de iluminación de alta potencia. La iluminación debe ser lo más homogénea posible para simplificar el procesamiento de las imágenes, lo cual se consigue de forma más sencilla con iluminación lineal. Otro aspecto a tener

en cuenta al utilizar tiempos de exposición bajos, es la necesidad de utilizar fuentes de iluminación estables que no presenten rizado, ya que en caso contrario cada línea de la imagen recibiría una intensidad lumínica diferente dificultando enormemente su procesamiento posterior. Además, todos estos sistemas tienen otro elemento hardware común, un codificador incremental (*encoder*), que es el encargado de enviar la señal de disparo o *trigger* para la adquisición de cada línea de la imagen.

Aunque cada sistema de inspección puede tener ciertas peculiaridades que se abordarán en esta tesis, todos ellos tienen los siguientes elementos comunes:

- Una o varias cámaras, habitualmente cámaras lineales, con sus correspondientes ópticas para la adquisición de imágenes a la resolución requerida
- Un sistema de iluminación cuya distribución sea homogénea a lo largo de la línea de la imagen
- Un codificador incremental que sincronice el movimiento de la superficie a inspeccionar, o el de la plataforma de inspección, con la adquisición de las imágenes
- Equipos informáticos para el control de la adquisición, el almacenamiento de la información y el procesamiento de las imágenes
- Un interfaz gráfico que permita la gestión y el control de todo el proceso de adquisición y procesamiento de las imágenes
- Un interfaz gráfico que permita la visualización de los resultados obtenidos

El procesamiento de las imágenes también consta de tres etapas comunes. Un preprocesamiento, una segmentación y la clasificación:

- La etapa de preprocesamiento tiene por objetivo preparar la imagen para su posterior procesamiento. En esta etapa se aplican filtros para eliminar el ruido de la imagen y, fundamentalmente, se corrigen los efectos de la iluminación.
- La etapa de segmentación busca la detección y agrupación de los píxeles que contienen defectos en la imagen. El objetivo es determinar cuáles son los píxeles de la imagen que se alejan del patrón normal de la superficie.
- La etapa de clasificación establece a qué grupo de defectos pertenece cada uno de los grupos de píxeles detectados en la etapa anterior.

Cuando el nivel de gris de la imagen sin defectos es estable y los defectos presentan un alto contraste respecto al nivel de gris de la imagen sin defectos, llevar a cabo la segmentación es un proceso sencillo que se realiza mediante una umbralización del nivel de gris de la imagen, es decir, determinando cuál es el rango de valores de nivel de gris que corresponde a la superficie libre de defectos y cuál corresponde a los defectos. Sin embargo, en muchas ocasiones esto no ocurre, siendo

el caso más habitual justo el contrario, es decir, que los niveles de gris de la imagen sin defectos presente una gran variabilidad, y que el contraste respecto a las zonas con deterioros sea muy pequeño o incluso nulo. Evidentemente, la aplicación de métodos de umbralización en estos casos no permite la correcta segmentación de la imagen, por lo que será preciso aplicar otras técnicas o estrategias de procesamiento de imágenes. El objetivo de este proyecto es la búsqueda de estas técnicas, que serán propias de cada problema específico y serán analizadas más adelante.

Aunque algunos sistemas de inspección visual de superficies continuas solo buscan determinar cuál es la superficie que contiene defectos y cuál está libre de defectos, en un gran número de ocasiones conocer cuál es el tipo de defecto detectado es de vital importancia para poder tomar las acciones apropiadas y/o poder determinar las causas que lo provocan. Esta clasificación suele realizarse agrupando los píxeles que pertenecen a las superficies etiquetadas como defectuosas y, extrayendo de ellos un vector de características que los describan, distinguir unos tipos de defectos de otros. La elección de las características más apropiadas para la correcta clasificación de los defectos es un aspecto crítico y abierto aún a la investigación. En otras ocasiones la clasificación viene determinada a partir de la propia detección, desarrollándose tantos algoritmos de detección como tipos de defectos sea preciso detectar.

A continuación se van a analizar tres tipos concretos de inspección de superficies continuas. Por un lado se va a abordar un problema de inspección industrial, la inspección de chapa laminada, mientras que por otro lado se va a abordar la inspección de dos tipos de infraestructuras lineales, la superficie de la carretera y el revestimiento de túneles de hormigón. La inspección de estos tres tipos de superficies van a ser el banco de pruebas sobre el que se van a aplicar los resultados de investigación de esta tesis. En este capítulo se describen los principales defectos que aparecen sobre estas superficies, las dificultades que surgen para su inspección y cómo se ha abordado la adquisición y el procesamiento de imágenes en trabajos anteriores.

2.3 Inspección de bobinas de chapa laminada

El acero es el material más utilizado en la ingeniería y en la construcción tanto por la cantidad empleada como por la variedad de usos que se le puede dar, representando el 95% de la producción total, [Chr08]. El acero laminado es el más utilizado de todos ellos, siendo empleado en multitud de productos, destacando la industria del automóvil y la de electrodomésticos. Según la asociación mundial del acero (*World Steel Association*) la producción bruta de acero en el año 2013 fue de

1.582 millones de toneladas, [Wor14], más que la producción total de todos los demás metales juntos.

El acero laminado es habitualmente suministrado en bobinas de varios kilómetros de longitud y un ancho de aproximadamente dos metros. Estas bobinas son procesadas en líneas de corte de alta velocidad por empresas dedicadas expresamente a esta tarea. Los defectos superficiales presentes sobre estas bobinas deben ser detectados y clasificados con precisión para que los responsables de calidad de la empresa puedan decidir cuando el producto es apropiado para un determinado uso o cliente.

La calidad superficial de estos productos tiene una gran importancia por lo que tradicionalmente es inspeccionada por operarios al final de las líneas de producción. La inspección de estos productos es habitualmente llevada a cabo por personal especializado pero no alcanza los requerimientos de fiabilidad, repetitividad y velocidad necesarios. Estas líneas de corte pueden llegar a alcanzar velocidades de hasta 200m/s, por lo que la inspección visual del 100% del producto por los operarios es una tarea imposible si se quiere alcanzar unos índices de productividad aceptables. En la práctica, se realiza la inspección visual de los primeros metros de la bobina a baja velocidad, obteniéndose una idea general del resto de la bobina. Si se aprecian ciertos defectos en el comienzo de la bobina, ésta puede llegar a ser inspeccionada en su totalidad a baja velocidad para obtener un conocimiento más exacto del estado superficial de la misma, con el consiguiente retraso en la producción de la línea.

Para poder inspeccionar la producción total, sin reducir la velocidad de procesamiento de la bobinas, sería necesario la implantación de sistemas de inspección automática que detectaran y clasificaran los defectos presentes en las bobinas y almacenaran esta información. Para conseguir este objetivo se propone la utilización de técnicas visión artificial. La visión artificial no consiste solamente en diseñar algoritmos de procesamiento de imágenes, sino que requiere la integración de cámaras, iluminación, equipos informáticos e interfaces hombre-máquina. Alcanzar el objetivo de desarrollar un sistema automático de detección y clasificación de defectos en la superficie de bobinas de acero laminado, supone un enorme reto debido a la enorme variedad de defectos que pueden aparecer sobre la superficie de la bobina, la naturaleza reflectante de la superficie del producto, la presencia de lubricantes que enmascaran la presencia de defectos y las altas velocidades a las que son procesadas las bobinas.

La documentación publicada sobre la inspección superficial de acero laminado fundamentalmente consiste en trabajos de investigación realizados en instituciones académicas, plantas de producción de acero laminado y fabricantes de equipos de inspección superficial. Un número destacable de trabajos ha sido publicado conjuntamente por instituciones académicas o centros de investigación junto a plantas

de producción, indicando que existe una buena colaboración entre las partes. Durante los últimos años un porcentaje significativo de los trabajos publicados provienen de China, que concuerda con el hecho de ser actualmente el mayor productor y demandante de acero.

Algunas publicaciones informan sobre el trabajo realizado en la detección y clasificación de defectos mediante la utilización de sistemas comerciales, pero no suelen ofrecer detalles sobre las técnicas utilizadas, probablemente por motivos de protección de propiedad industrial.

2.3.1 Complejidad de un sistema automático para la inspección de chapa laminada

Un sistema de inspección en tiempo real de la superficie de acero laminado tiene que enfrentarse a un gran número de retos. Los principales son:

- **Emplazamiento peligroso:** El lugar de la instalación de los equipos (el sistema de iluminación, las cámaras y otros equipos electrónicos auxiliares) es peligroso. La presencia de alta temperatura ambiente, polvo, aceite, agua y vapores es muy habitual. Además, el sistema de iluminación y las cámaras requieren la protección frente a golpes y vibraciones. Más aún, pesados equipos son movidos dentro y fuera de la instalación durante el proceso de trabajo y mantenimiento, tanto diario como periódico. Todos estos factores provocan la necesidad del uso de apropiadas medidas de protección tanto físicas y como ambientales de los equipos instalados.
- **Velocidad de funcionamiento:** Durante la producción normal, la velocidad a la que se procesan las bobinas de chapa laminada es generalmente alta. Estas velocidades pueden alcanzar los 200m/min. Por tanto, estas altas velocidades requieren equipos de adquisición y procesamiento de imágenes adecuados y *software* de procesamiento de imágenes que precisen pequeños tiempos de ejecución.
- **Gran variedad de defectos superficiales:** La variedad de defectos superficiales en chapa laminada es muy elevada, depende del tipo de acero utilizado y no existe ningún estándar de clasificación. Es más, sus características y la forma de clasificarlos varían de unas empresas a otras, e incluso de unos operarios a otros dentro de la misma empresa. Además, estos defectos pueden cambiar la forma de manifestarse cuando se producen cambios en el proceso de producción. En la **Figura 2** se encuentran algunos ejemplos de defectos que pueden aparecer sobre la superficie de una chapa laminada.
- **Necesidad de un gran número de cámaras:** Para la inspección de una bobina de acero laminado, dos conjuntos de sistemas de inspección son

necesarios, uno para la cara superior y otro para la cara inferior. Cada uno de estos sistemas requiere varias cámaras para capturar todo el ancho de la bobina, lo que hace que la captura de las imágenes, su sincronización y su procesamiento en tiempo real sea una tarea desafiante.

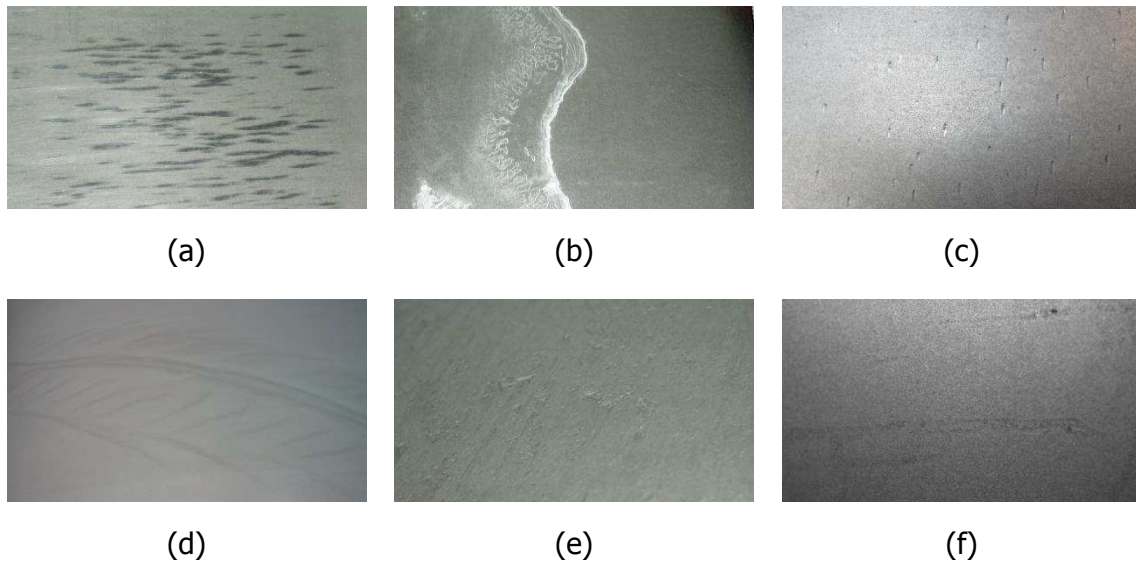


Figura 2. Imágenes de algunos de los defectos que pueden aparecer en la superficie de una bobina de chapa laminada. (a) Marcas de asentamiento; (b) Moho blanco; (c) Impresiones; (d) Helechos; (e) Viruela; (f) Grupos de cinc

2.3.2 Elementos clave de la estructura hardware de un sistema de inspección superficial de chapa laminada

Un sistema de inspección automática de defectos consiste en una o varias fuentes de iluminación, una o varias cámaras, un sistema de procesamiento de imágenes y un interfaz con el operario.

2.3.2.1 Sistema de adquisición de imágenes

Para obtener una imagen de calidad de la superficie es necesario iluminarla de forma adecuada y uniforme. En la práctica, cuanto mayor es la calidad de la iluminación menor es el tiempo y la complejidad del procesamiento de las imágenes. Se pueden encontrar diversas publicaciones sobre como iluminar superficies metálicas, [Mue00], [Seu01], [Per03], [Per04], y más concretamente sobre acero laminado, [Liu10].

Normalmente, en la iluminación de metales se utilizan dos tipos de iluminación conocidas como *campo brillante* y *campo oscuro*. En la iluminación de *campo brillante*, la iluminación y el sensor de la cámara están situados de tal manera que la cámara captura la mayor parte de la luz emitida al ser reflejada por la superficie del metal.

Algunos defectos provocan que la cantidad de luz reflejada disminuya significativamente, de tal manera que la superficie con defectos aparece más oscura. En contraposición, cuando se utiliza la iluminación de *campo oscuro*, el ángulo de incidencia de la luz respecto a la normal a la superficie es muy grande, y la cámara se sitúa justo en la normal a la superficie. Esto implica que la superficie del acero libre de defectos se ve oscura, pero algunos defectos provocan un brillo en la imagen. La iluminación de campo oscuro requiere una mayor intensidad en la iluminación, aproximadamente 8 veces más según [Mue00].

Desafortunadamente, no todos los defectos se pueden detectar únicamente utilizando solo la iluminación de *campo brillante* o solo la de *campo oscuro*. En muchos casos se utilizan dos grupos de sistemas de cámara-iluminación para cubrir ambos métodos de iluminación, [Sas07], [Wu08], [Zha10]. En [Wu08] se describe una instalación en una planta de acero de China donde se utilizan 20 cámaras matriciales para adquirir las dos caras de las bobinas mediante los dos métodos de iluminación descritos anteriormente. Sin embargo, teniendo en cuenta la complejidad del sistema y los costes de instalación y mantenimiento, la mayoría de los sistemas sitúan la iluminación y las cámaras en un lugar intermedio entre el *campo oscuro* y el *campo brillante*.

2.3.2.2 Fuente de iluminación

La fuente de iluminación debe cumplir dos características fundamentales:

- **Homogeneidad:** La intensidad de la iluminación debe ser lo más homogénea posible en toda la superficie iluminada. Conseguir la perfecta homogeneidad es en la práctica imposible debido a que la distancia entre la fuente de iluminación y los distintos puntos del objeto a iluminar no es la misma, por lo que siempre es necesario realizar una corrección posterior.
- **Sin rizado:** Se denomina rizado a la evolución temporal de la intensidad de la iluminación y aparece cuando la fuente de iluminación es alimentada con corriente alterna. También puede aparecer cuando la iluminación se alimenta con corriente continua producida por un convertidor de corriente continua a corriente alterna. Cuando la velocidad de adquisición de las cámaras es elevado es un factor crítico a tener en cuenta.

La longitud de onda de la fuente de iluminación no es un factor crítico a la hora de iluminar metales ya que estos se comportan de forma similar a distintas longitudes de ondas del espectro visible. Sin embargo, cuando el metal se encuentra lubricado, lo ideal es buscar longitudes de onda próximas al infrarrojo o en el infrarrojo cercano, ya que los aceites son más transparentes en esas longitudes de onda, [Med11].

2.3.2.3 Tipo de cámaras

Habitualmente se utilizan cámaras con sensor CCD de alta resolución. Aunque se pueden utilizar tanto cámaras matriciales como lineales, lo habitual es usar cámaras lineales ya que es más sencillo iluminar de forma homogénea una línea que un área y además esa línea se puede iluminar con una mayor intensidad, lo que permite alcanzar mayores velocidades y/o resoluciones de la imagen. Además, así se evitan los eventuales solapamientos y huecos entre imágenes que se producen con cámaras matriciales.

La resolución de las imágenes, en la mayoría de las ocasiones, es la misma en la dirección de avance de la bobina y en la dirección perpendicular al avance, como en [Yaz08], [Wu08] y [Cor10], con resoluciones de 0,37mm, 0,5mm y 1mm respectivamente, mientras que otros trabajos la resolución de avance es menor que la transversal, como en [Sas07], con una resolución transversal de 0,17mm mientras que en la dirección de avance es de 1,25mm.

2.3.2.4 Hardware de procesamiento de imágenes

Las imágenes capturadas por las cámaras son enviadas a un sistema de procesamiento de imágenes en paralelo de alta velocidad dedicado a cada cámara y localizado próximo a ella, [Wu08]. El procesamiento en paralelo permite que la inspección sea realizada en tiempo real mediante la detección de las zonas candidatas a contener defectos que deben ser procesadas a continuación. El sistema de procesamiento puede estar contenido en la propia cámara, ser un procesador dedicado tipo FPGA o un procesador de carácter general con un hardware especial para la adquisición de imágenes. Esta parte del sistema es de vital importancia tanto para conseguir que el procesamiento se realice en tiempo real como para que la detección y clasificación de defectos alcance niveles aceptables de precisión. Esta información es enviada a un nivel superior en el que se realiza un nuevo procesamiento para la clasificación de las zonas candidatas de ser defectos y se almacena la información. Finalmente, debe existir un tercer nivel hardware en el que el usuario pueda interactuar con el sistema de inspección, tanto para la visualización de los resultados como para la configuración del sistema.

2.3.3 Procesamiento de imágenes

El objetivo último de la inspección superficial es determinar cuáles son las zonas de la superficie del acero que tiene defectos y cuáles no, así como determinar a qué categoría pertenecen los defectos detectados. Normalmente el proceso comienza con la segmentación de la imagen, agrupando los píxeles en zonas de interés, para a continuación extraer un conjunto de características de cada zona para clasificar esas zonas por tipos de defectos. Los métodos de clasificación pueden ser supervisados o

no supervisados. Los métodos supervisados agrupan los defectos en categorías preestablecidas mientras que los no supervisados agrupan los defectos con características similares sin asignarlas a un grupo de defectos concreto. Sin embargo, los defectos presentan gran similitud entre distintos tipos de defectos y gran diversidad dentro del mismo tipo de defectos, lo que hace que la clasificación sea un área de investigación a la que se estén dedicando muchos esfuerzos.

2.3.3.1 Detección de defectos

La detección de defectos suele comenzar con la selección de las zonas de la imagen candidatas a ser defectos mediante algún método de segmentación. Suele ocurrir que algunas de las zonas seleccionadas pertenezcan a la superficie libre de defectos. A estas zonas se les conoce como pseudodefectos, [Zha10], [Cho14]. Cuando la cantidad de pseudodefectos es pequeña este procesamiento puede servir para determinar la calidad del producto procesado. Sin embargo, cuando la cantidad de pseudodefectos es elevada se debe realizar una clasificación de las zonas candidatas para determinar cuáles son realmente defectos y cuáles no. En algunos casos se realiza una clasificación específica para separar los candidatos entre defectos y no defectos, mientras que en otros casos se realiza la clasificación global de los defectos considerando que uno de los tipos es el grupo de pseudodefectos.

Para la detección de defectos se utilizan técnicas en el dominio espacial basadas en el análisis del nivel de gris de las imágenes, así como técnicas basadas en el análisis del dominio frecuencial, aunque estas últimas son más complicadas de implementar y su procesamiento es más lento, por lo que resulta difícil su utilización en aplicaciones de tiempo real.

En [Bla08] se detectan seis tipos de defectos en chapa laminada en frío. Para ello primero se elimina el ruido utilizando el filtro de Wiener, a continuación, la posición aproximada de los defectos es calculada utilizando el detector de bordes de Sobel, para finalmente determinar la localización exacta de los defectos mediante un método de crecimiento basado en lógica difusa.

Otros métodos se basan en buscar la variación del nivel de gris normal de la chapa, como en [Yaz08] y en [Yan10]. Estos métodos asumen que la superficie sin defectos tiene un nivel de gris homogéneo y que varía de forma lenta. La imagen estándar es generada a partir de una serie de imágenes consecutivas que cumplen con un criterio de variación de intensidad a nivel de pixel. Cuando una nueva imagen presenta una variación significativa respecto a la intensidad media de las imágenes anteriores, se considera que esas zonas de la imagen son candidatas a ser defectos. En [Con07], se busca la variación de la imagen dividiéndola en sub-imágenes de 20 x 20 píxeles y comparando la distribución de los niveles de gris con un patrón.

La transformada de Hough es utilizada cuando los defectos tienen una forma muy definida, [Sha08], [Mar10]. En [Mar10] se utiliza la transformada de Hough para detectar tres tipos de defectos, soldaduras oblicuas, soldaduras perpendiculares y el agujero utilizado para marcar las soldaduras. Después de homogeneizar la iluminación y filtrar la imagen para eliminar el ruido, se aplica un detector de bordes seguido de una segmentación. La transformada de Hough es utilizada para detectar líneas y círculos e identificar los citados tres tipos de defectos. Obtienen una exactitud del 98% con 300 muestras, aunque este tipo de defectos es relativamente sencillo de detectar.

La medida de la entropía es utilizada para la detección de defectos en [Yaz08]. El procedimiento utilizado comienza restando una imagen patrón de las nuevas imágenes. Para cada píxel se calcula la entropía en un entorno de 9×9 y se realiza la segmentación de esta imagen utilizando el método de Otsu. Finalmente se eliminan los píxeles no unidos mediante una dilatación seguida de una erosión de la imagen binaria. En [Zha10] se utiliza también la entropía para diferenciar las zonas de la imagen con defecto de las que no tienen defecto, definiendo lo que denominan exceso de entropía y segmentando la imagen utilizando los valores de exceso de entropía. Sin embargo, en este trabajo no se concreta para qué tipo de defectos es válido ni cuáles son los resultados obtenidos.

En [Wei08] se utilizan características sencillas de la imagen para compararlas con los valores normales encontrados en la chapa laminada no defectuosa. La imagen es dividida en sub-imágenes de 40×40 en las que se evalúa la diferencia entre el píxel con mayor intensidad y el de menor, la media y la varianza. Utilizan una función discriminante para separar los defectos de los no defectos, obteniendo una exactitud del 91%. En el trabajo no se indica a qué tipo de defectos se aplica el algoritmo propuesto.

En [Cor10] se busca diferenciar las zonas sin defecto de la chapa laminada de las zonas con defecto mediante el uso de descriptores morfológicos y *curvelets* con el fin de describir la textura de la imagen. En este trabajo se utilizan 30 descriptores morfológicos, la mitad son aperturas (una erosión seguida de una dilatación) y la otra mitad son cierres (una dilatación seguida de una erosión) de diferentes tamaños y formas. Las *curvelets* son filtros pasa banda con orientación. Se utilizan 26 *curvelets* de distintas orientaciones y frecuencias. Utilizando todos los filtros obtienen una exactitud del 94%, que puede ser también alcanzada seleccionando únicamente los 20 mejores filtros. No se indica cuáles son los tipos de defectos que se utilizan para el estudio.

Algunos defectos sobre la superficie de chapa laminada son producidos por los rodillos utilizados para dar el grosor a la chapa durante la laminación y para desenrollar las bobinas para ser cortadas. Estos defectos tienen la peculiaridad de presentar un ciclo con un periodo que coincide con el perímetro del rodillo. En [Bul13] se propone

un sistema para la detección de estos defectos periódicos, en el que todos los defectos detectados se consideran candidatos a pertenecer a un defecto periódico. Para que un defecto pertenezca a una misma serie de defectos se establecen una serie de tolerancias relativas a la posición en la dirección transversal del defecto, la distancia entre dos defectos consecutivos y la forma del defecto.

Recientemente, en [Cho14] se presentó un método para la detección de grietas en las proximidades de los bordes de la chapa laminada utilizando filtros de Gabor. Para ello utilizan cuatro cámaras, dos en cada borde, una para ver la cara superior y otra para ver la cara inferior. Son cámaras lineales que toman una imagen cada vez que la chapa avanza 0,5mm, con una resolución en la dirección perpendicular al avance de 0,05mm/pixel. Las grietas que busca detectar aparecen en la dirección de avance de la bobina y presenta un bajo contraste con la bobina sin defectos. Para resolver el problema utilizan dos filtros de Gabor orientados en la dirección esperada de las grietas pero con dos frecuencias distintas que permitan detectar frecuencias altas y bajas, es decir, grietas de distinta anchura. La salida de la imagen filtrada por estos dos filtros se combina usando una media ponderada y se aplica un método de binarización para marcar los píxeles candidatos a ser defectos. Estos píxeles se unen formando regiones y se analizan sus características para determinar cuáles son defectos y cuales son falsos positivos.

Las técnicas descritas para la detección de defectos sobre la superficie de chapa laminada permiten la detección de un número limitado de tipos de defectos y con unas características bien definidas. La presencia de suciedad y lubricantes sobre la superficie de la chapa laminada así como las vibraciones que se producen durante el procesado del material no se suelen considerar.

Diferenciar las zonas libres de defectos de los defectos es una tarea compleja. En esta tesis, con el objetivo de conseguir detectar la mayor parte de los defectos presentes, se van a adoptar estrategias de detección de defectos conservadoras, es decir, técnicas para la detección de un gran porcentaje de los defectos presentes en la superficie de la chapa laminada a consta de producir un gran número de falsos positivos o pseudodefectos, que serán reclasificados como tal en la etapa de clasificación.

2.3.3.2 Clasificación de defectos

La clasificación de los defectos en chapa laminada es importante para identificar y poder corregir la causa de los defectos. Para ello se utilizan distintos tipos de clasificadores, como son las redes neuronales, las máquinas de soporte vectorial, los clasificadores de vecinos cercanos, lógica difusa, etc.

Las redes neuronales utilizadas en la clasificación de defectos superficiales suelen estar compuestas de una capa de entrada, una capa oculta y una capa de salida. El número de neuronas de la capa de entrada coincide con el número de características utilizadas para la clasificación mientras que la capa de salida coincide con el número de tipos de defectos. La elección del número de neuronas de la capa oculta va a determinar el comportamiento del clasificador. Si su número es demasiado bajo el clasificador no funcionará correctamente, mientras que si su número es elevado se ajustará excesivamente a los datos y perderá generalidad. La técnica de entrenamiento habitual de estas redes es el método *BackPropagation*.

En [Yaz08] se utiliza una red neuronal para la clasificación de 5 tipos de defectos: pegaduras, moho blanco, falta de recubrimiento, suciedad y manchas en el decapado. Utiliza una red neuronal con 9 entradas, 13 neuronas ocultas y 5 salidas entrenadas por el método *BackPropagation*, obteniendo una exactitud superior al 97%. Las entradas son medidas estadísticas del nivel de gris de las zonas detectadas como defectos en un paso previo, aunque no se dan valores de comportamiento sobre la etapa previa ni se tiene en cuenta la presencia de pseudodefectos en las muestras utilizadas para el entrenamiento.

En [Pen09] también se clasifican 5 tipos de defectos: latigazos, marcas de rodillo, moho blanco, agujeros y grietas. Los defectos han sido previamente detectados por una binarización y la muestra disponible consiste en 60 defectos de cada tipo. No se ofrece información sobre cómo se realiza el proceso de entrenamiento con esas muestras. La principal aportación de este trabajo es la utilización de una variación en el algoritmo *BackPropagation* que para el caso de estudio consigue mejorar los tiempos de entrenamiento de la red neuronal.

En [Mas12] se utiliza una red neuronal convolucional (*Convolutional Neural Network*, CNN) que es un tipo de red neuronal donde las neuronas individuales responden a regiones solapadas en el campo visual. Las imágenes son filtradas con diferentes filtros 2D y las imágenes filtradas son utilizadas como entradas a la red neuronal. Las imágenes son redimensionadas a un tamaño de 150 x 150 píxeles y son divididas en sub-imágenes utilizando la media o el valor máximo de cada sub-imagen filtrada como valor de entrada a la red neuronal. El resultado es un error del 7% en la clasificación de 7 defectos, pero no se describe qué defectos son. En el trabajo presentado se supone que no existen defectos no detectados ni falsos positivos en la fase de detección previa a la clasificación.

El método de las redes neuronales también se utiliza para la separación de defectos frente a no defectos, como en [Kan05]. En este caso se utiliza una red neuronal sencilla con 5 neuronas de entrada, cuyas características han sido seleccionadas mediante el análisis de componentes principales, y una única neurona de salida cuya

función es una sigmoidea tangente hiperbólica y cuyos valores pueden variar entre -1 y 1. En este trabajo el método solo se aplica a dos imágenes, con un entrenamiento previo de la red con dos sub-imágenes de 10 x 15 píxeles extraídas de cada una de las imágenes utilizadas para la verificación.

El método de los vecinos cercanos fue utilizado en [Dup97] para la clasificación de 8 defectos, clasificando correctamente el 79,3% de los defectos. Para obtener estos resultados utiliza 22 características, y mide la distancia de cada nuevo caso con todos los individuos de la base de datos de defectos conocidos. La base de datos consta de 1.525 imágenes de las cuales 607 son utilizadas para el entrenamiento.

Las máquinas de soporte vectorial (*Support Vector Machine*, SVM) han sido utilizadas tanto para diferenciar defectos de pseudodefectos, como para realizar clasificación multiclase. Para aplicar la clasificación multiclase se requieren tantos clasificadores binarios como defectos se desea distinguir. Cada clasificador binario etiqueta a un defecto como positivo, mientras que el resto de los defectos son etiquetados como negativos. En [Suv12] se clasifican 4 tipos de defectos utilizando 4 clasificadores binarios SVM aplicados a cada uno de los píxeles de la imagen considerados defectuosos. Para determinar qué tipo de defecto contiene la imagen se utiliza el voto por mayoría, asignando la imagen al tipo de defecto más votado por los clasificadores binarios. Consiguen el 100% de acierto en tres de los defectos y el 82% en el cuarto, aunque solo utilizan 31 imágenes, por lo que no parece que el conjunto de imágenes utilizado para la verificación sea lo suficientemente significativo.

También se han realizado trabajos en los que se comparan distintos métodos de clasificación, como en [Yun11], donde se comparan los métodos de vecinos cercanos, redes neuronales y máquinas de soporte vectorial con un método desarrollado por los autores basado en el algoritmo de AdaBoost. Para la comparación se utilizan 400 imágenes para el entrenamiento y 100 para la verificación, conteniendo todas ellas algún defecto. Se busca distinguir entre seis tipos de defectos: borde dentado, rayones, soja, inclusiones, soldaduras y pliegues. El algoritmo de vecinos cercanos utiliza un valor de $k=3$. La red neuronal tiene una capa de entrada de 11 neuronas, una capa oculta de 8 y una capa de salida de 6. Sobre el método SVM no se ofrecen detalles de su implementación. El algoritmo de AdaBoost propuesto evalúa 160.000 características, pero no se especifica cuantas son seleccionadas al aplicarlo. Los resultados obtenidos con las 100 imágenes de verificación oscilan entre el 80% alcanzado por el método de vecinos cercanos y el 94% por el método de AdaBoost propuesto. El número de imágenes de verificación no parece lo suficientemente significativo para poder obtener resultados concluyentes.

La clasificación de defectos se realiza fundamentalmente utilizando clasificadores supervisados como las redes neuronales y el algoritmo de vecinos cercanos. No existe ningún procedimiento para determinar a priori cuál es el mejor clasificador para cada caso, aunque sí que se puede afirmar que la combinación de varios clasificadores mejora los resultados obtenidos por un único clasificador. Por otro lado, la elección de las características que utiliza el clasificador es fundamental para obtener los resultados deseados y es el aspecto que se va a tratar de mejorar en esta tesis. Además, en el entrenamiento de los clasificadores se debe prestar especial atención al problema del sobreentrenamiento, que puede provocar la falta de generalidad de los clasificadores.

2.4 Inspección de la superficie de carreteras

La red de carreteras pavimentadas es extremadamente extensa, alcanzando los 4,4 millones de kilómetros en la Unión Europea (EU27), [Eur12]. En Estados Unidos alcanza los 5 millones de kilómetros de longitud y el gasto de mantenimiento fue de 22.654 millones de euros en 2008, [OEC11]. La Unión Europea recomienda un gasto anual del 2% del valor de la red en su mantenimiento, [Lax01]. Todas estas cifras indican que las decisiones que se toman sobre cómo se gasta el presupuesto de mantenimiento tienen una gran importancia.

El mantenimiento y la rehabilitación de las carreteras requieren disponer de información precisa sobre el estado de la red vial. Para optimizar las labores de mantenimiento de la red vial se utilizan sistemas de gestión del pavimento, [Lep99]. Estos sistemas toman decisiones sobre qué acciones de mantenimiento deben llevarse utilizando modelos del estado de la carretera para evitar la subjetividad de los inspectores, [Bia12], y para determinar el periodo óptimo de mantenimiento, [Zha12]. Por tanto, la recolección y evaluación de grandes volúmenes de datos sobre el estado del pavimento es uno de los componentes principales para que los sistemas de gestión del pavimento sean efectivos. El deterioro de la superficie de la carretera debido a su uso, juega un papel fundamental en la predicción del nivel de servicio de la carretera.

La inspección visual de la superficie de la carretera es el método más común utilizado para evaluar el grado de degradación de la misma. Para realizar esta tarea, los operarios caminan por el arcén de la carretera y anotan los defectos que observan. Además de ser una tarea muy peligrosa, los resultados que se obtienen no alcanzan los requerimientos de valoración del estado de la carretera deseados, ya que estos resultados no son ni fiables ni repetibles. Más aún, la velocidad a la que estos datos son adquiridos no superan los 5km/h. Aunque el sistema visual humano tiene una gran habilidad para adaptarse a nuevas y variables tareas, ante un mismo estado de la

carretera dos operarios pueden clasificarlo de forma diferente, debido a la subjetividad humana y a la fatiga visual que se produce al realizar trabajos monótonos.

En este contexto, las empresas dedicadas al mantenimiento de carreteras desean tener a su disposición un sistema de gestión automático del estado del firme. El mayor reto para conseguirlo es automatizar el proceso de evaluar el deterioro del firme. Las grietas son el principal defecto que aparece sobre la superficie de la carretera y es por ello que un gran número de técnicas se han propuesto para la detección y clasificación de este tipo de defectos.

2.4.1 Principales deterioros en pavimentos bituminosos

En 1989 la Dirección General de Carreteras (DGT) del Ministerio de Obras Públicas y Urbanismo (MOPU) publicó el catálogo de deterioros de firmes en carreteras, [MOP89], el cual sigue siendo la principal fuente de referencia para la clasificación de defectos superficiales en carreteras en España. Este catálogo distingue entre los defectos que pueden presentarse en pavimentos bituminosos y en pavimentos rígidos.

Según este catálogo, los pavimentos bituminosos, que son el objeto de estudio de este trabajo, se clasifican en cuatro grandes grupos: deformaciones, roturas, desprendimiento y exudaciones. En la **Figura 3** se puede ver la descripción de cada uno de estos tipos de defectos.

DESCRIPCIÓN DE LOS DETERIOROS

PAVIMENTOS BITUMINOSOS

DEFORMACIONES	ROTURAS	DESPRENDIMIENTOS	EXUDACIONES
Rodera	Fisura long. central	Firme brillante	Exudación
Hundimiento	Fisura long. lateral	Descarnadura	Flujo de ligante
Blandón	Fisura transversal	Áridos pulimentados	Mancha de humedad
Cordón longitudinal	Fisura errática	Peladura	Ascensión de finos
Arrollamiento transversal	Fisura parabólica	Bache	
Firme ondulado	Piel de cocodrilo	Fallo envuelta	
Ondulación	Cuarteo malla gorda	Desintegración	
Huella	Fisuras en rodadas	Verruga	
Protuberancia	Fisuras en borde de calzada	Estriado	
Burbuja	Fisura reflejada		
	Fisura curva		
	Fisuras finas		

Figura 3. Descripción de los deterioros en firmes según el "Catálogo de deterioros en firmes" publicado por el MOPU en Abril de 1989, [MOP89]

El grupo de roturas, también conocido como grietas o fisuras, es el tipo de deterioro que aparece con mayor frecuencia en la superficie del pavimento, siendo la causa principal de su aparición la fatiga de las capas asfálticas superiores. En las primeras fases de desarrollo de las grietas suelen aparecer aisladas o como un serie de grietas interconectadas, mientras que en fases posteriores suelen aparecen conectadas y con una concentración de grietas elevada. Otras importantes causas de su aparición son las condiciones atmosféricas (especialmente las bajas temperaturas), la baja calidad de las mezclas asfálticas y la debilidad de las capas inferiores del pavimento. En la **Figura 4** se encuentran algunas imágenes de estos tipos de defectos que se pueden encontrar en las carreteras.

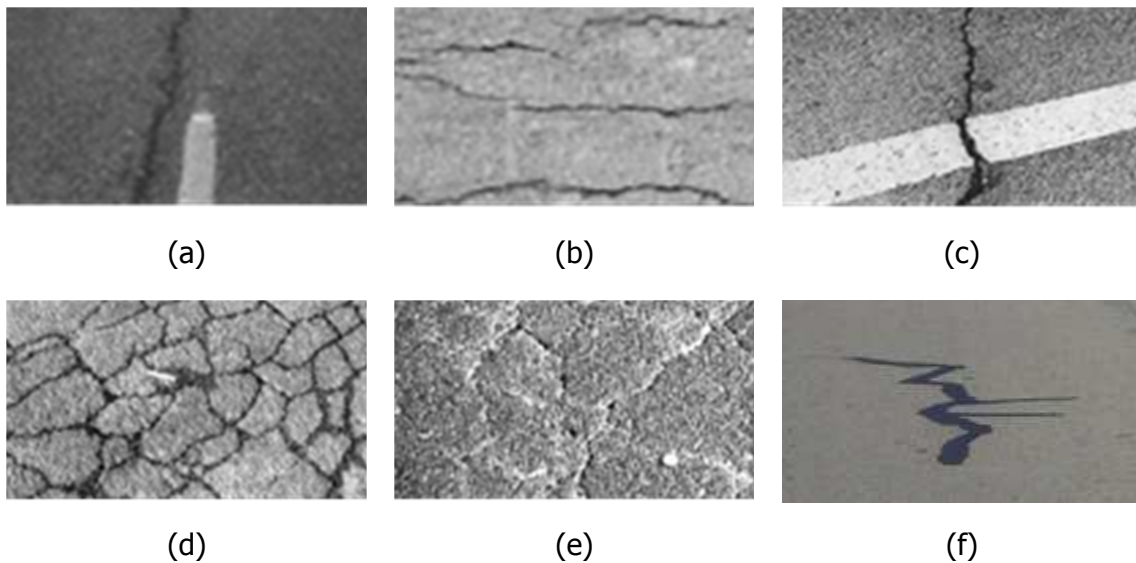


Figura 4. Imágenes de algunos de los defectos que pueden aparecer en la superficie de la carretera. (a) Grieta longitudinal; (b) Fisuras finas; (c) Grieta transversal; (d) Piel de cocodrilo; (e) Cuarteo de malla gorda; (f) Fisura sellada

En este trabajo se va a prestar especial atención a la detección y clasificación de grietas. Desde el punto de vista de la apariencia visual estas grietas se clasifican en grietas longitudinales y transversales. Cuando ambos tipos de grietas se encuentran entremezcladas el tipo de defecto se denomina piel de cocodrilo o cuarteo.

2.4.2 Sistemas de adquisición de imágenes de la superficie de la carretera

Los procedimientos para la inspección visual del pavimento de la carretera requieren la recolección periódica de los deterioros que aparecen sobre la superficie, constituyendo una de las fuentes de información más importantes para el análisis de la condición del pavimento, permitiendo la evaluación de la calidad estructural y funcional de la carretera. Esta recolección periódica de datos pretende la identificación

del alcance y la gravedad de las diferentes tipos de deterioros superficiales del pavimento de la carretera.

Tradicionalmente, las inspecciones visuales de la superficie de la carretera son realizadas por operarios humanos que recorren la carretera y van anotando los deterioros que encuentran. En algunos casos se utiliza la ayuda de equipamiento informático, como es el caso de los sistemas DESY, [Pin03], y VIZIROAD, [Mih11], en los que el operario pulsa un botón específico cada vez que encuentra un tipo de defecto. En la **Figura 5** se muestran estos sistemas, los cuales permiten la incorporación de imágenes adquiridas por cámaras digitales, aunque el responsable de la detección y clasificación de los defectos sigue siendo el operario.



Figura 5. Sistemas de ayuda para la recogida de datos de defectos en la carretera. A la izquierda VIZIROAD. A la derecha DESY.

La incorporación de equipos de video o fotografía embarcados en vehículos ha sido propuesta por diferentes fabricantes, permitiendo un aumento en la velocidad de inspección e incrementando la seguridad de los operarios. Estos nuevos procedimientos permiten la posibilidad de revisar los datos en la oficina tantas veces como sea necesario sin necesidad de desplazarse de nuevo a la carretera.

El sistema Roadrecon de PASCO fue el primer sistema automático de recogida de imágenes de la carretera, [Gra93]. En la **Figura 6** se encuentra una imagen del vehículo de inspección. Consiste en una cámara de 35mm montada en una estructura voladiza sobre el techo de un autobús que adquiere una película continua de la superficie de la carretera. Las imágenes deben ser tomadas de noche con luz artificial para controlar la iluminación. El vehículo podía circular hasta 80km/h y tomar imágenes de la carretera de 4,8m de ancho y 15m de largo. La interpretación de las imágenes se realizaba de forma manual.



Figura 6. Sistema de inspección de carreteras Roadrecon de PASCO. Fuente Highway & Airport Research Lab, <http://pavement.hanyang.ac.kr/research/pms.html#>

El sistema Komatsu fue desarrollado a finales de los 80 por un consorcio japonés, [Fuk90]. Podía circular a 10km/h y capturar imágenes de 2.048 x 2.048 píxeles. Solo podía trabajar de noche para tener controladas las condiciones de iluminación. Tampoco realizaba ningún análisis automático de las imágenes.

Earth Technology Corporation desarrolló a principios de los 90 el sistema conocido como *Pavement Condition Evaluation Services* (PCES), [Wan00]. Fue el primer sistema en utilizar cámaras lineales. Éstas tenían una resolución de 512 píxeles para capturar 2,5m de ancho de la carretera. Podía circular de día utilizando luz artificial realizando la captura a la velocidad de la vía. Asimismo realizaba un procesamiento en tiempo real de los datos.

Uno de los sistemas comerciales más populares en los Estado Unidos desarrollado en los años 90 fue el ARAN, *Automated Road Analyzer*, desarrollado por Frugo Roadware, [Fug15], y del que se muestra una imagen en la **Figura 7**. Este sistema utiliza dos cámaras matriciales e iluminación estroboscópica sincronizada con la adquisición para adquirir el ancho completo de un carril de carretera (4,3m) con una resolución de 2mm x 2mm por píxel. Así mismo se suministra con un *software* de análisis de imágenes, WiseCrax, que realizaba el análisis en la oficina, aunque requería un alto grado de intervención humana. Este sistema ha seguido evolucionando con el tiempo, convirtiéndose en un sistema modular que puede incorporar en el vehículo todo tipo de sensores de estado de la carretera, como perfilómetros, cámaras para ver el entorno de la carretera, GPS, texturómetros, etc. Aunque aún se encuentra disponible la versión de cámaras matriciales, actualmente también puede incorporarse

el equipo LRIS, que utiliza cámaras lineales e iluminación láser infrarroja, que se describe más adelante.



Figura 7. Sistema de inspección de carreteras ARAN de Fugro Roadware. [Fug15]

En el Reino Unido, a finales de los 90, el laboratorio de investigación del transporte desarrollo el sistema HARRIS (*Highways Agency Road Research Information System*). Dispone de 3 cámaras lineales de 512 píxeles cada una, que permite adquirir un ancho de 2,9m de la carretera con una resolución de 2mm x 2mm. Utiliza iluminación halógena y podía circular a una velocidad de hasta 80km/h, [Pyn99].

International Cybernetics Corporation (ICC) también desarrolló su propio sistema de recogida de imágenes de carretera. El sistema contaba con cámaras matriciales de color con una resolución de 1.300 x 1.024 píxeles para capturar tanto la carretera como los laterales de la misma con un intervalo de unos 8 metros. Además disponía de una cámara lineal de 2.048 píxeles y luz artificial. Las imágenes estaban georreferenciadas y se mostraban al usuario con un *software* específico de visualización, pero no permitía el procesamiento de las imágenes. Actualmente ICC también ha optado por sustituir el equipo de adquisición de imágenes por el LRIS que se describe más adelante, [Int15].

El sistema RoadCrack de CSIRO, [CSI15], es un sistema automático de detección de grietas en carretera desarrollado y utilizado principalmente en Australia. El sistema

necesita ser embarcado en un camión ya que requiere una gran infraestructura y un generador propio para obtener la potencia necesaria para alimentar los equipos. El sistema consta de 4 cámaras lineales, iluminación y reflectores para concentrar la luz, un codificador incremental y 13 ordenadores para procesar las imágenes en tiempo real. La anchura del pavimento que analiza es de 2,4m y detecta grietas superiores a 1mm de anchura a una velocidad de 105km/h. No almacena la información capturada.

Investigadores de la Universidad de Texas, en Austin (EE.UU), desarrollaron un vehículo de inspección denominado *TxDOT Pavement Survey Vehicle*. El sistema consta de una cámara lineal de 2.048 píxeles de resolución, una tarjeta capturadora de imágenes, un ordenador y un *software* de análisis de imágenes denominado Vcrack, [Hua06]. El sistema captura imágenes de 3 metros de ancho, con una resolución en la dirección de avance de 1,75mm a una velocidad de hasta 110km/h. Inicialmente utilizaba luz natural, aunque posteriormente se mejoró el sistema incorporando iluminación LED lineal para que la iluminación fuera más constante y poder utilizarse durante la noche. Permite la clasificación de distintos tipos de grietas, realizando la clasificación en tiempo real si la velocidad del vehículo es menor de 70km/h.

En Suecia, Ramboll desarrolló el sistema PAVUE, [Sjö95], un sistema que utiliza 4 cámaras matriciales analógicas de 581 x 752 píxeles e iluminación estroboscópica embarcadas en la parte posterior de un vehículo. Las imágenes son almacenadas con una resolución de tan solo 400 líneas debido a las limitaciones de almacenamiento de los equipos utilizados. Adquirían 3,2m del ancho de la carretera y 1m en la dirección de avance, lo que supone una resolución de 2,2mm a una velocidad máxima de 90km/h. Producía mapas de grietas sincronizados con la distancia recorrida, pero para el procesamiento era necesario el ajuste manual de los parámetros de análisis. El sistema se ayudaba de 6 distanciómetros láser, que consideraba que había una grieta cuando la distancia variaba súbitamente, calculando el ancho y la profundidad de la misma. La precisión y fiabilidad de este sistema no era elevada ya que el láser se encontraba en solamente seis puntos, por lo que solo detectaba las grietas que aparecían debajo de esas localizaciones, y el cálculo del ancho de la grieta se veía distorsionado por la dirección de la grieta.

La Universidad de Arkansas desarrolló el sistema conocido como *Digital Highway Data Vehicle* (DHDV), [Wan05], un sistema con una cámara de video y 4 fuentes de iluminación estroboscópica embarcados en la parte trasera del vehículo. La imagen tenía una resolución de 1.300 x 1.024 píxeles y podía circular a 95km/h.

En la Universidad de Utah, desarrollaron un sistema de inspección de carreteras denominado PicCrack, [Lee04], que se muestra en la **Figura 8**. El sistema consistía en una cámara matricial de video digital que se encontraba suspendida a 2,7m de altura detrás del vehículo. La cámara tenía una resolución de 776 x 582 píxeles, para adquirir

un ancho de 3,3m y una longitud de 2,7m, los cual supone una resolución aproximada de 5 x 5 mm por píxel. Además de la baja resolución de la imagen obtenida, las lentes presentaban una elevada distorsión en los bordes.



Figura 8. Sistema de inspección de carreteras PicCrack. Fuente: <http://www.go-explore-trans.org/david-lee/2/>

El sistema de adquisición LRIS (*Laser Road Imaging System*), fabricado por INO, [INO15], en exclusiva para PAVEMETRICS, [PAV15], es un sistema comercial que utiliza cámaras lineales de alta resolución e iluminación láser de alta potencia. En la **Figura 9** se muestra una imagen del mismo. Dispone de dos cámaras lineales de 2.048 píxeles cada una para adquirir 4m del ancho de la carretera. La velocidad máxima de adquisición de las cámaras es de 28.000 líneas por segundo, pudiendo circular a una velocidad máxima de 100km/h, por lo que la resolución de las imágenes es de 1 x 1 mm por píxel. El sistema de iluminación lineal láser tiene una longitud de onda de 808nm. Un filtro de esa misma longitud de onda impide que el resto de las longitudes de onda incidan en el sensor de las cámaras, de tal manera que hace al sistema prácticamente inmune a la luz exterior, pudiendo utilizarse el sistema tanto de día como de noche.

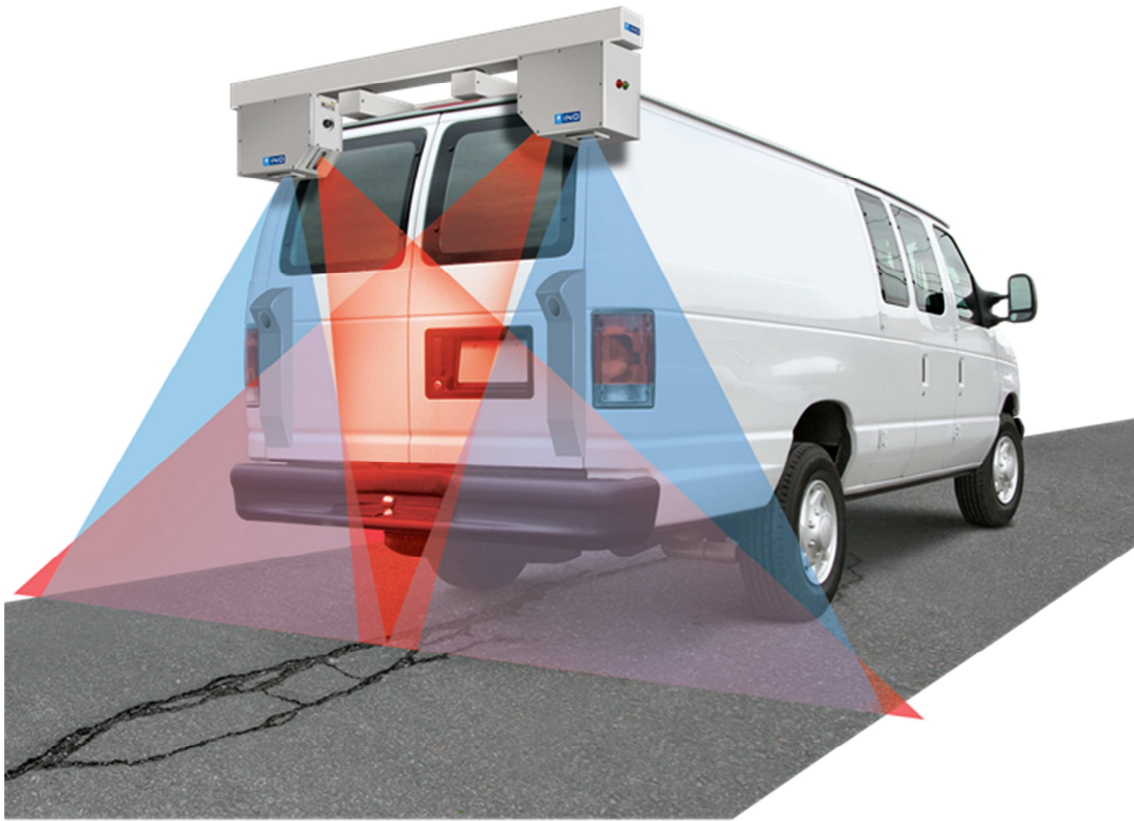


Figura 9. Laser Road Imaging System (LRIS), [INO15]

A continuación, en la **Tabla 1** se muestra una comparativa de los sistemas de inspección de carreteras descritos. En dicha tabla se comparan, cuando los datos se encuentran disponibles, el tipo de cámara, el ancho de carretera inspeccionada, la resolución de la imagen, la velocidad máxima de adquisición, la fuente de iluminación empleada (indicándose entre paréntesis si es posible utilizarlo de día y/o de noche) y si el procesamiento de las imágenes se realiza de forma automática o manual.

Los datos mostrados en la tabla ponen de manifiesto que el sistema comercial que mejores prestaciones ofrece es el sistema LRIS y es por ello que varios de los fabricantes que siguen comercializando sistemas de inspección de carreteras han optado por sustituir el sistema de adquisición de imágenes original por el LRIS. Aun así, el sistema presenta ciertas limitaciones para poder apreciar defectos que se podrían poner de manifiesto con iluminación oblicua, como ocurre con algunas grietas en las que el fondo de las mismas es muy similar a la carretera sin defecto.

Tabla 1. Comparativa de sistemas de inspección de carreteras

Sistema	Tipo de Cámara	Ancho	Resolución	Velocidad	Iluminación	Procesamiento
RoadRecon	35mm	4,8m	-	80km/h	Luz artificial (solo noche)	Manual
Komatsu	Matricial 2.048x2.048	-	-	10km/h	Luz artificial (solo noche)	Manual
PCES	2 cámaras lineales de 512 píxeles	2,5m	2,5x2,5mm	100km/h	Luz artificial (día y noche)	Tiempo real
ARAN	Cámaras matriciales	4,3m	2x2mm	Velocidad de la vía	Estroboscópica (día y noche)	Oficina con intervención humana
HARRIS	3 cámaras líneas de 512 píxeles	2,9m	2x2mm	80km/h	Halógena (día y noche)	-
ICC	1 cámara lineal de 2048 píxeles	-	-	-	Luz artificial	Mostraba imágenes pero no procesaba
RoadCrack	4 cámaras lineales	2,4m	1x1mm	105km/h	Reflectores para concentrar la luz	Procesa en tiempo real pero no almacena la información
TxDOT Pavement Survey Vehicle	1 cámara lineal de 2.048 píxeles	3m	1,5x1,75mm	110km/h	Luz natural e iluminación LED	En tiempo real si la velocidad es menor de 70km/h
PAVUE	4 cámaras matriciales analógicas de 581x752 píxeles	3.2m	2,2x2,2mm	90km/h	Estroboscópica (día y noche)	Ajuste manual de parámetros
DHDV	Cámara de video de 1.300x1.024 píxeles	2.25m	1,75x1,75mm	95km/h	Estroboscópica (día y noche)	-
PicCrack	Cámara matricial de 776x582 píxeles	3,3m	5x5mm	-	-	-
LRIS	2 cámaras lineales de 2.048 píxeles	4m	1x1mm	100km/h	Láser 808nm (día y noche)	No incluido

2.4.3 Procesamiento de imágenes

Las técnicas para la detección de deterioros en la superficie de la carretera son cada vez numerosas, impulsadas recientemente por la aparición de equipos comerciales para la adquisición de imágenes. La mayor parte de las publicaciones tratan la detección y la clasificación de grietas, dado que es el tipo de defecto más común presente en las carreteras.

En [Mah91] se encuentra la primera referencia a la detección de grietas en la superficie de la carretera. Utilizaba una cámara matricial con una resolución de 512 x 512 píxeles y luz estroboscópica. Las imágenes adquiridas tenían distinta resolución, variando entre 1 y 5 mm por píxel, y su objetivo fundamental era detectar las grietas más severas. Para detectar las grietas utilizaban una binarización cuyo umbral es determinado automáticamente utilizando una técnica que denominaron el histograma gradiente, que consistía en aplicar un gradiente sencillo a la imagen, $g_{i,j} = ((I_{i,j} - I_{i-1,j})^2 + (I_{i,j} - I_{i,j-1})^2)^{1/2}$. A continuación, se calcula la media ponderada del nivel de gris de la imagen usando como factor de ponderación para cada píxel el valor del gradiente anteriormente calculado. De la misma forma se calcula la desviación estándar. El umbral para binarizar cada imagen es la media menos la desviación estándar calculada en el paso anterior. A continuación, se extraen los píxeles que se encuentran unidos y se eliminan los que tienen un área pequeña. Esta técnica solo permite detectar las grietas más severas.

Son muchas las técnicas utilizadas desde entonces para la detección y clasificación de defectos en carreteras, las cuales han sido recopiladas, clasificadas y comparadas por algunos autores. Uno de los trabajos más completos es el realizado por Chambon y Moliard, [Cha11], en el que se clasifican las principales técnicas de procesamiento de imágenes de la superficie de la carretera en 5 categorías:

1. Basados en el **análisis del histograma**: Estos son los métodos más antiguos pero los más populares. Consisten en seleccionar un umbral de binarización basándose en el análisis del histograma de la imagen, realizando hipótesis de distribución normal de los niveles de gris y/o aplicando técnicas de umbralización adaptativas o locales. Estas técnicas son simples y su aplicación no consume mucho tiempo de procesamiento, pero suelen ofrecer un elevado número de falsos positivos. En realidad, estos métodos asumen que la distribución del nivel de gris de la imagen es diferente en la zona con grietas es diferente de la zona libre de defectos y puede separarse fácilmente. En muchas ocasiones esta hipótesis no es válida, por lo que los resultados que ofrecen estas técnicas tampoco lo son.
2. Basados en herramientas de **análisis morfológico**: Requieren un umbral inicial para la binarización de la imagen, para a continuación aplicar una serie de filtros morfológicos que reducen el número de falsos positivos

significativamente. Puede considerarse como una técnica de post-procesamiento para la categoría anterior.

3. Basados en una fase de **aprendizaje**: Incluyen todos los métodos que requieren un entrenamiento para su aplicación, siendo el método más habitualmente empleado las redes neuronales.
4. Basados en técnicas de **filtrado** de imagen: Esta familia de técnicas incluyen entre otras la detección de bordes, el filtrado adaptativo, filtros de Gabor, filtros de respuesta de impulso finito, derivadas y ecuaciones diferenciales. Suelen partir de la hipótesis de que las grietas corresponden a variaciones en la textura normal de la carretera.
5. Basados en el análisis de **modelos**: Se basan en el análisis de propiedades globales de la imagen frente a las propiedades locales. Normalmente realizan un análisis multiescala de la textura de la imagen combinado con algoritmos de trayectorias mínimas, o realizan la detección local de puntos de interés combinado con contornos geodésicos.

También es habitual la clasificación de las técnicas de procesamiento según las fases en las que se lleva a cabo, como la utilizada en [Oli09a] en la que se dividen los pasos del procesamiento en tres fases:

1. **Preprocesamiento**: que tiene por objeto reducir el ruido de la imagen y la falta de uniformidad en la distribución de la iluminación en la imagen.
2. **Detección**: que engloba a las técnicas que buscan la detección de los píxeles candidatos a contener grietas.
3. **Clasificación y caracterización**: en la que se determina el tipo de grieta y se miden las características de la misma, como la longitud y la apertura.

Estos mismos autores presentaron en [Oli08] un método para la homogeneización de las imágenes, con el fin de evitar que la no uniformidad de la iluminación durante la fase de adquisición de imágenes afectase al resultado. Para ello se divide la imagen en sub-imágenes, se calcula el valor medio de cada sub-imagen y se realiza una preselección de las celdas candidatas a contener grietas debido a que su nivel de gris se encuentre por debajo de un umbral calculado a partir de los valores de media y desviación de las celdas de alrededor. Las celdas que presuntamente no contenían grietas eran multiplicadas por un factor de corrección resultado de dividir el valor objetivo entre el valor medio buscado. Las celdas candidatas a contener grietas son multiplicadas por un factor obtenido como media de los factores de corrección de las celdas que las rodeaban. El buen funcionamiento del método está supeditado a que el método de preselección de las celdas con grietas funcione correctamente, de lo que no se ofrecen resultados.

Otra de las problemáticas que pueden aparecer en la imagen es la presencia de sombras. Zou et al. presentaron recientemente un método para reducir este problema [Zou12]. Las sombras reducen el contraste en la zona de la imagen donde están presentes por lo que simplemente incrementar la intensidad de la imagen no es suficiente. Además, las sombras normalmente presentan una penumbra alrededor, por lo que la intensidad de la sombra no es homogénea. En este trabajo se propone la discriminación de cada zona de sombra según su intensidad mediante el cálculo de geodésicas y la aplicación de una corrección de la iluminación que iguala las desviaciones estándar de la zona de sombras con la zona sin sombras. Aunque el resultado es aceptable, la solución no es tan buena como la que se consigue con el equipo LRIS que utiliza iluminación de alta potencia para minimizar la influencia del sol en las imágenes.

Para la detección de grietas el método más simple y común es la segmentación mediante la utilización de un umbral que discrimina los píxeles que corresponden a grietas de los que no.

En [Zha09] se propone una técnica de detección de grietas por segmentación, aplicando después a la imagen operaciones morfológicas. Con esta técnica se detectan grietas longitudinales, transversales y de cocodrilo, pero no se detalla el método de segmentación ni se dan resultados objetivos de los resultados obtenidos.

Una técnica más elaborada para la identificación de grietas se puede encontrar en [Qin08], donde la segmentación es realizada mediante el método que los autores denominan *Neighboring Difference Histogram Method* (NDHM). Este método compara cada píxel con el histograma de los píxeles de alrededor construyendo una nueva imagen con esta información para que los nuevos valores de los píxeles con defectos se diferencien en mayor medida de los píxeles que no contienen defectos. Los resultados obtenidos mejoran los que consiguen los métodos de segmentación de Otsu y de Kapur, pero no se da información de cómo se han obtenido las imágenes y solo se ofrece información sobre la aplicación del método en dos imágenes.

Oliveira y Correia han presentado diversos trabajos sobre detección y clasificación de grietas en la superficie de la carretera. En [Oli09a] proponen un método de detección de grietas en imágenes obtenidas por un equipo LRIS y por una cámara convencional. En primer lugar, se homogenizan las imágenes para obtener un valor medio del nivel de gris de 125, para a continuación aplicar el método de Otsu modificado, [Don08], para buscar los píxeles con menor intensidad y catalogarlos como píxeles pertenecientes a grietas. Dividen la imagen en bloques de 75 x 75 píxeles y calculan la entropía de la imagen binarizada. Una vez más utilizan el método Otsu modificado para calcular el umbral que separa los bloques que contienen grietas de los que no. Finalmente clasifican las grietas en longitudinales, transversales y mezcla.

Estos mismos autores, en [Oli09b] profundizan sobre las técnicas de clasificación no paramétricas analizando tres clasificadores: las ventanas de Parzen, el clasificador lineal de mínimos cuadrados de Fisher y el clasificador de vecinos cercanos.

2.5 Inspección del revestimiento de túneles de hormigón

Los túneles son obras subterráneas lineales utilizadas principalmente en las infraestructuras para el transporte de personas y mercancías. Este tipo de infraestructura es utilizado principalmente por las redes de carreteras, ferrocarriles y metros. Solo en España la Red Nacional de Carreteras cuenta con más de 200km de túneles, los túneles ferroviarios superan los 800km de longitud total y la de red de metro alcanza los 475km. A nivel mundial existen 68 túneles carreteros o ferroviarios de más de 10km de longitud.

La correcta conservación de túneles implica prever actuaciones que se anticipen a la aparición de defectos muy graves o irreparables. Este planteamiento posibilita la política de intervenir en el momento más adecuado provocando la menor alteración posible en la capacidad funcional de la vía de transporte.

Para poder actuar de forma eficiente es imprescindible conocer el estado superficial del túnel con la mayor precisión y objetividad posible. Este conocimiento pasa por el uso de adecuados dispositivos de inspección, detección y evaluación del estado del túnel.

Los métodos tradicionales de inspección de túneles están basados en los criterios subjetivos del observador (mediante una inspección visual se va rellenando una ficha de inspección) y en equipos que, normalmente, necesitan instalación y contacto con la superficie del túnel (testigos de yeso, fisurómetros, etc.)

Para conseguir mejorar estos métodos sería necesario disponer de un sistema automático de adquisición de imágenes del 100% de la superficie del túnel. El problema que se plantea a continuación es la gestión de una masiva cantidad de información que imposibilita su evaluación manual. El procesamiento automático de imágenes adquiridas en la superficie del túnel se presenta como una posible solución para lo que es necesario el desarrollo de algoritmos de procesamiento de imágenes capaces de detectar la aparición de fisuras y grietas en la superficie del túnel.

2.5.1 Defectos en túneles de hormigón

Son múltiples los defectos que pueden aparecer en la superficie del revestimiento de un túnel de hormigón. En esta sección se describen los defectos de mayor interés, tanto por su frecuencia de aparición como por su nivel de riesgo potencial para los usuarios, para los trabajadores o para la propia infraestructura, agrupados en cuatro tipos: mecánico, químico, relacionados con la presencia de agua y relacionados con las juntas.

2.5.1.1 Defectos de tipo mecánico

Las lesiones y sus síntomas en el hormigón de origen mecánico, están originadas por falta de resistencia a esfuerzos de tracción o compresión en el hormigón y de tracción en el acero, debido a las acciones que soportan los elementos estructurales y las deformaciones que sufren. Destacan las siguientes:

- **Fisuras:** Las fisuras en el hormigón, son roturas que aparecen en la superficie del mismo debido principalmente a la existencia de tensiones superiores a su capacidad de resistencia. Cuando la fisura atraviesa de lado a lado el espesor de una pieza, se convierte en grieta.

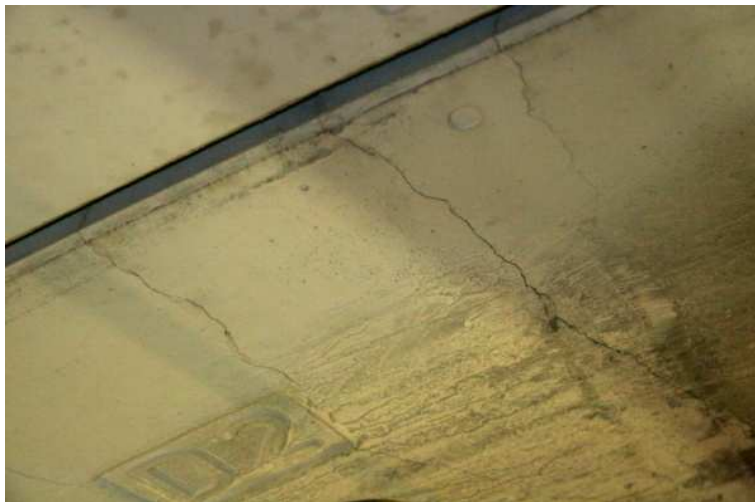


Figura 10. Fisuras en la superficie de un túnel

Las fisuras pueden aparecer por diferentes motivos: retracción de áridos, variaciones bruscas de temperatura, sobrecargas, movimientos del terreno, etc.

- **Desprendimientos o rotura de dovelas:** Se produce cuando se desprende una porción de la dovela causado principalmente por daños provocados durante la perforación por el sistema de empuje hidráulico que se apoya en las dovelas.



Figura 11. Desprendimiento en una parte de una dovela

- **Roturas de esquina y desconchones:** Las primeras como su mismo nombre indica aparecen en las proximidades de los vértices de las esquinas y son de forma triangular, mientras que los desconchones se producen alejados de las esquinas y tienen planta curva. En muchas ocasiones, las grietas que delimitan la incidencia presentan un salto, lo cual denota que se ha producido un cierto movimiento relativo. Algunas veces se ha producido ya un desprendimiento del hormigón, quedando la armadura de intradós al descubierto. Se clasifican de la siguiente forma:
 - Sin peligro de caída de hormigón
 - Con peligro de caída de hormigón
 - Con presencia de agua
 - Sin presencia de agua
- **Armadura vista:** en este defecto la armadura interna del hormigón asoma a la superficie.

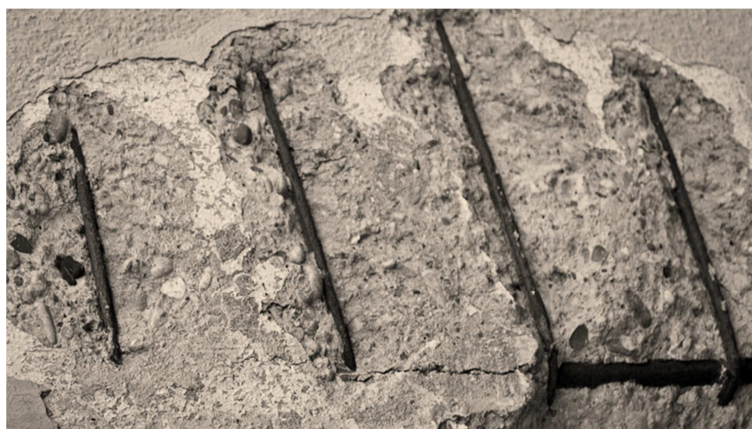


Figura 12. Armadura vista en dovela de hormigón

2.5.1.2 Defectos de tipo químico

Los defectos de tipo químico más importante son los siguientes:

- **Carbonatación / Eflorescencia:** La carbonatación es un proceso lento que ocurre en el hormigón, donde la cal apagada (hidróxido cálcico) del cemento reacciona con el dióxido de carbono del aire formando carbonato cálcico [$\text{Ca(OH)}_2 + \text{CO}_2 \rightarrow \text{CaCO}_3 + \text{H}_2\text{O}$]. Esta reacción, necesariamente se produce en medio acuoso, ya que el dióxido de carbono reacciona con el agua formando ácido carbónico, ya que éste reaccionará con el hidróxido de calcio, obteniendo como resultado el carbonato de calcio y agua. Dado que la carbonatación provoca una bajada de pH (ácido) esto puede llevar a la corrosión de la armadura y dañar la construcción. En la **Figura 13** se muestra un ejemplo de la pared de un túnel de hormigón en el que se ha producido el fenómeno de la eflorescencia.



Figura 13. Eflorescencia en el interior de un túnel

- **Oxidación:** se refiere a la oxidación del acero que forma la armadura interna del hormigón.

2.5.1.3 Defectos relacionados con la presencia de agua

Estas humedades son de naturaleza muy variada y pueden presentarse en diversas zonas y con distintas intensidades. En la **Figura 14** se muestran algunos ejemplos de este tipo de defecto que se clasifican de la siguiente forma:

- **Humedad**
- **Goteo**
- **Filtración pequeña (rezume)**
- **Filtración importante (chorro)**



Figura 14. Diversos grados de humedades en el interior de un túnel.

Los efectos para el túnel son diversos:

- Liberan presión hidrostática sobre el trasdós del revestimiento.
- Los goteos sobre la catenaria y las vías pueden llegar a atacar estos elementos metálicos.
- Lavado de material y generación de huecos en trasdós.
- Transmiten posibles agentes agresivos a zonas del túnel sin patologías o los hacen penetrar a zonas internas del anillo (revestimientos armados).
- Efecto antiestético y que puede alarmar al usuario.

2.5.1.4 Defectos relacionados con las juntas

En el proceso constructivo de un túnel mediante una máquina tuneladora se van colocando dovelas prefabricadas según avanza la excavación del túnel. Las dovelas que conforman el diámetro completo del túnel forman un anillo. La separación entre dovelas y anillos es lo que se conoce como juntas. Los principales defectos que aparecen en las juntas son los siguientes:

- **Apertura:** Excesiva separación en las juntas.
- **Cejas o resaltes:** Es el nombre comúnmente usado para denominar el decalaje de colocación de anillos consecutivos (desalineaciones y discontinuidades). Suelen producirse cuando el confinamiento no es uniforme. El problema que presentan es que la transmisión de las compresiones de los gatos de un anillo a otro se realiza a través de una superficie determinada que viene a ser de un 60-80% del canto, por lo que si no se produce una colocación a tope, se aumenta la excentricidad y, por tanto, la posibilidad de transmisión de esfuerzos fuera del núcleo central. Lógicamente, se producen durante el montaje y no tienen ningún efecto posterior fuera del estético.

2.5.2 Caracterización de las fisuras

El objetivo del sistema de inspección que se va a desarrollar es la detección de grietas y fisuras. Es un tipo de defecto que es posible identificar mediante inspección visual pero que puede pasar desapercibido en una inspección rutinaria. Además es un defecto que aparece en las fases previas del deterioro del túnel, por lo que su detección puede permitir actuar de forma preventiva antes de que se agrave el problema.

Las fisuras se clasifican en tres categorías de daños en función de la apertura de la fisura y del daño aparente que tenga cada dovela.

- **Tipo 0:** Se engloban en esta categoría las fisuras con una apertura inferior a 0,1mm y deben ser detectadas para cumplir con los criterios de durabilidad más exigentes que se piden en la actualidad.
- **Tipo A:** Se engloban en esta categoría las de carácter meramente estético y son aquellas que por el autosellado de la junta o por su apertura, no alcanzan el acero del intradós. Se limita esta categoría de fisuras a las que presentan aperturas entre 0,1-0,4mm.
- **Tipo B:** Fisuras que, pasando la primera capa de armaduras, están estabilizadas y tienen apertura en cabeza menor de 2 mm. La razón de haber escogido este valor estriba en el cálculo de la apertura de fisura necesaria para que la armadura del intradós plastifique (0,4-2mm).
- **Tipo C:** Dovelas multifisuradas.
 - Fisuras tipo **C1:** se trata de dovelas multifisuradas sin riesgo de caída o desprendimiento de hormigón. Las fisuras son del tipo A (<0,4 mm), normalmente con humedades y/o rezume de agua.
 - Fisuras tipo **C2:** se trata de dovelas multifisuradas con riesgo de desprendimiento de hormigón.

2.5.3 Sistemas comerciales para la inspección de túneles

Apenas existen sistemas comerciales para la inspección automática de la superficie de túneles mediante imágenes, aunque sí se encuentran algunos equipos para la adquisición de datos tridimensionales de la superficie del túnel, como el *T-Sight5000* de MERMEC Group [MER15] o el *Túnel Scanning* de Dibit [Dib15]. En ambos casos la resolución de escaneo es de aproximadamente 5mm, lo que no permite la detección de pequeñas grietas o fisuras objeto de estudio en esta tesis.

Existen dos sistemas comerciales que utilizan imágenes para la detección de grietas, aunque no se ofrece la información suficiente para conocer su fiabilidad ni cuáles son las técnicas empleadas para la detección. En ambos casos se considera que el sistema

de detección es semiautomático, es decir, requiere intervención humana para la clasificación del contenido de las imágenes.

El primero de ellos, es el sistema LTSS (*Laser Tunnel Scanning System*) de Pavemetrics, [PAV15b], captura la imagen y el modelo tridimensional de 12m de perímetro del túnel desplazándose a 30km/h con una resolución en la dirección perpendicular de 1mm, por lo que suele requerir varias pasadas para adquirir el perímetro completo del túnel, y con una resolución en la dirección de avance de hasta 1mm. En [Lau14] se describen algunas características del sistema, el cual consta de 6 cámaras e iluminación láser de alta potencia. LTSS es una herramienta para catalogar el contenido del túnel, realizando la detección automática de algunas condiciones del mismo, como el estado de las juntas, la presencia de grietas y la presencia de humedad, aunque no se ofrecen detalles de cómo se realiza esa evaluación. En la **Figura 15** se muestra una imagen del sistema.

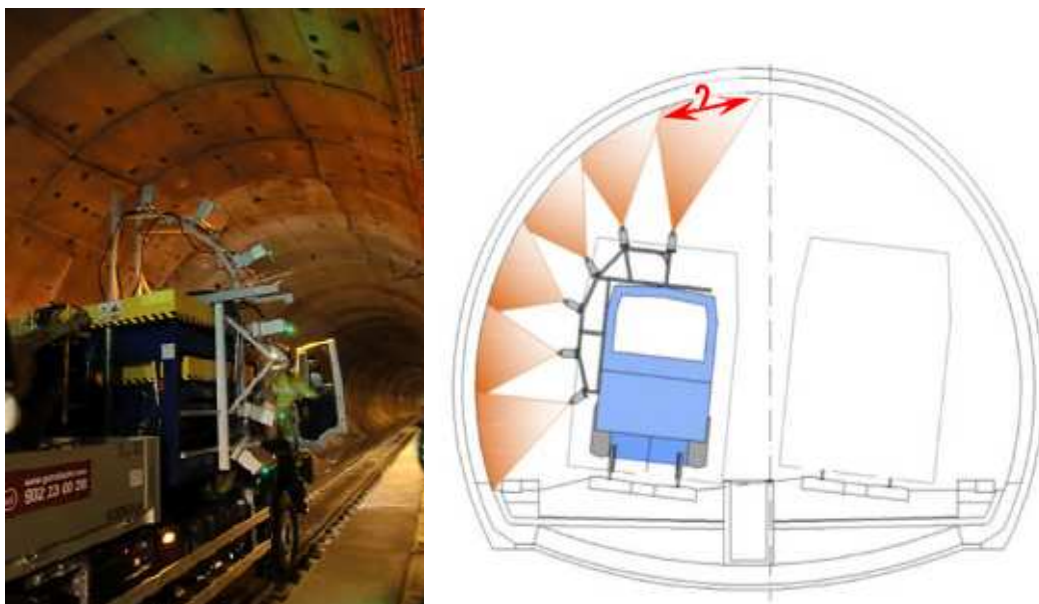


Figura 15. Sistema de escaneado de túneles LTSS de Pavemetrics

Por otro lado, la empresa Terra Vermessungen ha desarrollado el sistema de detección de grietas tCrack, [Ter15]. Este sistema, que se puede ver en la **Figura 16**, permite la inspección de túneles de 10m de diámetro analizando 22m de su perímetro, con una resolución de 0,5mm, lo que permite almacenar y clasificar grietas de 0,1mm de ancho, [Mül12]. Para ello se utilizan 10 cámaras de Basler del modelo Pilot de 5Mpixeles, cámaras cuya resolución es 2.456 x 2.058 píxeles pero que permiten la captura de un tamaño inferior de imagen. El sistema tCrack captura una imagen cada 10mm a una velocidad de 0,7m/s, aunque no se especifica el alto de la imagen capturada. Incluso suponiendo que no hay solapamiento en las imágenes, la resolución

en la dirección transversal debería ser de 0,9mm, pero incluso llegando a resoluciones de imagen de 0,5mm, difícilmente se pueden detectar grietas de 0,1mm de ancho, tal y como se establece en la descripción del producto.



Figura 16. Sistema de detección de fisuras en túneles tCrack de terra vermessungen

2.5.4 Procesamiento de imágenes

Se pueden encontrar un número significativo de trabajos en los que se aborda la detección de grietas en edificios e infraestructuras de hormigón, especialmente en puentes. En [Abd03] se comparan cuatro métodos para la detección de grietas en el hormigón de puentes. Utilizan los detectores de bordes de Sobel y Canny, la transformada de Fourier y la wavelet de Haar. Concluyen que la transformada wavelet de Haar es la que mejores resultados obtiene.

En [Nis12] se describe un método para la detección de grietas en estructuras de hormigón. El método se basa en la utilización de un gran número de filtros que, aplicados a una base de datos procesada manualmente, selecciona, mediante algoritmos genéticos, la combinación óptima de filtros para obtener el mejor resultado posible.

En [Riv15] se describe un método para la detección de grietas en superficies de hormigón mediante cámaras matriciales de color de 18Mpixeles. El sistema es capaz de

detectar grietas con una abertura superior a los 0,25mm. En primer lugar, el color de la imagen es corregido en las zonas sobreexpuestas y en las zonas con sombras. A continuación se transforma la imagen a escala de gris para aplicar un algoritmo de detección de bordes. La aplicación, desarrollada en Matlab, permite elegir entre los algoritmos de detección de bordes de Roberts, Sobel, Prewitt, Canny, Laplace y Gaussiano. Con las imágenes que se utilizan en este trabajo, han llegado a la conclusión que el filtro de Prewitt es el más apropiado. Después aplican operaciones morfológicas para eliminar el ruido y conectar las grietas. Posteriormente se segmenta la imagen, se agrupan los píxeles en grietas, y se extraen características de forma y color de las agrupaciones de píxeles para determinar cuáles corresponden realmente a grietas. Finalmente, se miden las propiedades de las grietas, como su longitud y abertura. El método propuesto requiere cierta intervención del operador y ha sido diseñado con los datos de 1 pared de hormigón de 3x3m y verificado en otras 8 paredes. Los datos globales de cada zona de la imagen obtenidos manualmente se comparan con los obtenidos por el método propuesto, pero no se da ninguna medida estadística para su evaluación.

Sin embargo, la utilización de filtros para detección de bordes es al menos cuestionable, especialmente para imágenes tan altamente texturizadas. En la sección 5.1 se tratará en más profundidad esta circunstancia.

El método de percolación o filtración (*percolation*), basado en un fenómeno físico, [Sta94], es utilizado por varios autores con algunas modificaciones. El algoritmo básico consiste en aplicar a cada píxel de la imagen el siguiente procedimiento. El píxel a evaluar se añade a la región de percolación y se comprueba si el nivel de gris de alguno de los píxeles adyacentes es menor que el nivel de gris del píxel inicial. Si así ocurre, esos píxeles son añadidos a la región de percolación, en caso contrario sólo se añade el píxel cuyo nivel de gris tenga el menor valor. Se repite el proceso con todos los píxeles adyacentes de la región de percolación ajustando el umbral al valor máximo valor del nivel de gris de la región de percolación actual, finalizando cuando la región crezca hasta un tamaño determinado. Si a partir de ese tamaño no se encuentran píxeles que cumplan la condición del umbral, se detiene el proceso, en caso contrario continúa hasta alcanzar el tamaño máximo establecido. Cuando se detiene la búsqueda se analiza la forma de la región de percolación. Si su forma es alargada se considera que es una grieta, en caso contrario se descarta. En [Yam10] se propone una pequeña modificación del algoritmo para acelerar el proceso que consiste en realizar cálculos intermedios para detener la búsqueda lo antes posible, sin necesidad de cumplir las condiciones de parada del algoritmo original, y en marcar los píxeles descartados como píxeles de fondo, de tal manera que si se incorporan a una región de percolación posterior, se detenga también la búsqueda.

En [Fuj11] se propone un método para la detección de grietas en estructuras de hormigón que utiliza dos etapas de preprocesamiento y dos etapas de detección. La primera etapa del preprocesamiento consiste en aplicar el filtro de la mediana y a continuación aplicar el filtro Hessiano para enfatizar líneas, como son las grietas. Para la detección se utiliza un filtro basado en la probabilidad de que un píxel sea considerado grieta, realizando una umbralización local en el entorno de los píxeles asignados como grieta en la primera fase. El algoritmo ha sido probado en 60 imágenes adquiridas con una cámara de 640 x 480 píxeles con una resolución de 0,1mm y cuyo objetivo es la detección de grietas de apertura inferior a 0,2mm. Consiguen una sensibilidad de 0,80 y una especificidad de 0,992, aunque utilizan los mismos datos para el entrenamiento que para la verificación de resultados.

En [Zha15] se representan las imágenes como superficies geodésicas. Consideran que la distancia entre dos puntos de la imagen es la distancia que habría que recorrer por la superficie geodésica. Se seleccionan como puntos candidatos a pertenecer a grietas los que se encuentran en los valles de la superficie geodésica, ya que corresponden a los puntos más oscuros de la imagen. Se conectan dichos puntos cuando las distancias geodésicas son pequeñas y se comprueba que tengan forma de grieta para eliminar las zonas de la imagen que, aunque son oscuras, no son grietas. El algoritmo ha sido probado en 80 imágenes de columnas de hormigón de 200 x 200 píxeles en las que las grietas tienen una anchura entre 1 y 3 píxeles, aunque los autores no ofrecen información del tamaño real de estas grietas. Han comparado los resultados con los obtenidos por una segmentación manual, pero comparando el número de píxeles marcados como defecto en una región en torno a cada píxel, para tratar de limitar el error humano al marcar una grieta sinuosa. La sensibilidad del método varía entre 0,8 y el 0,9 dependiendo del tamaño del entorno en el que se comparan los píxeles marcados.

Los trabajos en los que se detectan grietas en el interior de túneles se diferencian respecto a los anteriores principalmente en la utilización de luz artificial. Algunos trabajos se basan en técnicas de umbralización, como [Miy07], que comparan el valor de cada píxel con la media de su columna. A este grupo también pertenece el trabajo de Fujita et al., [Fuj11] descrito anteriormente. El principal inconveniente de estas técnicas es cómo elegir el valor del umbral. Normalmente están basados en conocimiento previo, pero en esos casos no suelen ser métodos generalizables y no suelen funcionar adecuadamente al ser aplicados en túneles reales, donde pueden aparecer sombras y humedades.

Otras técnicas para la detección de grietas en túneles están basadas en modelos, como la descrita anteriormente en [Yam10] que utilizaba el concepto de percolación, o como la presentada en [Paa06], donde se asume que las grietas están formadas por segmentos de líneas unidas entre sí. En este caso, la resolución de las imágenes es de

2,5mm por píxel y el algoritmo es considerado semiautomático ya que necesita la intervención humana para elegir las semillas de cada grieta.

También existen técnicas basadas en el reconocimiento de patrones, como la utilizada en [Liu02], donde se utiliza el algoritmo SVM (*Support Vector Machine*) para determinar si las grietas preseleccionadas realmente son grietas. En [Zha14] se propone un sistema de inspección de túneles de hormigón compuesto por 9 cámaras lineales CMOS de color que se transforman a imágenes de grises para su procesamiento. A las imágenes de grises, primero se les aplica un filtro de suavizado para eliminar el ruido y a continuación el filtro *top-hat* para realzar las grietas, ya que se supone que son más oscuras que el resto de la imagen. La imagen filtrada es segmentada calculando el umbral de segmentación para cada imagen en función de la distribución de los niveles de gris de la imagen filtrada. Los píxeles contiguos son agrupados y se calculan sus características de forma y nivel de gris para determinar cuáles son las agrupaciones que realmente corresponden a grietas y cuáles no. Para ello se aplican 4 técnicas de reconocimiento de patrones, dos redes neuronales, el algoritmo SVM (*Support Vector Machine*) y el algoritmo de vecinos cercanos. La exactitud de estos algoritmos es aproximadamente del 90%, calculada con 200 agrupaciones de píxeles candidatas a ser grietas de las que solo 64 eran realmente grietas. Las imágenes fueron adquiridas en un túnel de metro con 9 cámaras lineales de 12.288 píxeles que cubren 270° del túnel, pero no se ofrece el diámetro del túnel para poder conocer su resolución.

2.6 Conclusiones

En este capítulo se han descrito las principales características de las superficies continuas, cuáles son los principales retos para su inspección automática y cómo ha sido abordada en otros trabajos. Estas superficies, presentes principalmente en procesos industriales y en infraestructuras lineales, tienen un ancho acotado, una longitud indeterminada y deben ser inspeccionadas en movimiento. La velocidad de inspección es crítica, tanto en las aplicaciones industriales donde se debe inspeccionar a la velocidad del proceso de producción, como en las infraestructuras lineales donde se debe circular a la velocidad de la vía para evitar cortes en la circulación y poner en peligro tanto a usuarios como a operarios. Por tanto, el primer reto que se debe abordar al diseñar un sistema de inspección superficial continua es plantear **una arquitectura eficiente para la adquisición y el tratamiento de la información** obtenida. En el Capítulo 3 se describen las principales etapas de la inspección de superficies continuas, que comienza con la adquisición de las imágenes utilizando cámaras lineales, opción más recomendable que las cámaras matriciales ya que es más

sencillo iluminar de forma homogénea una línea que un área, y continúa con el procesamiento de las imágenes.

El procesamiento de las imágenes en un sistema de inspección superficial tiene por objetivo detectar anomalías en la superficie y clasificarlas por tipos. Estos defectos se caracterizan por alejarse de una homogeneidad esperada, aunque en muchas ocasiones la zona sin defectos presenta una gran variabilidad y su contraste respecto a las zonas con deterioros es muy pequeño. En estos casos las técnicas de segmentación más habituales no obtienen el efecto deseado, por lo que es preciso abordar el problema desde otro enfoque. Estos defectos también presentan una gran variabilidad entre defectos de un mismo tipo, por lo que su clasificación tampoco resulta evidente. En la mayoría de los trabajos analizados, la clasificación utiliza los resultados de una fase de detección previa en la que se han discriminado solo los defectos objeto de estudio. En una aplicación real, debido a las características de estas superficies, la fase de detección cataloga algunas zonas de la imagen como defectuosas sin serlo realmente, a las que se denomina pseudodefectos. Estos deben ser contemplados a la hora de realizar la clasificación. Por todo ello, en el Capítulo 4 se van a describir algunas técnicas para la detección y clasificación de defectos superficiales y cómo evaluar objetivamente los resultados obtenidos, mientras que en el Capítulo 5 se va a proponer un **nuevo enfoque para afrontar la detección y la clasificación de defectos** analizando las anomalías en el dominio de la frecuencia espacial presentes en las imágenes.

3 ARQUITECTURA PARA LA INSPECCIÓN DE SUPERFICIES CONTINUAS

3.1 Etapas para la inspección visual de superficies continuas

En la bibliografía se encuentran múltiples definiciones de las etapas de un sistema de visión por computador. El número de etapas en que se separan las distintas acciones que debe realizar un sistema de inspección visual automático varía de un autor a otro. En algunas ocasiones no se contemplan todas las acciones necesarias y en otras no son lo suficientemente generalistas. Por ejemplo, en [Mal03] se describen 4 etapas para un sistema de inspección industrial: Adquisición de la imagen, procesado de la imagen, extracción de características y toma de decisiones, mientras que en [Piu05] se describen 5 etapas: formación de la imagen, preprocesamiento, segmentación, extracción de características y clasificación.

En este trabajo se considera que la arquitectura de un sistema de inspección visual automático consta de tres grandes bloques o etapas, abordándose en cada una de ellas la totalidad de las acciones que se requieren. En primer lugar, se debe realizar la adquisición de la imagen, a continuación, la imagen debe procesarse y finalmente, se debe utilizar la información obtenida de forma que ésta sea útil para el usuario final de la aplicación. En la **Figura 17** se encuentra un esquema en el que se muestran estos tres grandes bloques o etapas y las principales acciones que se abordan en cada uno de ellos.

La inspección visual automática de superficies continuas tiene ciertas peculiaridades que serán consideradas en la definición de la arquitectura. La principal viene determinada por la definición de superficie continua que, como se describió en la sección 2.1, es aquella que tiene un ancho determinado y una longitud indeterminada, desconocida y mucho mayor que su ancho. Para inspeccionar este tipo de superficies, aunque sería posible realizarlo con cámaras matriciales, es recomendable la utilización de cámaras lineales. La adquisición de imágenes con cámaras lineales requiere la aplicación de elementos y técnicas específicas para este tipo de cámara que se tendrán en consideración en esta sección al definir la arquitectura del sistema de inspección visual.

Otra de las características de la inspección de estas superficies, que también se describió en la sección 2.1, es que el objetivo de la inspección de su superficie es la

búsqueda de zonas que se alejen de la homogeneidad esperada, lo que supondrá la utilización de técnicas de procesamiento de imágenes específicas para este objetivo.

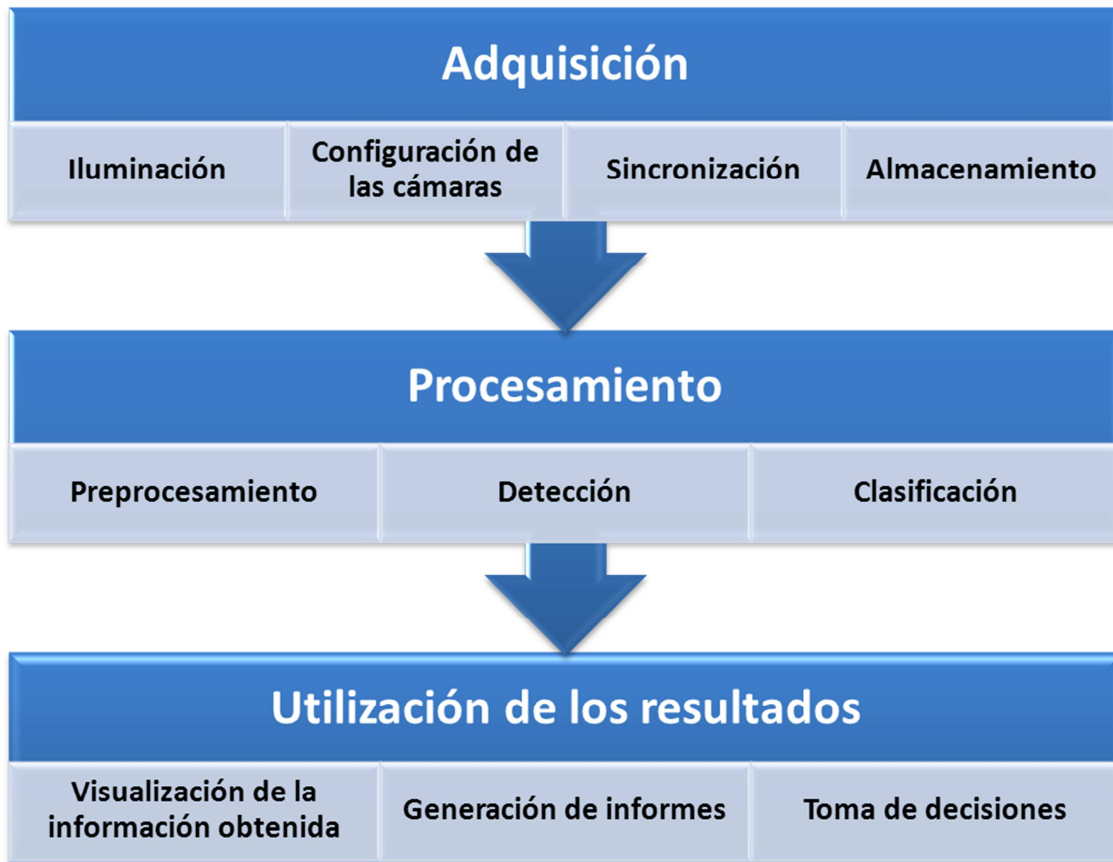


Figura 17. Etapas para la inspección visual automática

A continuación se describen los tres grandes bloques o etapas establecidos para la inspección de superficies continuas, detallando cada una de las acciones necesarias que se llevan a cabo en cada uno de ellos.

3.2 Adquisición

La inspección de superficies continuas, aunque es posible realizarla con cámaras matriciales, es recomendable llevarla a cabo mediante cámaras lineales. Una cámara lineal construye la imagen capturándola línea a línea. La cámara consta de un sensor lineal, que suele oscilar entre los 512 y los 12.000 píxeles, que para realizar la captura del objeto, éste debe desplazarse con respecto a la cámara, o bien la cámara debe desplazarse respecto al objeto.

Esta tecnología fue desarrollada para aplicaciones de inspección de materiales fabricados en continuo, como papel, tela, planchas metálicas, etc. Sin embargo, en la actualidad se está imponiendo en muchos otros procesos de inspección que requieren alta resolución y velocidad a un precio competitivo.

La adquisición de imágenes requiere la utilización de una iluminación adecuada y la configuración de la cámara acorde a la iluminación utilizada y los objetivos de inspección buscados. Cuando se utilizan cámaras lineales, la iluminación debe conseguir iluminar de la forma más homogénea posible una única línea de la superficie a inspeccionar, que abarque todo su ancho. Es necesario sincronizar con precisión la adquisición de cada línea de la imagen con el movimiento relativo de la superficie inspeccionada respecto a la cámara, para lo que se suele utilizar un codificador incremental. Finalmente, se debe almacenar y utilizar la gran cantidad de información adquirida.

3.2.1 Iluminación

La principal ventaja de utilizar cámaras lineales es que solo es necesario iluminar una delgada línea del objeto a inspeccionar. Gracias a ello, se reduce drásticamente la cantidad de energía necesaria y es más sencillo conseguir iluminar de forma homogénea la zona a inspeccionar. La iluminación de una línea se realiza fundamentalmente mediante fuentes de iluminación led que concentran la luz mediante ópticas en una línea del ancho deseada. Para conseguirlo, el sistema de iluminación debe encontrarse a la distancia adecuada del objeto a inspeccionar y se debe alinear con el sensor de la cámara con gran precisión. También son muy efectivas las fuentes de iluminación láser, con la ventaja de que concentran la luz a cualquier distancia, aunque su coste es mayor y pueden suponer un peligro si el rayo incide directamente sobre el ojo de las personas.

Aunque la iluminación lineal puede resultar aproximadamente homogénea a lo largo de la línea, siempre se van a producir pequeñas diferencias en la intensidad de la imagen a lo largo de la misma, debido, fundamentalmente, a que la distancia que recorre la iluminación desde la fuente hasta el sensor no es exactamente la misma en todo el ancho de la imagen, el ángulo de incidencia es diferente a lo largo de la línea iluminada y a que los sensores de la cámara no reaccionan de la misma forma ante el mismo estímulo de iluminación. Es por ello, que resulta recomendable que el primer paso de procesamiento de la imagen sea la homogeneización de la iluminación. Este procedimiento consiste en modificar los valores de la imagen en cada píxel de la cámara para que la respuesta de la cámara sea la misma en toda la longitud de la línea del sensor al capturar la imagen de un objeto de color homogéneo.

3.2.2 Configuración

La inspección superficial tiene por objetivo detectar defectos que se muestran en la imagen con una intensidad por encima o por debajo que la que presenta la superficie libre de defectos. Por tanto, es conveniente configurar la cámara de tal manera que el valor medio de una imagen libre de defectos sea el valor central del rango de valores permitido. Es decir, si el rango de valores de una cámara se encuentra entre 0 y 255, la imagen de la superficie libre de defectos debería tener un valor en torno a 128. De esta manera, tanto los defectos que son más oscuros que la superficie sin defectos, como los que son más brillantes, son observados correctamente en la imagen.

Para conseguir este objetivo, las cámaras disponen de tres parámetros configurables: el tiempo de exposición, la apertura del diafragma y el valor de la sensibilidad. Estos parámetros, que determinan el nivel de intensidad de cada píxel de la imagen, también afectan a otros factores como la nitidez o el nivel de ruido. A continuación, se va a describir cómo afectan estos parámetros a la calidad de las imágenes, particularizando para el caso de la adquisición de imágenes de superficies continuas, y cómo se debe proceder para configurar correctamente los parámetros de adquisición de una cámara.

El **tiempo de exposición** es el intervalo de tiempo en que el sensor de la cámara está captando energía lumínica. En una cámara lineal está limitado por el tiempo que transcurre entre la señal de captura de dos líneas consecutivas, que a su vez depende de la velocidad a la que se mueve la superficie a inspeccionar respecto a la cámara. Si una señal de captura de línea llegara a la cámara antes de finalizar el tiempo de exposición, ésta sería ignorada y la imagen tendría una resolución en la dirección de avance inferior a la esperada. En los transitorios de algunos procesos productivos, principalmente arranques y paradas, pueden producirse velocidades instantáneas muy elevadas que provocan este efecto. Por otro lado, la inspección de superficies continuas se realiza en movimiento, por lo que durante el periodo de tiempo que el obturador se encuentra abierto, la superficie se desplaza. Este desplazamiento está acotado y siempre es inferior a la resolución de la imagen. Por tanto, cuanto mayor sea el tiempo de exposición, mayor es la energía captada, pero menor es la nitidez de la imagen.

La **apertura del diafragma** determina la cantidad de luz que atraviesa la óptica de la cámara por unidad de tiempo. Cuanto mayor sea la apertura, mayor es la cantidad de luz que llega al sensor en un mismo intervalo de tiempo, y, por tanto, es necesario menor tiempo de exposición para conseguir los mismos niveles de exposición en la imagen. Sin embargo, al aumentar la apertura de la óptica se disminuye la profundidad de campo, es decir, el rango de distancias en las que el objeto se encuentra enfocado. Si la superficie inspeccionada se encuentra a una distancia que no

varía con el tiempo, se puede aumentar la apertura de la óptica para reducir el tiempo de exposición necesario. Este es el caso de muchas aplicaciones industriales en los que la superficie inspeccionada y la cámara se encuentran siempre en la misma posición. Sin embargo, si la distancia entre la cámara y la superficie inspeccionada varía a lo largo del tiempo, se debe reducir la apertura del diafragma para mantener la imagen enfocada en todo momento. Este es el caso, por ejemplo, de la inspección de carreteras en el que las irregularidades del terreno provocan que la suspensión del vehículo aleje y acerque la cámara de la superficie de la carretera.

Por último, se encuentra la **sensibilidad de la cámara** o ganancia. Este parámetro determina como de sensible es el sensor de la cámara ante un mismo estímulo de luz. Es un valor que se puede ajustar para conseguir el valor de intensidad en la imagen deseado, pero tiene la importante contrapartida de que cuanto mayor sea el valor de ganancia mayor es el nivel de ruido en la imagen. Por tanto, es recomendable establecer ganancias bajas para reducir el nivel de ruido de la imagen. Por otro lado, el nivel de ruido para una determinada sensibilidad es un factor clave a la hora de valorar la calidad de una cámara.

Por tanto, el primer paso a la hora de configurar los parámetros de adquisición es determinar cuál es la máxima apertura del diafragma que permite obtener una imagen enfocada en el caso más desfavorable. También es necesario calcular cuál es el tiempo de exposición máximo que se puede alcanzar para no perder ninguna línea de la imagen. Por último, se selecciona la combinación de tiempo de exposición – ganancia que proporcione una imagen nítida y con un nivel de ruido aceptable. Si el resultado obtenido no es satisfactorio, es imprescindible aumentar la potencia de la fuente de iluminación.

La configuración del tiempo de exposición y de la ganancia es conveniente realizarla cada vez que se comienza la inspección de un nuevo tipo de superficie. Esto se consigue mediante un proceso recursivo en el que se adquieren imágenes de una zona de la superficie a inspeccionar y se ajustan estos parámetros hasta conseguir el efecto deseado.

3.2.3 Sincronización

Para realizar la adquisición de imágenes con cámaras lineales, el desplazamiento relativo entre la cámara y el objeto debe ser conocido. Para ello existen dos opciones, o bien la velocidad es constante y conocida, o bien, se dispone de un elemento sensor para disparar la adquisición de cada línea en el instante oportuno. No es habitual que la velocidad sea constante y conocida, y si por cualquier circunstancia la velocidad cambiara, el sistema dejaría de funcionar correctamente. Por otro lado, la segunda

opción permite trabajar en los transitorios, es decir, cuando arranca o se detiene la línea de producción o la plataforma de inspección. Por tanto, utilizar un elemento sensor para disparar la adquisición de cada línea es, sin duda, la opción más aconsejable de las dos.

El elemento sensor que permite disparar la adquisición de cada una de las líneas de la imagen es un codificador incremental, el cual genera un pulso cada vez que el objeto (o la cámara) se desplaza una distancia prefijada, por lo que determina la resolución de la imagen en la dirección de avance. Esta resolución suele hacerse coincidir con la resolución transversal, aunque no es imprescindible. También es posible utilizar un *encoder* de cuadratura, el cual genera dos señales desfasadas, lo que permite saber en qué sentido se produce el desplazamiento y actuar en consecuencia, aunque no suele ser necesario su utilización en la mayoría de las aplicaciones de inspección de superficies continuas.

Cuando el sistema de inspección utiliza varios dispositivos (varias cámaras lineales o una o más cámara junto a otro tipo de dispositivos como GPS, cámaras de entorno, etc.), la forma de sincronizar la adquisición de todos ellos consiste en comenzar la toma de datos con la señal del codificador incremental apagada. Cuando todos los dispositivos se encuentran listos para adquirir, se activa la señal del codificador incremental. De esta forma, se consigue que el primer pulso que recibe cada dispositivo sea el mismo para todos ellos. Para detener la adquisición se realiza el proceso inverso. En el momento de finalizar la inspección, se corta la señal del codificador incremental, para a continuación, detener la adquisición de los dispositivos. De esta manera, el último pulso que recibe cada dispositivo es el mismo para todos ellos.

3.2.4 Almacenamiento

El proceso de adquisición finaliza con el almacenamiento de la información. Ésta puede ser almacenada en la memoria principal del ordenador (memoria RAM) para su procesamiento inmediato o en la memoria secundaria (disco duro) para su posterior procesamiento o consulta de la información almacenada. La primera opción se utiliza si se requiere disponer de los resultados de la inspección de forma inmediata y el almacenamiento del total de los datos brutos adquiridos no es necesario. La segunda opción se utiliza cuando el procesamiento de la información no requiere realizarse de forma inmediata o cuando es necesario almacenar toda la información adquirida para su posterior consulta. En muchas ocasiones se realizan los dos tipos de almacenamiento, de tal forma que se procesan los datos de forma inmediata y se almacenan los datos brutos para su posterior consulta.

Cuando la información se almacena de forma temporal en la memoria principal para su procesamiento inmediato, es necesario que éste se realice a gran velocidad, de tal manera que la velocidad de procesamiento sea igual o mayor que la de adquisición. Para ello se requieren equipos con alta potencia de cálculo y suele ser habitual que se realice de forma distribuida, es decir, la información se procesa en distintos ordenadores para alcanzar la velocidad necesaria. La arquitectura del sistema se complica, ya que los distintos ordenadores encargados del procesamiento deben compartir información y comunicarse entre ellos. En esta tesis se propone la inspección en tiempo real de chapa laminada utilizando una arquitectura para el procesamiento distribuido en tiempo real. En la sección 6.2 se describe con detalle su implementación.

Para el almacenamiento de la información en la memoria secundaria se requieren equipos con gran capacidad de almacenamiento. La velocidad de procesamiento en este caso no es crítica, pero sí lo es la forma en que se almacena la información. Para alcanzar altas velocidades de almacenamiento, los datos adquiridos se transfieren en bloques de gran tamaño a un único archivo, ya que el almacenamiento de pequeños archivos supone una importante reducción de la velocidad de escritura en memoria secundaria. Como el flujo de datos es muy elevado no es posible almacenar toda la información en un único dispositivo, por lo que se distribuye entre varios y, en ocasiones, ubicados en distintos equipos. Para diseñar la arquitectura del sistema es necesario conocer las limitaciones de velocidad de almacenamiento de los equipos empleados. Es preciso conocer la velocidad de escritura y la velocidad de transferencia del bus de datos del microprocesador para diseñar de forma eficiente la arquitectura para el almacenamiento de los datos.

3.3 Procesamiento

En esta etapa los datos adquiridos son procesados de forma automática para su interpretación. El procesamiento puede realizarse al mismo tiempo que se capturan los datos o posteriormente, dependiendo de las necesidades de la aplicación. Cuando el proceso se realiza inmediatamente después de la adquisición de los datos, los algoritmos empleados deben ser los suficientemente rápidos para que la velocidad de adquisición no supere a la de procesamiento. El procesamiento de las imágenes se puede dividir en tres grandes etapas, que se describen a continuación.

3.3.1 Preprocesamiento

La primera etapa del procesamiento de una imagen es el preprocesamiento, que tiene por objetivo mejorar la imagen para que las etapas posteriores de detección y clasificación funcionen más eficientemente. En esta etapa es habitual la aplicación de

filtros para la eliminación del ruido de la imagen o para el realce de alguna de sus características. En el caso de inspección de superficies continuas mediante cámaras lineales, la principal operación que se debe realizar sobre la imagen es su homogeneización. Los píxeles de cada columna de la imagen han sido adquiridos con el mismo elemento sensor y en condiciones de iluminación muy similares. Sin embargo, cada elemento sensor de la cámara puede reaccionar de forma diferente, y las condiciones de iluminación de la zona capturada por cada pixel pueden ser diferentes. Para conseguir una imagen homogénea se deben corregir estas diferencias, de tal forma que todos los píxeles de la imagen presenten el mismo nivel de gris ante una misma superficie.

3.3.2 Detección

La etapa de detección en la inspección de superficies continuas consiste en determinar qué zonas de la imagen contienen defectos y cuáles no. La detección puede realizarse a nivel de pixel o dividiendo la imagen en pequeñas porciones. Las técnicas empleadas para la detección pueden ser muy variadas, desde simples algoritmos de segmentación en los que los píxeles cuyos niveles de gris se encuentren fuera de un rango determinado son considerados defectos, hasta complejos algoritmos en los que se evalúan las interrelaciones con los píxeles vecinos para determinar si se trata de zonas defectuosas o no. A continuación, los píxeles o zonas consideradas defectos se agrupan para formar una única entidad o defecto para su posterior clasificación. Los algoritmos de detección de defectos pueden configurarse para que su detección sea más o menos conservadora, es decir, el algoritmo de detección puede detectar todos los defectos presentes pero asignar zonas defectuosas que realmente no lo son, o marcar como defectos solo las zonas que realmente tienen defectos pero habrá defectos que no sean considerados como tal. Lo habitual es buscar un punto intermedio entre estos dos casos extremos. En la etapa de clasificación se debe tener la presencia de defectos que realmente no son defectos, también conocidos como pseudodefectos.

3.3.3 Clasificación

La etapa de clasificación consiste en asignar cada defecto a un tipo de defecto. Se debe tener en cuenta que la fase de detección puede haber etiquetado como defecto zonas que realmente no lo son o que no pertenecen a ninguno de los tipos esperados. La clasificación analiza un vector de características de cada defecto para asignarlo a un tipo. Existe una gran diversidad de técnicas de clasificación, algunas de las cuales se describen con más detalle en la sección 4.5. La elección del algoritmo de clasificación y de las características utilizadas para su clasificación determina el grado de acierto del clasificador. Estas características miden la forma y la distribución del nivel de gris de la

imagen en el interior del defecto. La distribución del nivel de gris de la imagen es medida por simples estadísticos como la media, la desviación estándar, el valor máximo, etc., o mediante el análisis de la imagen filtrada por filtros que realcen alguna característica que defina a los defectos, como es el caso de la utilización de filtros de Gabor que va a ser fruto de un profundo estudio en esta tesis.

3.4 Utilización de los resultados

La última fase de la inspección visual es la utilización de los resultados obtenidos. El resultado de la inspección puede ser la ejecución de acciones automáticas que actúen sobre el producto inspeccionado, como en el caso de la inspección de chapa laminada que se presenta en esta tesis, en el que ante la presencia de algunos tipos de defectos se detiene el procesado de la bobina, o informar de la calidad del producto inspeccionado para ayudar a la toma de decisiones a los operarios, como es el caso de la inspección superficial de la carretera, en el que dependiendo del estado de una carretera se acometen determinadas obras de rehabilitación.

3.4.1 Visualización de la información obtenida

Uno de los principales usos que se le puede dar a la información obtenida en una inspección visual es la visualización de las imágenes adquiridas y el resultado del procesamiento en cualquier momento posterior a su inspección. Esta información puede ser incorporada a sistemas de trazabilidad de tal manera que se puede conocer el estado superficial de un producto en un momento dado. En el caso de inspección de carreteras o túneles, el responsable de la infraestructura puede visualizar su estado sin necesidad de desplazarse, comprobar su evolución en sucesivas campañas de inspección y determinar las acciones a tomar para la conservación de la infraestructura. En el caso de productos industriales, permite conocer la calidad del producto en distintos puntos de la fase de producción, permitiendo conocer la causa de los defectos detectados y actuar para mejorar el proceso productivo, conocer la calidad del producto antes de su procesado para poder reclamar al proveedor, o conocer la calidad del producto al final de su procesado para su utilización más apropiada o su desecho si no cumple con los estándares de calidad deseados.

El entorno de visualización debe permitir la navegación por toda la superficie inspeccionada. Al tratarse de una superficie continua, se dispone de una imagen con un ancho fijo y una longitud mucho mayor y distinta para cada superficie inspeccionada. Para navegar por la superficie es necesario poder desplazarse por toda la imagen de forma continua y acceder a la posición deseada. La utilización de cámaras lineales simplifica la formación de la imagen global respecto a los sistemas que utilizan

cámaras matriciales que requieren la utilización de complejas técnicas de *stitching* para juntar las imágenes, [Jah11].

El entorno de visualización también debe poder mostrar otra información adicional de forma sincronizada con la imagen principal. En esta tesis, se han desarrollado los entornos de visualización adecuados para los tres sistemas de inspección implementados. Los tres tienen elementos comunes para la visualización y funcionalidades específicas para cada una de las aplicaciones. En el Capítulo 5, se describen los entornos de visualización implementados.

3.4.2 Generación de informes

Los resultados de la inspección deberían ser mostrados de forma que se adecue a las expectativas de las personas que van a utilizar esa información. Aunque el sistema de visualización de la información ofrece información completa del estado de la superficie inspeccionada, es conveniente disponer de informes que resuman con mayor o menor detalle el estado de la superficie inspeccionada. Estos informes pueden ser utilizados sin la necesidad de tener acceso a toda la información almacenada, utilizando *software* comercial o en formato impreso. En esta última etapa de la inspección, la información debe ser generada de forma que sea útil para la persona que va a utilizar esa información, por lo que la generación de estos informes debe ser realizada atendiendo sus requisitos.

3.4.3 Toma de decisiones

El resultado del procesamiento de la información puede ser utilizado para la toma de decisiones. Si el procesamiento de las imágenes se realiza al mismo tiempo en el que se toman las imágenes, la información extraída puede interactuar con la línea de producción, deteniendo la producción ante la presencia de ciertos defectos críticos, tanto por su cantidad como por su gravedad. Los resultados de la inspección también pueden ayudar a la toma de decisiones sobre las actuaciones que deben llevarse para reparar una superficie o para modificar un proceso productivo que este provocando un determinado defecto. Para ello, la información se muestra al usuario de forma sencilla, resumida y adecuada a sus necesidades, pero facilitándole el acceso a la información detallada cuando así lo requiere.

3.5 Conclusiones

En este capítulo se han descrito las etapas de un sistema de inspección de superficies continuas. Éstas son generalizables a otro tipo de inspecciones visuales,

pero el objetivo de este capítulo es detallar las particularidades que afectan a la inspección en continuo.

En primer lugar, se aborda la **adquisición de las imágenes**. El sistema de adquisición de imágenes recomendado para la inspección de superficies continuas es la cámara lineal, dispositivo que captura la imagen línea a línea mientras se desplaza el objeto respecto a la cámara. La captura tiene que sincronizarse con dicho desplazamiento, para lo cual se genera una señal de disparo de la adquisición cada vez que se produce un desplazamiento establecido, el cual determina la resolución de la imagen en la dirección de avance. La iluminación es un elemento clave en un sistema de inspección visual, ya que es la principal responsable de realzar los defectos presentes en la superficie. Una de las ventajas de utilizar cámaras lineales es que solo es necesario iluminar correctamente una línea, lo que reduce la potencia de la iluminación requerida y facilita su distribución homogénea sobre la superficie inspeccionada. Para conseguir imágenes nítidas deben configurarse los parámetros de la cámara, estableciendo tiempos de exposición muy breves, pero manteniendo la imagen enfocada en todo momento y los niveles de ruido bajos. Estas inspecciones generan un elevado flujo de datos que deben ser almacenados y procesados. En el Capítulo 6 se detalla cómo se han aplicado estas técnicas a tres sistemas de inspección de superficies continuas implantados en situaciones reales de funcionamiento.

La segunda etapa es el **procesamiento de las imágenes**, que se ha dividido en tres pasos. En el primero, se realiza el preprocesamiento de las imágenes, cuyo objetivo más destacado en la inspección de superficies continuas es homogeneizar la imagen a lo largo del ancho de una línea. A continuación, se realiza la detección, separando los píxeles de la imagen que contienen defectos de los que no. Cuando la diferencia entre los niveles de gris de la superficie sin defectos respecto a los defectos es pequeña, como ocurre en las aplicaciones que se proponen en esta tesis, es recomendable que esta etapa se realice de forma conservadora, de tal forma que el mayor porcentaje posible de defectos alcancen la fase de clasificación. Los píxeles catalogados como defectuosos se agrupan para realizar la clasificación, cuyo objetivo es separar los defectos por tipos. A la etapa de clasificación pueden llegar agrupaciones de píxeles que no corresponden con ninguno de los defectos esperados, hecho que se debe tener presente y que no suele ser considerado por la mayoría de los autores. Existe también la posibilidad de realizar la detección y la clasificación en un único paso, diseñando clasificadores específicos para cada tipo de defecto. Distintas alternativas son presentadas en el Capítulo 4 para abordar la detección y la clasificación, aunque todas ellas precisan la medición de alguna variable que describa la imagen, aspecto que será tratado en el Capítulo 5.

Para finalizar, los resultados de procesamiento de la imagen son utilizados para **mostrar visualmente la información extraída, generar informes y tomar decisiones** sobre el producto inspeccionado o el proceso de producción. La visualización de imágenes de superficies continuas tiene la peculiaridad de tener que mostrar una imagen con un ancho acotado pero una longitud mucho mayor y distinta para cada superficie inspeccionada. Cada uno de los sistemas de inspección desarrollados en esta tesis dispone de un *software* específico, cuyo núcleo es común a todos ellos, que aborda los aspectos descritos. Una descripción más detallada de la funcionalidad de este *software* se encuentra en el Capítulo 6.

4 TÉCNICAS DE DETECCIÓN Y CLASIFICACIÓN DE DEFECTOS

4.1 Introducción

En este capítulo se analizan las técnicas utilizadas para la detección y la clasificación de defectos superficiales. La detección de defectos puede considerarse un caso especial de clasificación en la que se separan los píxeles de la imagen en dos grupos, el grupo de píxeles con defectos y el grupo de píxeles libre de defectos. Al tener dos posibles grupos la detección se puede considerar un clasificador binario.

Cuando se habla de clasificación de defectos se presupone que en un paso anterior del procesamiento de la imagen en el que se han detectado cuales son los píxeles con defectos y estos se han agrupado siguiendo criterios de proximidad y/o similitud. Entonces, la clasificación de estos píxeles se realiza utilizando características de esas agrupaciones de píxeles y el resultado esperado de esta clasificación es asignar cada uno de estos grupos de píxeles a un tipo de defecto.

A la hora de evaluar un clasificador es ampliamente utilizada la técnica de considerar a un clasificador como un conjunto de clasificadores binarios. De esta manera se analizan tantos clasificadores binarios como tipos de defecto y, para cada clasificador binario, se considera que está bien clasificado si el clasificador asigna el defecto al grupo correcto y estará mal clasificado cuando es asignado a cualquier otro grupo.

Es por ello que resulta de especial interés analizar cuáles son las técnicas para evaluar el comportamiento de los clasificadores binarios y poder así analizar el rendimiento de los clasificadores.

Teniendo en cuenta estas consideraciones, esta sección se divide en 5 secciones. En primer lugar, se describe cómo se va a evaluar el desempeño de las técnicas de detección y clasificación que se van a implementar. Después se describen algunos métodos para el cálculo de los umbrales de clasificadores binarios. A continuación se analiza el algoritmo de entrenamiento AdaBoost, que permite combinar distintos clasificadores binarios para obtener un único clasificador que mejora el comportamiento de un único clasificador binario. En la siguiente sección se describen los clasificadores no binarios basados en redes neuronales, vecinos más cercanos y el clasificador de Naïve Bayes, utilizados más adelante en esta tesis. Finalmente se

analizan otras técnicas para combinar varios clasificadores no binarios y mejorar los resultados globales.

4.2 Evaluación del desempeño de la detección y clasificación de defectos superficiales

Cuando se aplica un algoritmo para la detección o para la clasificación de defectos se debería plantear en primer lugar cómo evaluar su desempeño, es decir, cómo medir de una forma objetiva de qué manera se ha comportado el algoritmo y poder así compararlo con otros métodos y elegir el que mejores resultados obtiene. En principio puede parecer algo trivial, sin embargo, existen diversas formas de abordar el problema que pueden dar resultados distintos.

En primer lugar, se debe establecer qué es la detección de defectos superficiales y qué es la clasificación de defectos superficiales.

La detección de defectos superficiales consiste en determinar las zonas de la imagen que tienen un cambio significativo con respecto a las características habituales de la superficie.

La clasificación de defectos superficiales consiste en asignar cada una de las zonas de la imagen a un grupo o tipo de defecto. Esta clasificación suele partir de una detección previa, en la que se han establecido que píxeles de la imagen son defectos y cuáles no y estos se han agrupado siguiendo criterios de proximidad y/o similitud.

Un clasificador binario es aquel cuyo resultado solo puede tener dos posibles valores. La detección de defectos puede considerarse un caso especial de clasificación binaria en la que se clasifican los píxeles de la imagen en dos grupos, el grupo de píxeles con defectos y el grupo de píxeles libre de defectos. Al tener dos posibles grupos la detección se puede considerar un clasificador binario.

Un clasificador con múltiples salidas puede ser considerado un conjunto de clasificadores binarios. Cada defecto tendrá asociado un clasificador binario con dos posibles salidas, o bien la imagen contiene el defecto o no lo contiene.

De esta manera se analizan tantos clasificadores binarios como tipos de defecto y, para cada clasificador binario, se considera que está bien clasificado si el clasificador asigna el defecto al grupo correcto y está mal clasificado cuando se asigna a cualquier otro grupo.

La selección de la técnica a utilizar para la evaluación del desempeño dependerá de la aplicación concreta que se aborde. El problema se divide en dos casos diferenciados: clasificadores binarios y clasificadores no binarios.

4.2.1 Clasificadores binarios

Un clasificador binario divide un conjunto de muestras en dos grupos diferentes. Habitualmente estos clasificadores buscan determinar la presencia o ausencia de un determinado objeto, por lo que cuando este objeto aparece en la imagen, a la muestra se la llama positivo y cuando no negativo. Las muestras pueden ser imágenes completas, zonas de las imágenes, o incluso puede aplicarse la clasificación a nivel de píxeles.

Por tanto, este tipo de clasificadores pueden ser aplicados para detectar la presencia o ausencia de un tipo de defecto superficial. Cuando la característica que se mide es una variable continua, como son las características que se pueden extraer de una imagen filtrada por un filtro de Gabor que se describen en la sección 5, la selección del valor umbral que determina el límite entre la presencia o ausencia de un tipo de defecto tiene una gran influencia en la precisión del sistema. La elección del umbral puede ser muy sencilla, pero es deseable que sea elegida con el objetivo de optimizar el desempeño del sistema.

Un clasificador binario, j , queda expresado con la siguiente ecuación:

$$h_j(x) = \begin{cases} 1 & \text{si } p_j f_j(x) < p_j \lambda_j \\ 0 & \text{en cualquier otro caso} \end{cases} \quad \text{Ec. 1}$$

siendo $h_j(x)$ el clasificador, $f_j(x)$ la característica que se mide, λ_j el valor umbral y p_j es un valor de paridad que indica la dirección de la inecuación. La selección del umbral es una tarea que tiene un alto impacto en el desempeño del clasificador que puede ser llevada a cabo atendiendo diversos métodos, tal y como se explicará en el apartado 4.3.

Para comparar los resultados que obtiene un clasificador al utilizar distintos umbrales se utilizan medidas como la sensibilidad, especificidad, exactitud y precisión, así como la exactitud ponderada. Para calcular estas magnitudes, utilizadas habitualmente en la evaluación de clasificadores, se utiliza la matriz de confusión que aparece en la siguiente tabla.

4. Técnicas de detección y clasificación de defectos

Tabla 2. Matriz de confusión. Relación entre los clasificadores y la clasificación realizada por un experto

		Clasificación del experto	
		Positivos	Negativos
Clasificación del clasificador	Positivos	Verdaderos positivos (TP)	Falsos Positivos (FP)
	Negativos	Falsos negativos (FN)	Verdaderos negativos (TN)

A partir de esta tabla se calculan los valores de sensibilidad, especificidad, exactitud y precisión aplicando las siguientes ecuaciones.

$$\begin{aligned} \text{Sensibilidad} &= \frac{TP}{TP+FN} & \text{Especificidad} &= \frac{TN}{TN+FP} \\ \text{Exactitud} &= \frac{TP+TN}{TP+TN+FP+FN} & \text{Precisión} &= \frac{TN}{TN+FN} \end{aligned} \quad \text{Ec. 2}$$

La sensibilidad, también llamada ratio de acierto o ratio de verdaderos positivos, es la proporción de positivos correctamente clasificados del total de positivos.

La especificidad es la proporción de negativos correctamente clasificados del total de negativos. Su valor complementario (1 – especificidad) es llamado ratio de falsos positivos o ratio de falsas alarmas.

La exactitud es la proporción de muestras correctamente clasificadas del total de muestras.

La precisión es la proporción de positivos correctamente clasificados del total de muestras clasificadas como positivos.

Además, se utiliza la exactitud ponderada (*balanced accuracy*), que es la media entre la sensibilidad y la especificidad, [Bro10]. Esta medida resulta más útil que la exactitud para conjuntos de datos en los que el número de positivos y negativos no está equilibrado, ya que le otorga el mismo peso al conjunto de positivos y al conjunto de negativos en vez de darle el mismo peso a cada una de las muestras.

$$\text{Exactitud ponderada} = \frac{\text{Sensibilidad} + \text{Especificidad}}{2} \quad \text{Ec. 3}$$

Su valor complementario ($1 -$ exactitud ponderada) se le llama error ponderado.

La curva ROC (*Receiver Operating Characteristic*) es una herramienta muy útil para visualizar y evaluar el comportamiento de clasificadores binarios, [Faw06]. Esta curva representa el comportamiento de un clasificador cuando se varía el umbral de clasificación. En ella se representa el ratio de verdaderos positivos (sensibilidad) en el eje de ordenadas, frente al ratio de falsos positivos ($1 -$ especificidad) en el eje de abscisas, para todos los valores posibles del umbral.

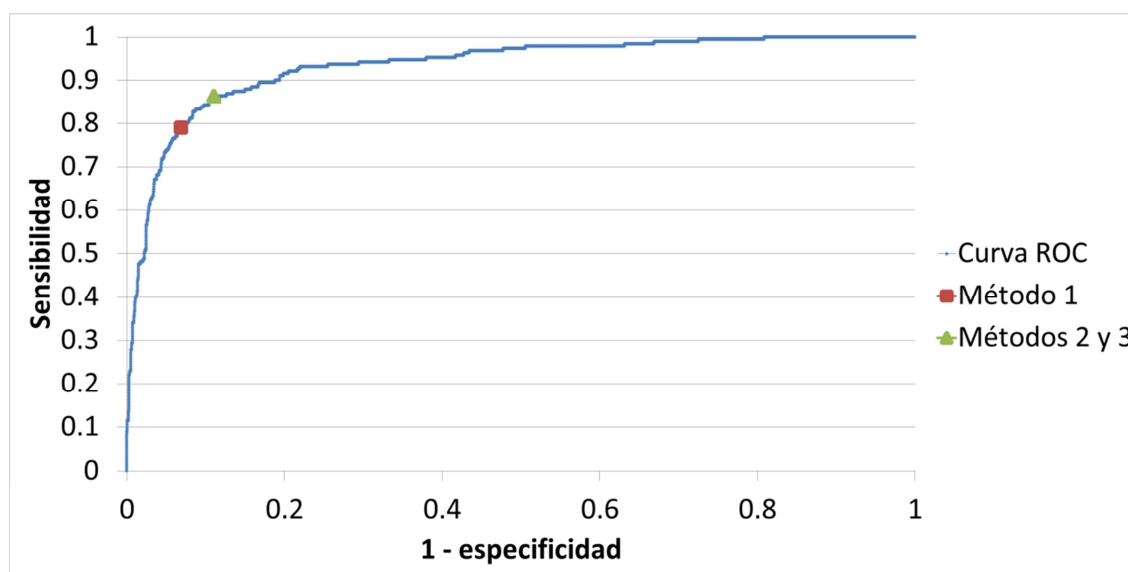


Figura 18. Ejemplo de curva ROC. Esta curva corresponde a uno de los clasificadores utilizados en la sección 7.2 para la detección grietas longitudinales en la superficie de la carretera. Se han marcado los puntos de la recta correspondientes a los umbrales obtenidos al utilizar los métodos de cálculo de umbrales descritos en la sección 4.3

Para generar esta curva de forma precisa, todas las muestras son ordenadas por su valor de salida. A continuación se genera una lista de valores umbrales correspondientes al valor medio de la salida de dos muestras consecutivas. Con cada uno de estos umbrales se calcula la sensibilidad y la especificidad, que se utiliza para generar el gráfico o realizar los cálculos que fueran necesarios.

Esta información no solo sirve para evaluar un clasificador sino que también permite fijar el umbral de un clasificador con el objetivo de alcanzar las especificaciones que se tengan por objetivo, como puede ser alcanzar un valor determinado de especificidad o sensibilidad.

Estas curvas también pueden utilizarse para evaluar un clasificador mediante el cálculo de área bajo la curva ROC, conocido este valor como AUC (*Area Under the ROC Curve*). El valor de esta área puede variar entre 0 y 1, obteniendo el valor 1 el

clasificador perfecto. En la práctica ningún clasificador tendrá una AUC inferior a 0,5, que es el que obtiene un clasificador aleatorio. Aunque este valor puede resumir muy bien el comportamiento de un clasificador, cuando las curvas ROC se cruzan no siempre ofrece el resultado esperado, [Han09], ya que en estas circunstancias el clasificador con menor AUC podría alcanzar mejores resultados en la zona ROC que nos interesa que el clasificador con la AUC superior. Es por ello que en este trabajo no se va a utilizar el análisis AUC.

4.2.2 Clasificadores no binarios

Los clasificadores no binarios son aquellos que ofrecen como salida la clase a la que pertenece una determinada muestra. Las muestras pueden ser imágenes, trozos de imágenes o, como es habitual en la clasificación de defectos superficiales, una zona de la imagen que ha sido previamente marcada como defecto por un método de detección de defectos.

A la hora de evaluar un clasificador no binario se suele utilizar la exactitud global, que consiste en dividir el número de muestras correctamente clasificadas por el número total de muestras.

A pesar de que este criterio de evaluación resulta sencillo en su principio e implementación, resulta poco adecuado por considerar de manera conjunta las precisiones individuales de las clases, por lo tanto, se recomienda que sea acompañado por el análisis de la Matriz de Confusión, tal y como se muestra en la **Tabla 3**. Ésta es ampliamente utilizada para la presentación y el análisis del resultado de una clasificación, por su capacidad de plasmar los conflictos entre las clases. La matriz de confusión, no sólo muestra el porcentaje correcto de clasificación, sino además la fiabilidad para cada una de las clases y las principales confusiones entre ellas.

La diagonal principal de la matriz contiene las muestras que fueron clasificadas correctamente, mientras que el resto de los valores representa las muestras mal clasificadas y cuál fue el error cometido, lo que permite realizar un análisis más riguroso de cómo se comporta el clasificador.

Esta matriz permite analizar el comportamiento del clasificador para cada defecto de forma independiente. De forma análoga a los clasificadores binarios, han sido definidos los verdaderos positivos y negativos y los falsos positivos y negativos:

- Los verdaderos positivos para el defecto i , son el valor contenido en la casilla C_{ii} de la **Tabla 3**.
- Los verdaderos negativos para el defecto i , son la suma de los C_{jk} , para todo j distinto de i y para todo k distinto de i .

- Los falsos positivos para el defecto i , es la suma de los C_{ij} siempre que j sea distinto de i
- Los falsos negativos para el defecto i , es la suma de los C_{ji} siempre que j sea distinto de i

Estos valores permiten evaluar el rendimiento de la clasificación de cada defecto de forma análoga a la efectuada para los clasificadores binarios, calculando la sensibilidad, la especificidad, la exactitud, la precisión y la exactitud ponderada.

Tabla 3. Matriz de confusión de un clasificador no binario

		Clasificación del clasificador				
		Defecto 1	...	Defecto i	...	Defecto n
Clasificación del experto	Defecto 1	C_{11}		C_{1i}		C_{1n}
	...					
	Defecto i	C_{i1}		C_{ii}		C_{in}
	...					
	Defecto n	C_{n1}		C_{ni}		C_{nn}

Aunque se ha considerado que todas las posibles clases son defectos, en muchas ocasiones se considera que una de las posibles salidas del clasificador es la clase "sin defecto". Desde un punto de vista formal puede tratarse como un defecto más. En esta clase suelen estar englobadas todas las muestras que no pertenecen al resto de las clases, bien porque no son defectos o bien porque corresponden a un tipo de defecto que no se tuvo en cuenta en el entrenamiento.

4.3 Cálculo de umbrales para clasificadores binarios

Como se describió en la sección 4.2.1, un clasificador binario divide un conjunto de muestras en dos grupos diferentes. La ecuación general que describe a un clasificador binario es la definida en la Ec. 1.

El cálculo del umbral, λ_j , en un clasificador binario es una de las tareas más importantes y va a permitir obtener un clasificador con el objetivo deseado. Para ello existen diversas estrategias que se pueden plantear. Tres de los métodos más habituales son los siguientes:

- **Método 1:** Para calcular el umbral con este método, en primer lugar, se calcula el valor medio de las muestras positivas y el valor medio de las muestras negativas. La media de los dos valores anteriores será el umbral seleccionado, tal y como se muestra formalmente en la siguiente ecuación:

$$\lambda_j = \frac{1}{2} \left(\frac{1}{m} \sum_{p=1}^m f_j(x_p | y_p = 1) + \frac{1}{l} \sum_{n=1}^l f_j(x_n | y_n = 0) \right) \quad \text{Ec. 4}$$

donde m y l son el número de muestras positivas y negativas respectivamente. F_j es el j -ésimo clasificador.

- **Método 2:** El segundo método propuesto para seleccionar el umbral consiste en minimizar la distancia de la recta que une la esquina superior izquierda del espacio ROC (0, 1) con el valor del ROC del clasificador para cada posible valor del umbral, es decir, minimizar la distancia con el clasificador ideal.
- **Método 3:** El tercer método propuesto consiste en maximizar la suma de la sensibilidad y la especificidad, que equivale a maximizar la exactitud ponderada.

A modo de ejemplo, en la **Figura 18** de la sección 4.2, se representa la curva ROC de uno de los clasificadores descritos en la sección 7.2 para la clasificación de grietas longitudinales en la superficie de la carretera. Los puntos correspondientes a los valores seleccionados por los métodos anteriormente descritos aparecen marcados. En la sección 7.2 se ofrece más información sobre estos resultados.

4.4 Combinación de clasificadores binarios mediante el algoritmo de entrenamiento AdaBoost

La combinación de varios clasificadores se comporta mejor que un solo clasificador, [Kon07]. Es por ello que se ha buscado la combinación de varios clasificadores binarios para conseguir mejores resultados en la clasificación. El algoritmo de AdaBoost (Contracción de *Adaptive Boosting*) fue originalmente presentado por Freund and Schapire en 1995, [Fre95]. Posteriormente ha sido utilizado en multitud de ocasiones con pequeñas variaciones, como por ejemplo en [She04] y [Cor12].

El algoritmo de entrenamiento AdaBoost crea un único clasificador a partir de la combinación de varios clasificadores. El clasificador resultante es conocido como clasificador fuerte mientras que los clasificadores utilizados para formar este clasificador son llamados clasificadores débiles. El proceso para conseguir el clasificador fuerte es un proceso iterativo en el que T clasificadores débiles son seleccionados en T rondas. En cada iteración, todas las características son examinadas de forma exhaustiva para encontrar el mejor clasificador débil, que es aquel que tenga el menor error ponderado de clasificación. A continuación, ese error es utilizado para actualizar los pesos de cada muestra, de tal manera que el peso de las muestras clasificadas erróneamente aumente su peso en el siguiente cálculo de errores. Finalmente, una combinación ponderada de los clasificadores débiles seleccionados forma el clasificador fuerte.

La metodología utilizada sigue el siguiente proceso:

1. Dado un conjunto de muestras (\mathbf{x}_i, y_i) , $i = 1, 2, \dots, N$, donde \mathbf{x}_i es un vector con el valor de las características de la i -ésima muestra, de las cuales m son muestras positivas ($y_i=1$) y l son negativas ($y_i=0$)
2. Se inicializan los pesos, w , para cada muestra i

$$w_{1,i} = \begin{cases} \frac{1}{2m} & \text{Si } i \text{ es una muestra positiva} \\ \frac{1}{2l} & \text{Si } i \text{ es una muestra negativa} \end{cases} \quad \text{Ec. 5}$$

3. Desde $t = 1, \dots, T$

- a. Se normalizan todos los pesos
- b. Para cada característica j , se entrena un clasificador débil, h_j , con error ε_j

$$\varepsilon_j = \sum_{i=1}^N w_{t,i} |h_j(x_i) - y_i| \quad \text{Ec. 6}$$

- c. Se elige el clasificador débil h_j con el menor error, ε_j
- d. Se actualizan sus pesos

$$w_{t+1,i} = w_{t,i} \beta_t^{1-e_i} \quad \text{con}$$
$$e_i = \begin{cases} 0 & \text{Si } x_i \text{ está clasificado correctamente} \\ 1 & \text{en cualquier otro caso} \end{cases} \quad \text{Ec. 7}$$
$$\text{y } \beta_t = \frac{\varepsilon_t}{(1-\varepsilon_t)}$$

4. Se crea el clasificador fuerte $H(x)$ combinando los clasificadores $h(x)$

$$H(x) = \begin{cases} 1 & \text{si } \sum_{t=1}^T \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^T \alpha_t \\ 0 & \text{en cualquier otro caso} \end{cases} \quad \text{Ec. 8}$$

Siendo $\alpha_t = \log\left(\frac{1}{\beta_t}\right)$

4.5 Técnicas para la clasificación de tipos de defectos

Las técnicas para la clasificación de defectos en clases son aplicadas en la fase final del proceso de inspección visual para asignar a qué tipo concreto de defecto corresponde cada una de las instancias detectadas. Por tanto, para aplicar estas técnicas se requiere de un paso previo en el que las imágenes hayan sido segmentadas y los píxeles considerados como defectos hayan sido agrupados, formando una única entidad. De estas entidades se extraen un conjunto de características que definen a cada una de las instancias y son utilizadas por las técnicas de clasificación que se van a describir en esta sección para determinar a qué tipo de defecto corresponden.

Estas técnicas están basadas en el aprendizaje automático supervisado, que es un tipo de clasificación que cuenta con un conocimiento a priori extraído de un conjunto de muestras o modelos previamente clasificados de forma manual por un experto.

En esta sección se van a describir tres de las técnicas de clasificación supervisada más populares: las redes neuronales artificiales, el clasificador de K vecinos más cercanos y el clasificador bayesiano. La aplicación de estas técnicas requiere la selección de un conjunto de características que deben medirse para cada una de las muestras, siendo un factor determinante en el desempeño del clasificador, y donde se pone el foco atención en el Capítulo 5 de la presente tesis.

4.5.1 Redes neuronales

Las redes neuronales artificiales (RNA) o *Artificial Neural Networks* (ANN) son básicamente sistemas especializados de cómputo que intentan emular las capacidades de procesamiento de los sistemas biológicos, principalmente del cerebro humano, a través de estructuras que imitan las naturales, cuyo componente principal son las neuronas. Por tanto, las neuronas artificiales tienen una cierta similitud en su comportamiento con las neuronas naturales: toman información proveniente de otras neuronas, o desde un dispositivo sensor, realizan algún tipo de operación simple sobre estos datos, y le entregan los resultados a otras neuronas.

La forma de operación de las redes neuronales consiste en agrupar estos pequeños elementos de proceso en capas, las cuales operan al mismo tiempo sobre las señales provenientes de capas anteriores, constituyendo un ámbito de procesamiento paralelo. Cabe destacar que estas estructuras muestran un paralelismo lógico si se observan todas las neuronas pertenecientes a una misma capa. Por otro lado, también se pueden detectar operaciones en serie cuando se considera el transvase de información de una capa a la siguiente.

A pesar de la sencillez de la arquitectura de las redes neuronales artificiales, éstas son capaces de ofrecer algunas características propias de los sistemas naturales, como pueda ser el aprendizaje, la generalización a partir de ejemplos previos y la extracción de datos importantes dentro de un cierto conjunto al mismo tiempo que se desechan otros irrelevantes.

Las redes neuronales son utilizadas para la clasificación de defectos superficiales, como en [Yaz08], [Pen09] y [Mas12], donde se utilizan algunas variantes de redes neuronales para la clasificación de defectos superficiales en chapa laminada. En la sección 7.1 se encuentran los resultados obtenidos en esta tesis al aplicar la clasificación mediante redes neuronales a defectos de chapa laminada.

4.5.1.1 El modelo biológico y la neurona artificial

Las redes neuronales artificiales están inspiradas en modelos biológicos. Esto es, el punto de partida para ciertos diseños se inspiró en el cerebro humano. Como también sabemos, el conocimiento existente sobre este órgano es limitado, así que se constituye únicamente como una referencia, y en ciertos casos las líneas de investigación se alejan de él. Cabe destacar que en los orígenes de las redes neuronales un gran número de investigadores procedía del campo de la neurobiología, por lo que existía una realimentación positiva entre los descubrimientos que se hacían en la neurobiología y el campo de las RNA. Una pequeña introducción al

funcionamiento de las neuronas naturales es aconsejable para entender posteriormente el porqué de las estructuras artificiales.

El sistema nervioso humano, constituido en esencia por un tipo de células especiales, llamadas neuronas, es de una complejidad extrema. Un número aproximado de 10^{11} neuronas están comunicadas entre sí por unas 10^{15} conexiones. Cada neurona comparte muchas características con el resto de células que constituyen el cuerpo humano, pero tiene capacidades propias únicas de recepción, procesamiento y transmisión de señales electroquímicas a través del sistema de comunicación cerebral.

El propósito principal para el que se diseñó la neurona artificial era imitar las características básicas de las neuronas naturales. Básicamente, se le aplica un conjunto de entradas, cada una de las cuales representa la salida de otra neurona. Cada entrada se multiplica por un factor determinado, que se denomina peso, y que se puede asimilar a la importancia relativa de cada sinapsis en la neurona biológica. Todas las entradas, multiplicadas por sus correspondientes pesos, son entonces sumadas, para determinar el nivel de activación que llega a la neurona. El conjunto de entradas es un vector columna \mathbf{X} compuesto por n elementos, x_1, x_2, \dots, x_n . Cada señal es multiplicada por su peso asociado, w_1, w_2, \dots, w_n , un vector fila denotado por \mathbf{W} .

Finalmente, las entradas multiplicadas por sus correspondientes pesos se aplican a un bloque que las suma, y al que se hace referencia mediante el símbolo Σ . De esta forma, queda determinado el nivel de activación de la neurona artificial, que se representa con un simple producto de matrices:

$$\mathbf{W} \vec{X} = w_1 x_1 + w_2 x_2 + \dots + w_n x_n = a \quad \text{Ec. 9}$$

donde a , es el nivel de activación.

El siguiente paso para modelar la neurona artificial, consiste en incluir un elemento más: la función de salida, conocida también como función de transferencia. La señal a , que se había obtenido anteriormente, suele ser procesada por una cierta función F para producir la salida final de la neurona, a la que se denomina y . Dicha función puede ser una simple función lineal, una función escalón o una tangente hiperbólica, que es la función más utilizada por ser una función monótonamente creciente, no lineal y derivable y además estar acotada.

A este modelo también se le puede añadir un elemento adicional, que en ciertos diseños no se contempla, denominado bias. El bias es una entrada permanentemente

puesta a uno, que se conecta al elemento sumador, con un cierto peso b . El modelo completo se encuentra en la **Figura 19**.

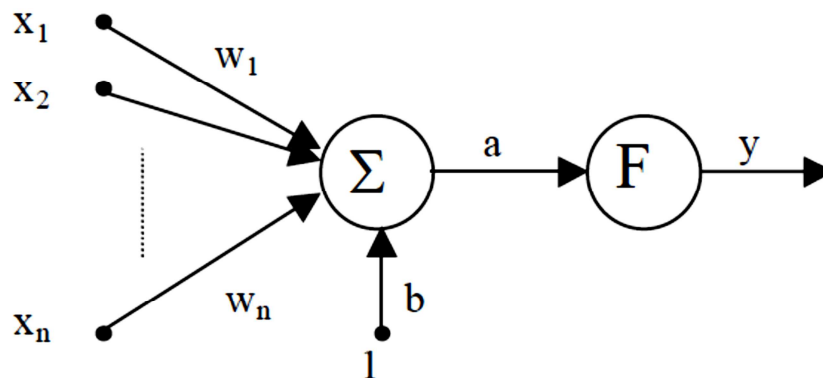


Figura 19. Modelo completo de una neurona artificial

4.5.1.2 Estructura de las redes neuronales artificiales

El primer estudio sistemático sobre las redes neuronales artificiales lo realizaron McCulloch y Pitts, en 1943, [McC43], los cuales publicaron su trabajo sobre el uso de estas redes en reconocimiento de patrones en 1947, [Pit47]. Su modelo de red era una única neurona como la representada en la **Figura 19**, cuya función de transferencia era la función escalón. A este tipo de neurona se le conoce con el nombre de *perceptrón*.

Para que la red neuronal cumpla su cometido es necesario entrenarla de manera adecuada. La demostración del teorema de aprendizaje del *perceptrón* la llevo a cabo Rosenblatt en 1962, [Ros62]. Llegó a la conclusión que una red neuronal solo puede llegar a aprender aquello que es capaz de representar. Para ampliar el rango de funciones que se pueden representar se introdujeron en las redes neuronales más neuronas dispuestas en diversas capas, utilizando diversas funciones transferencias.

En la **Figura 20**, se puede ver un ejemplo más complejo de la estructura de una red neuronal, la cual es descrita por los siguientes elementos:

- Entradas
- Pesos
- Capas ocultas
- Capa de salida
- Las funciones de transferencia

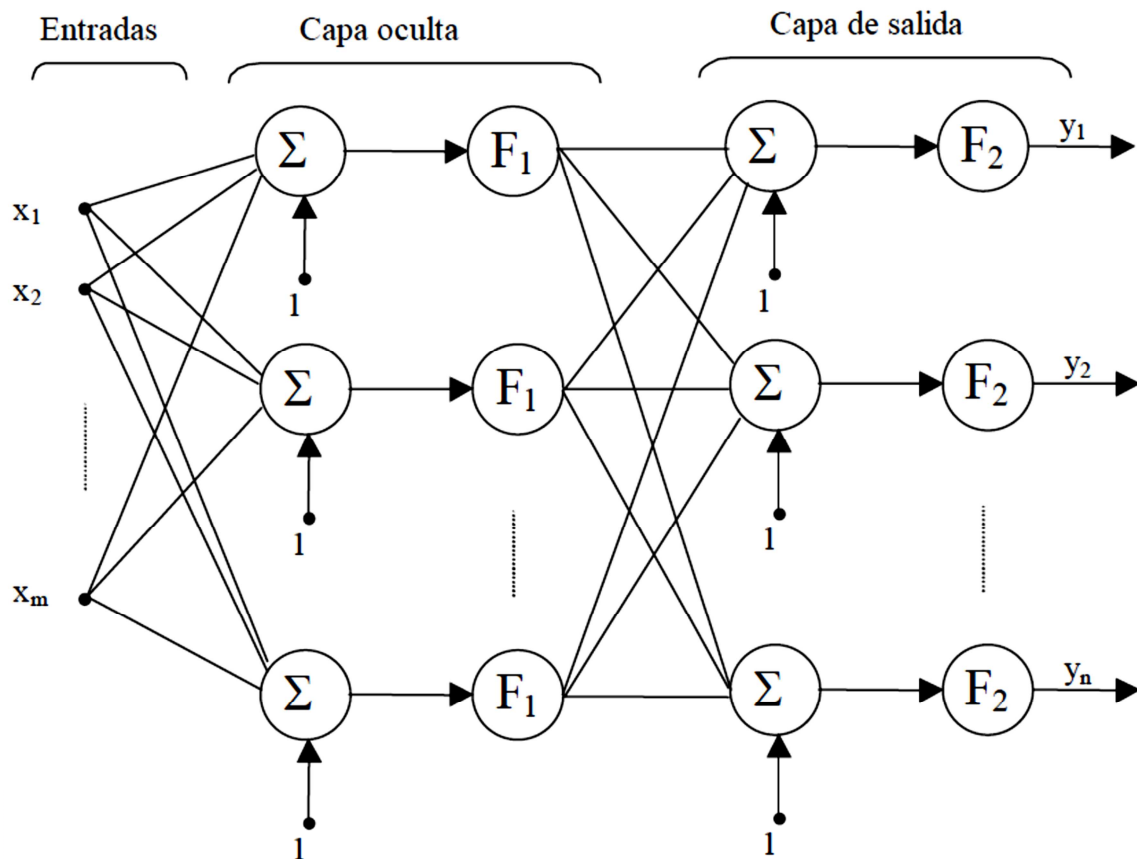


Figura 20. Modelo de red multicapa

4.5.1.3 Entrenamiento. El algoritmo *Backpropagation*

Backpropagation es un método sistemático para entrenar redes neuronales artificiales multicapa. Rumelhart, Hinton y Williams (1986), [Rum86], presentaron una descripción acertada del algoritmo *backpropagation*. Unos meses más tarde, otro investigador llamado Parker demostró haber realizado el mismo trabajo en 1982 y haberlo patentado, [Par82]. También se descubrió que la primera descripción del algoritmo databa de 1974, en la tesis doctoral de Werbos, [Wer74].

El algoritmo *Backpropagation* fue creado como una generalización de la regla de Widrow – Hoff, válida para redes neuronales lineales, adaptándola a redes multicapa con funciones de transferencia no lineales y diferenciables. La red se entrena mediante pares de entrenamiento, intentando aproximar una cierta función, o para resolver problemas de clasificación definidos por el usuario.

Hornik, Stinchcombe y White demostraron en 1989, [Hor89], que una red neuronal con señales de bias, una primera capa de función de transferencia sigmoide (por

ejemplo, la tangente hiperbólica), y una capa de salida lineal, es capaz de aproximar cualquier función que tenga un número finito de discontinuidades.

Las redes neuronales artificiales entrenadas mediante *backpropagation* tienden a mostrar salidas razonables cuando se le presentan entradas que nunca antes habían visto en el proceso de entrenamiento. Generalmente, un vector de entrada nuevo produce una salida de la red neuronal similar a la salida correcta asociada a un vector de entrada parecido al que presentamos, y que sí estaba en los pares de entrenamiento previamente aplicados. Esta propiedad de generalización hace posible el entrenamiento de la red mediante unos pocos pares de entrenamiento representativos, sin tener que pasar por todos los pares entrada / salida deseados.

Al estar compuesta la red neuronal por neuronas con función de transferencia no-lineal, la superficie de error que se genera, y sobre la cual se debe llegar al mínimo en el proceso de entrenamiento, es una superficie con muchos mínimos locales, al contrario que en el caso lineal que sólo se tenía un mínimo. El entrenamiento mediante *backpropagation* puede provocar que se alcance un mínimo local en la superficie de error, en lugar del mínimo global que sería deseable. El mínimo local puede ser satisfactorio para nuestras pretensiones, pero si no es así, se puede probar a aumentar el número de neuronas de la red o reiniciar el algoritmo partiendo de diferentes condiciones iniciales. De todas formas, el número de neuronas a añadir no se puede calcular de un modo analítico, siendo el mejor modo de proceder el de ensayo y error. La otra forma de solventar el problema, consiste en reiniciar el proceso de aprendizaje con unas condiciones iniciales distintas.

El objetivo del entrenamiento es ajustar los pesos de la red para que el conjunto de vectores de entrada a aplicar produzca el conjunto de vectores de salida adecuado. Una red tiene que entrenarse mediante la aplicación de un número determinado de pares de entrenamiento.

El algoritmo de entrenamiento sigue tres fases fundamentales: inicialización, hacia delante y hacia atrás. El proceso esquematizado se muestra en la **Figura 21**, se describe con detalle a continuación:

- 1. Inicialización:** La fase de inicialización se ejecuta una sola vez.
 - a. Construcción de la red:** Se define la estructura de la red, el número de entradas, de salidas, de capas ocultas y el número de neuronas de cada capa. No existe una metodología para elegir el número de neuronas de las capas ocultas, por lo que se suele emplear el método de ensayo y error.
 - b. Inicialización aleatoria de pesos y umbrales:** Se inicializan los pesos y los umbrales (bias) de cada neurona con un número aleatorio entre -0,5 y +0,5.

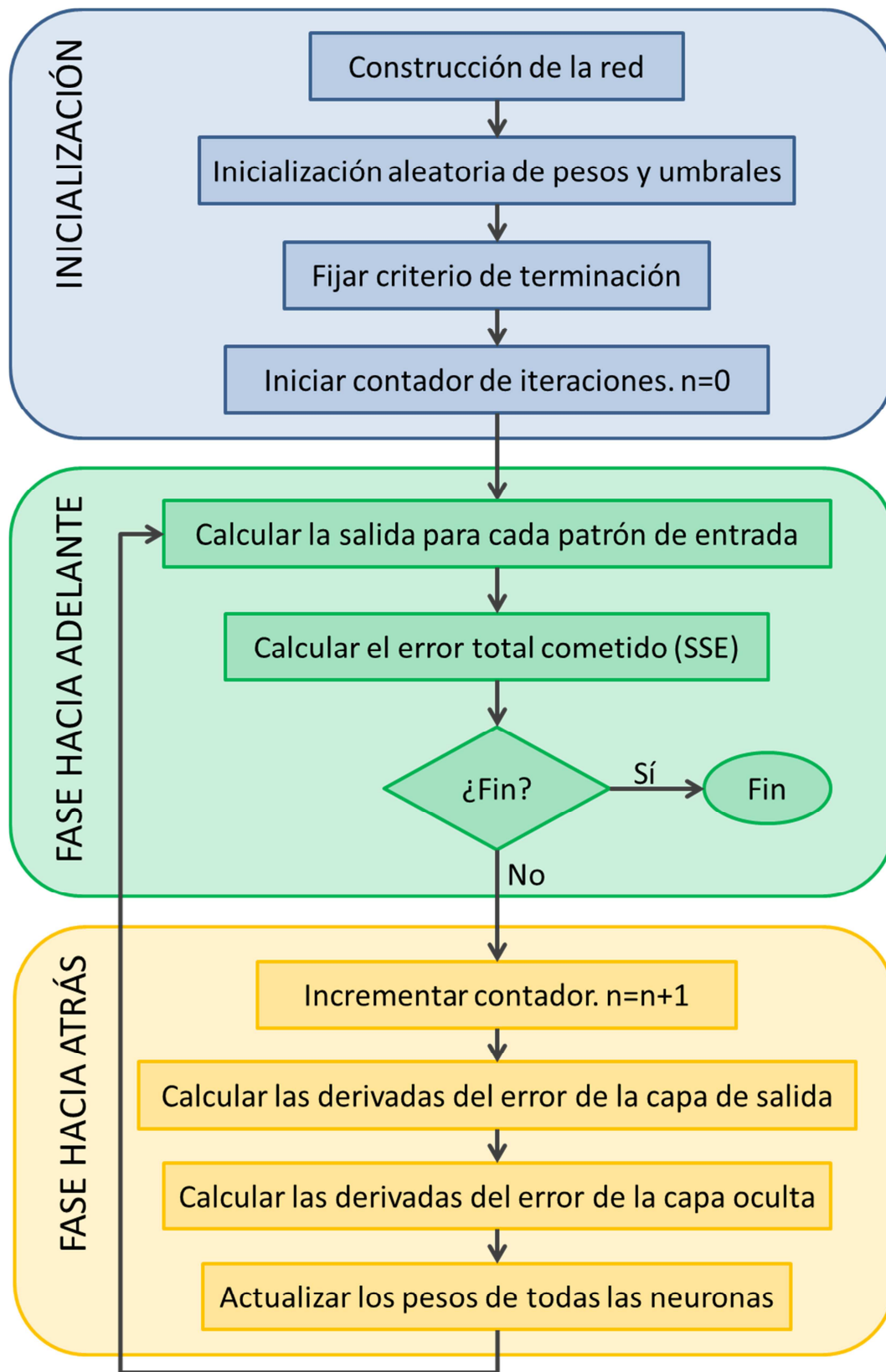


Figura 21. Algoritmo backpropagation

Lo más habitual es utilizar funciones de transferencia sigmoideal o lineales. El valor de la derivada de la función sigmoideal es $y_{pk}(1-y_{pk})$, mientras que el de la función lineal es 1.

- c. Calcular las derivadas del error de la capa oculta:** Los errores de las capas ocultas depende de los errores de las capas de salida y se calculan aplicando la siguiente formula:

$$\delta_j = f'(a_j) \sum_k \delta_k w_{kj} \quad \text{Ec. 12}$$

Donde j es la neurona oculta que se está evaluando y k son todas las neuronas de la capa superior a las que está conectada.

- d. Actualizar los pesos de todas las neuronas:** Se actualizan los pesos de todas las neuronas utilizando las deltas calculadas.

$$\Delta w_{ij}(n+1) = \eta \delta_i o_j + \alpha \Delta w_{ij}(n) \quad \text{Ec. 13}$$

Donde η es la razón de aprendizaje y siempre tiene un valor mayor que 0 y α es una constante entre 0 y 1, llamada momento que acelera el proceso de convergencia y disminuye posibles oscilaciones.

A continuación se actualizan los pesos y el proceso continúa en la "fase adelante".

4.5.2 Clasificador de K vecinos más cercanos

Al clasificador de K vecinos más cercano se le conoce con el nombre K-NN (*K – Nearest Neighbour*). Se trata de un método de clasificación supervisada que se basa en la idea básica de que una nueva muestra se va a clasificar en la clase más frecuente a la que pertenecen sus K vecinos más cercanos. Este método, también se aplica a la clasificación de defectos superficiales en chapa laminada en [Dup97].

Este método ha sido empleado recientemente utilizando filtros de Gabor como características para la clasificación de imágenes con distintos fines. En [Aba12] se clasifican las texturas de restos arqueológicos de cerámica mediante la aplicación del algoritmo de clasificación de vecinos cercanos (KNN) a imágenes filtrados por un banco de filtros de Gabor de 6 frecuencias y 4 orientaciones. En [Guo14] se presenta un método para la detección de vehículos en carretera mediante la utilización de bancos de filtros de Gabor. Utilizan el clasificador de vecinos cercanos (KNN) en el que el

vector características son los parámetros que describen la distribución de las magnitudes de las imágenes filtradas por cada filtro de Gabor. Para esa descripción se busca que esta curva de distribución se aproxime a la función gamma y se compara con los resultados que se obtienen cuando se utiliza la función Gaussiana.

En este apartado se va a describir el algoritmo básico, algunas variantes del mismo, y algunas técnicas de búsqueda para disminuir los tiempos de procesamiento al aplicar este algoritmo.

4.5.2.1 Algoritmo básico

El algoritmo K-NN se basa en una idea muy simple e intuitiva, así como fácil de implementar su algoritmo básico, por lo que su uso se ha extendido mucho especialmente en el campo de la detección de formas, aunque es perfectamente aplicable para resolver cualquier problema de clasificación en el que se pueda describir mediante características cuantificables el problema a resolver.

Para implementar el algoritmo es necesario disponer de una serie de muestras de las que se conoce a priori el grupo al que pertenecen y disponer de un grupo de características cuantificables que las describan. Ante una nueva muestra a analizar, se calculan las distancias del vector características de cada muestra de la base de datos con la nueva muestra. A continuación se ordenan las distancias calculadas y se seleccionan las K muestras más cercanas. La nueva muestra pertenecerá al grupo que aparece con mayor frecuencia en las K muestras seleccionadas. En caso de empate en la frecuencia de aparición se utiliza una regla heurística para dilucidar el grupo a asignar.

Existen algunas variantes de este algoritmo que permiten mejorar algunos problemas de clasificación que se van a presentar a continuación.

4.5.2.2 Variantes del algoritmo

En este apartado se exponen algunas variantes del algoritmo básico que permiten mejorar los resultados en algunos casos.

4.5.2.2.1 K-NN con rechazo

Esta variante se basa en la idea de que para poder clasificar un nuevo caso deben existir algunas garantías de que corresponde realmente a esa clase. La variante K-NN con rechazo introduce la posibilidad de que algunos casos queden sin clasificar. Existen dos alternativas para implementar esta variante:

- El número de votos de una determinada clase debe superar un umbral prefijado para ser asignada a una clase.
- La diferencia entre la mayor frecuencia y la siguiente supere un determinado umbral.

4.5.2.2.2 K-NN con distancia media

Esta variante en vez de considerar que un nuevo caso pertenece a la clase con mayor frecuencia, asigna al nuevo caso a la clase con menor distancia media dentro de los K casos seleccionados por el algoritmo. Es decir, se seleccionan los K casos de la misma forma en que se hace en el algoritmo básico, se calculan los valores medios del vector características de las muestras pertenecientes a cada una de las clases presentes en los K casos seleccionados, y, finalmente, se calcula la distancia de la nueva muestra a esos vectores de valores medios, asignando la nueva muestra al caso con menor distancia.

4.5.2.2.3 K-NN con distancia mínima

La variante K-NN con distancia mínima busca reducir el tiempo de computación y solo es válida para muestras con alto grado de homogeneidad dentro de la misma clase. Para implementarlo se calcula el baricentro de todos los casos de cada clase y selecciona para cada clase el caso más próximo a su baricentro. De esta manera se tiene un único caso representativo para cada clase. Para evaluar un nuevo caso se calcula su distancia a cada caso representativo y se le asigna la clase del caso representativo más próximo. Este procedimiento puede verse como un 1-NN en el que se dispone de un solo caso representativo para cada clase.

4.5.2.2.4 K-NN con pesado de casos seleccionados

Esta variante consiste en asignar distintos pesos a cada caso a la hora de evaluar la frecuencia de los K casos más próximos. Es habitual utilizar como peso el inverso de la distancia al nuevo caso, aunque es posible utilizar otras reglas para el cálculo de los pesos.

4.5.2.2.5 K-NN con pesado de variables

En el algoritmo básico, las distancias entre los casos conocidos y el nuevo caso se calculan por medio de la distancia euclídea dando el mismo peso a todas las variables. En esta variante se asignan diferentes pesos a las variables a la hora de calcular las distancias, ya que a veces existen variables que son irrelevantes para determinadas clases. Para ello se aplica la siguiente fórmula:

$$d(x_i, x) = \sqrt{\sum_{j=1}^n w_j (x_{i,j} - x_j)^2}$$

Ec. 14

donde x_i son los casos conocidos, x el caso nuevo y w_j el peso asignado a la variable j .

4.5.2.3 Técnicas de búsqueda de los K vecinos más cercanos

Una de las limitaciones que aparece a la hora de aplicar el algoritmo de los K vecinos más cercanos es el tiempo necesario para la búsqueda al evaluar un nuevo caso. El tiempo necesario para la búsqueda está principalmente influenciado por la cantidad de muestras disponibles y la dimensión del vector de características. Estas técnicas de búsqueda no solo se utilizan para la clasificación, que es el objeto de este trabajo, sino que han tenido un gran auge al ser ampliamente utilizadas para realizar consultas en bases de datos. De especial relevancia son los métodos de búsquedas en espacios vectoriales, como son los casos en los que los vectores de características son valores numéricos que describen a las muestras.

El método trivial para resolver el problema de la búsqueda de vecinos más cercanos, es la búsqueda exhaustiva o búsqueda por fuerza bruta, que consiste en recorrer la lista de todas las muestras. Para cada nueva muestra que se desea evaluar se calcula la distancia a cada muestra y se guardan aquellas que se encuentran entre las K más cercanas hasta el momento.

El coste en número de distancias a calcular es proporcional al número de muestras. Cuando la dimensión del vector de características es elevada el cálculo de la distancia tiene un coste muy elevado, por lo que no suele ser aplicable. Es por ello por lo que han surgido técnicas que preprocesan los datos que reducen el número de distancias a calcular. Dentro de estas técnicas están las que dividen el espacio en árboles de búsqueda como las siguientes:

- **KD-Tree:** Es el método para espacios vectoriales más popular, [Ben75]. Se trata de un árbol binario en el que cada nivel del árbol es dividido por un hiperplano perpendicular a una dimensión del vector. Existen infinitas posibilidades a la hora de elegir el hiperplano de separación, aunque la técnica habitual es seleccionarlo de tal manera que el árbol este balanceado, es decir, haya la mismas cantidad de muestras en cada rama del árbol. Esta técnica se implementa calculando la mediana de las muestras para cada variable del vector de características.
- **R-Tree:** En este método cada nodo del árbol utiliza hiper-rectángulos, [Gut84], de tal manera que cada nodo encierra a los nodos hijos que los

contiene. El espacio se divide forma jerárquica, aunque pueden estar superpuestos, es decir, nodos de un nivel pueden estar en la región delimitada por dos o más nodos de un nivel superior. Cada nodo almacena la información de cuáles son los nodos del siguiente nivel y el hiper-rectángulo que los delimita. Cada nueva entrada se realiza de tal manera que busca aumentar el tamaño del nodo lo menos posible. El tamaño del nodo está limitado, por lo que cuando se supera el tamaño este se divide en dos.

- **X-Tree:** Este método es similar al R-Tree, pero eliminando el solapamiento entre los nodos, [Ver96]. Cuando un nodo al dividirse no puede evitar el solapamiento entre nodos, se construye lo que se conoce como supernodo, de mayor capacidad que el tamaño inicial fijado. De esta manera se consigue reducir el número de nodos donde se debe buscar un nuevo elemento.

Para aplicar el algoritmo de búsqueda de los K vecinos más cercanos en los métodos basados en la utilización de árboles, se busca en primer lugar en la hoja del árbol en la que se insertaría un nuevo caso. A continuación se debe verificar que hojas del árbol no pueden contener un vecino más cercano, lo que se conoce como poda, utilizando para ello la jerarquía del árbol. Se busca en la hoja más próxima al nuevo caso que podría contener un caso más cercano que el que se encuentra en la posición K en este momento. El algoritmo se detiene cuando no existen nuevas hojas del árbol que pudieran contener un caso más próximo que el que se encuentra en ese momento en la posición K .

4.5.3 Clasificador bayesiano. Clasificador naïve Bayes

Este clasificador supervisado introduce un enfoque probabilístico en el aprendizaje. Además es un clasificador competitivo, ya que las distintas hipótesis compiten entre sí, venciendo la que tiene mayor probabilidad de ser cierta.

Este clasificador se conoce bajo diversos nombres que incluye los de idiota Bayes [Ohm88], naïve Bayes [Kon90], simple Bayes [Gam91] y Bayes independiente [Tod94].

A pesar de su utilización en el reconocimiento de patrones, [Dud73], el clasificador Naïve Bayes aparece por primera vez en la literatura del aprendizaje automático a finales de los ochenta [Ces87] con el objetivo de comparar su capacidad predictiva con la de métodos más sofisticados.

El clasificador naïve Bayes introduce una serie de simplificaciones que permiten la aplicación del algoritmo. Estas simplificaciones son las siguientes:

- No pueden darse dos casos al mismo tiempo. Es decir, una misma muestra no puede pertenecer simultáneamente a dos clases.

- Las variables utilizadas para la clasificación son condicionalmente independientes dada la clase. Es decir, que conocido a que clase pertenece una muestra, el conocimiento de cualquiera de sus características es irrelevante para el resto de las características.

Por tanto, la probabilidad de que una muestra C pertenezca a la clase c , conocido el valor de las variables X_1, \dots, X_n es igual a la probabilidad de que la muestra C pertenezca a la clase c , sin conocimiento previo, multiplicado por cada una de las probabilidades de que la variable X_i tome el valor x_i , condicionado a que la muestra C pertenezca a la clase c , tal y como muestra la siguiente ecuación.

$$P(C = c | X_1 = x_1, \dots, X_n = x_n) = P(C = c) \prod_{i=1}^n P(X_i = x_i | C = c) \quad \text{Ec. 15}$$

Al suponer que una muestra solo puede pertenecer a una clase, la muestra C pertenecerá a la clase c con mayor probabilidad, como refleja la siguiente ecuación:

$$c^* = \arg \max_c \left(P(C = c) \prod_{i=1}^n P(X_i = x_i | C = c) \right) \quad \text{Ec. 16}$$

En el caso de que las variables X_1, \dots, X_n son continuas, se considera que siguen una distribución normal para cada clase, c , la probabilidad condicionada se sustituye por la función de densidad normal de la variable X condicionada a que la muestra C pertenezca a la clase c , cuya ecuación es la siguiente.

$$P(X_i = x_i | C = c) = \frac{1}{\sqrt{2\pi}\sigma_i^c} \exp\left(-\frac{1}{2}\left(\frac{x_i - \mu_i^c}{\sigma_i^c}\right)^2\right) \quad \text{Ec. 17}$$

Por tanto, para implementar el algoritmo de naïve Bayes, cuando se dispone de un grupo lo suficientemente grande de muestras para su entrenamiento del que se conoce el valor de una serie de características que los describen, se deben realizar los siguientes cálculos previos:

- Calcular la probabilidad de que ocurra cada una de las clases, sin disponer de conocimiento previo, es decir, $P(C=c)$.
- Para cada clase c , calcular la media, μ_i , y la desviación estándar, σ_i , de cada una de las variables o características, X_i .

Para clasificar una nueva muestra, se calcula el valor de cada una de las características de esa muestra, x_i , y se calcula cuál es la clase a la que la nueva muestra pertenece con mayor probabilidad aplicando la siguiente ecuación, resultado de substituir la Ec. 17 en la Ec. 16:

$$c^* = \arg \max_c \left(P(C = c) \prod_{i=1}^n \frac{1}{\sqrt{2\pi}\sigma_i^c} \exp\left(-\frac{1}{2}\left(\frac{x_i - \mu_i^c}{\sigma_i^c}\right)^2\right) \right) \quad \text{Ec. 18}$$

Esta técnica es empleada en esta tesis para la clasificación de defectos en chapa laminada. Los detalles de su implementación y los resultados obtenidos se encuentran en la sección 7.1.

4.6 Combinación de clasificadores de tipos de defectos

La combinación de varios clasificadores obtiene mejores resultados que la utilización de un único clasificador, [Kon07]. La combinación de clasificadores se puede sustentar a partir de sus ventajas estadísticas, computacionales y representacionales [Kun04]. Las técnicas de combinación de clasificadores están orientadas a problemas de clasificación en las que se de alguna de estas situaciones [Jai00]:

- Tener la posibilidad de entrenar varios clasificadores
- Tener varios conjuntos de entrenamiento cuyos datos sean independientes
- Tener varios clasificadores con un elevado rendimiento en una región determinada del espacio de características.

Existen diversas arquitecturas de combinación, entre las que destacan:

- **En pila:** Todos los clasificadores se entrenan con datos que provienen de un único espacio de características, es decir, se dispone de un único conjunto de entrenamiento y se combinan las salidas de los distintos algoritmos de entrenamiento.
- **Paralela:** Se combinan los clasificadores que han sido entrenados en diferentes espacios de características, es decir, se dispone de varios conjuntos de entrenamiento que se utilizan para entrenar bien un mismo clasificador o varios clasificadores distintos.
- **Secuencial:** Los clasificadores son aplicados en cascada, es decir, cada clasificador opera sobre la salida del anterior.

Las estrategias de combinación de los clasificadores dependen de cómo es la salida de los clasificadores. Cuando la salida de los clasificadores es una etiqueta, como es el caso de los clasificadores expuestos anteriormente, se utilizan **métodos de nivel abstracto**. Las principales reglas de combinación en este caso son las siguientes:

- **Voto mayoritario simple:** Esta técnica es similar al voto que se realiza en unas elecciones políticas, es decir, cada clasificador tiene un voto y la etiqueta más votada resulta elegida. Habitualmente se suele considerar que la salida es confiable cuando se alcanza la mayoría absoluta.
- **Voto mayoritario por peso:** Esta técnica es similar al modo en el que los accionistas de una empresa realizan sus votaciones, es decir, cada clasificador emite un voto que es ponderado por un peso. El peso de cada clasificador se asigna dependiendo de la eficacia de cada clasificador al realizar una clasificación de una muestra de entrenamiento.
- **Reglas bayesianas:** Es la técnica más complicada de implementar y la menos popular. Se basa en la estimación de la probabilidad a posteriori del conjunto de clasificaciones realizadas por los clasificadores. Es decir, para cada posible combinación de salida de los clasificadores se calcula la probabilidad de que pertenezca a una clase dada y se asigna la clase más probable. Estas probabilidades se calculan aplicando la clasificación a un conjunto de datos de entrenamiento y calculando la frecuencia de cada clase a cada combinación de las decisiones de los clasificadores. Es necesario calcular k^{N+1} probabilidades, siendo k el número de clase y N el número de clasificadores.

En muchas ocasiones un clasificador asocia a cada clase un valor o una probabilidad que indica el grado de confianza en que la clasificación de un patrón corresponda a una determinada clase. Este valor se puede utilizar para realizar una ordenación de la probabilidad de que un patrón pertenezca a cada una de las clases. De manera análoga a los métodos de nivel abstracto se pueden utilizar las tres siguientes reglas de combinación para lo que se conoce como **métodos de nivel de rango**:

- **Regla de la Borda:** Para una nueva muestra, se ordenan las posibles clases de menor a mayor probabilidad para cada uno de los clasificadores. A continuación se asigna a cada clase la suma del valor de la posición que ocupa para cada uno de los clasificadores. La clase con mayor valor será considerada la ganadora.
- **Regla de la Borda con peso:** Análoga a la anterior, pero la suma se realiza asociando pesos a cada clasificador. El peso de cada clasificador se asigna dependiendo de la eficacia de cada clasificador al realizar una clasificación de una muestra de entrenamiento.

- **Reglas bayesianas:** Como en el caso de los métodos de nivel abstracto se asignan probabilidades a cada posible combinación de los órdenes de cada clasificador. Es un método muy complejo que apenas se utiliza.

Muchos clasificadores ofrecen para cada clase como salida valores de confianza o medidas de distancia a un patrón. Estas medidas pueden ser interpretadas como la probabilidad de que una muestra pertenezca a una clase dada. Para poder realizar la combinación de estos clasificadores es necesario normalizar la salida, lo cual no siempre es una tarea trivial cuando se trata de clasificadores con diferentes tipos de salida. Este tipo de combinación se conoce como **métodos de nivel de medidas:**

- **Regla del promedio simple:** Es una de las reglas de combinación más sencillas que consiste en promediar la salida de los distintos clasificadores y asignar la nueva muestra a la clase con un mayor valor promedio.
- **Regla del producto:** Se suele utilizar cuando los vectores de características de los clasificadores individuales no coinciden. Consiste en realizar la multiplicación de las salidas de cada uno de los clasificadores y asignar la nueva muestra a la clase con mayor valor.
- **Otros operadores estadísticos:** Al igual que en los casos anteriores se aplica un operador a la salida de los clasificadores individuales. Los más habituales y que mejores resultados presentan son el máximo, el mínimo y la mediana.
- **Operadores pesados:** A la regla del promedio simple y a la regla del producto se le puede aplicar pesos, dependiendo de la confianza de cada clasificador individual. En el primer caso se multiplicaría a cada salida por ese peso mientras que en el segundo se elevaría el valor a la potencia del peso.

4.7 Conclusiones

En este capítulo se han presentado algunas técnicas de detección y clasificación de defectos y la forma de evaluar su desempeño. La detección de defectos puede considerarse una **clasificación binaria** en la que los píxeles de la imagen pertenecen a la superficie con defecto o a la superficie libre de defecto. Para realizar esta separación es preciso caracterizar la imagen mediante alguna propiedad de la misma y determinar cuál es el valor del umbral que delimita los dos grupos. La exactitud ponderada se presenta como la forma más apropiada de medir la eficacia de un clasificador binario en el que el número de positivos y negativos no está equilibrado, tal y como ocurre en las superficies objeto de estudio en esta tesis, en las que la superficie libre de defectos es muy superior a la que contiene defectos, y así va a realizarse en esta tesis.

El resultado de la detección es un conjunto de agrupaciones de píxeles que corresponde a zonas con deterioros superficiales. Clasificar estos deterioros en **tipos de defectos** es una tarea compleja para la que se pueden utilizar diversos tipos de clasificadores, si bien es cierto que la combinación de varios de ellos va a proporcionar mejores resultados que un único clasificador. A la fase de clasificación también pueden llegar agrupaciones de píxeles que no contienen defectos o que contienen defectos que no estaban contemplados. Esta situación también es contemplada en la presente tesis.

Tanto la detección, como su posterior clasificación, requieren evaluar una o varias **características de las imágenes**. Algunos de los deterioros que pueden aparecer sobre las superficies propuestas en el Capítulo 2 resultan difíciles de distinguir de la superficie libre de defectos si se evalúa simplemente el valor del nivel de gris de la imagen, por lo que es preciso plantear un nuevo enfoque en la selección de las características. En el Capítulo 5 se va a proponer analizar las imágenes en el dominio espacio-frecuencial, extrayendo información de las mismas utilizando filtros de Gabor.

5 ANÁLISIS ESPACIO-FRECUENCIAL DE LAS IMÁGENES MEDIANTE FILTROS DE GABOR

5.1 Motivación

En la sección 4 se describieron un conjunto de técnicas para la detección y clasificación de defectos. Estas técnicas requieren la evaluación de una o varias características de la imagen. La elección de estas características es tan importante, o incluso más, que la propia técnica de detección o clasificación.

La detección de defectos consiste en dividir la imagen en zonas que contienen defectos y en zonas que no los contienen. Pueden tratarse como un caso de clasificación binaria en el que la salida del clasificador tiene dos opciones: defecto o no defecto. Para realizar la detección se utilizan técnicas de segmentación en las que la principal dificultad radica en determinar cuál es el valor del umbral que delimita los píxeles o zonas de la imagen que pertenecen a la zona con defecto de los que pertenecen a las zonas libres de defecto. La característica a evaluar puede ser el nivel de gris de la imagen, aunque es habitual realizar algún tipo de transformación en la imagen antes de aplicar el proceso de detección. En la sección 2 se describieron diversos trabajos en los que se utilizaban estas técnicas para la detección de defectos en los tres ámbitos en los que se centra esta tesis. Algunas de estas técnicas se basan en la segmentación de la imagen original, bien utilizando un umbral fijo como en [Zha09], un umbral que depende de la evolución de imágenes anteriores como en [Yaz08] y en [Yan10], o utilizando un umbral distinto para cada columna de la imagen como en [Miy07]. Sin embargo, otros autores segmentan la imagen modificada mediante algún filtro, como en [Zha10], [Yaz08] y [Oli09a] que utilizan la entropía, o en [Mar10] que utilizan la transformada de Hough, o filtros morfológicos como en [Zha14], así como distintos filtros de detección de bordes como en [Mah91], [Riv15] y [Fuj11].

Algunos de los defectos que aparecen en las superficies que se inspeccionan en esta tesis no pueden ser detectados con las técnicas descritas en el párrafo anterior con un grado de fiabilidad aceptable. Claro ejemplo de ello son las grietas que aparecen en la carretera y las grietas o fisuras que aparecen en la superficie de los túneles. En ambos casos el sistema de adquisición de imágenes ha sido configurado para que se provoque, en la medida de lo posible, una sombra en la zona de la grieta de tal manera que el nivel de gris de la imagen donde se encuentran las grietas presente un color más oscuro que el resto de la imagen. Diversos factores afectan a la intensidad

de esta sombra, pero incluso en los casos más favorables, estas grietas se encuentran en un fondo que contiene numerosas zonas con un nivel de gris similar, o incluso más bajo que el de las grietas, debido tanto al betún que contiene el asfalto, cuyo color es habitualmente negro, como a las propias sombras que se forman en este tipo de superficies, ya que tanto la carretera como el hormigón son habitualmente rugosos, y a manchas u otras variaciones en el nivel gris.

En esta sección se muestran algunos ejemplos de esta situación aplicados a imágenes adquiridas con los sistemas de inspección que se describen en la sección 6, tras un preprocesamiento para homogeneizar el nivel de gris de la imagen.

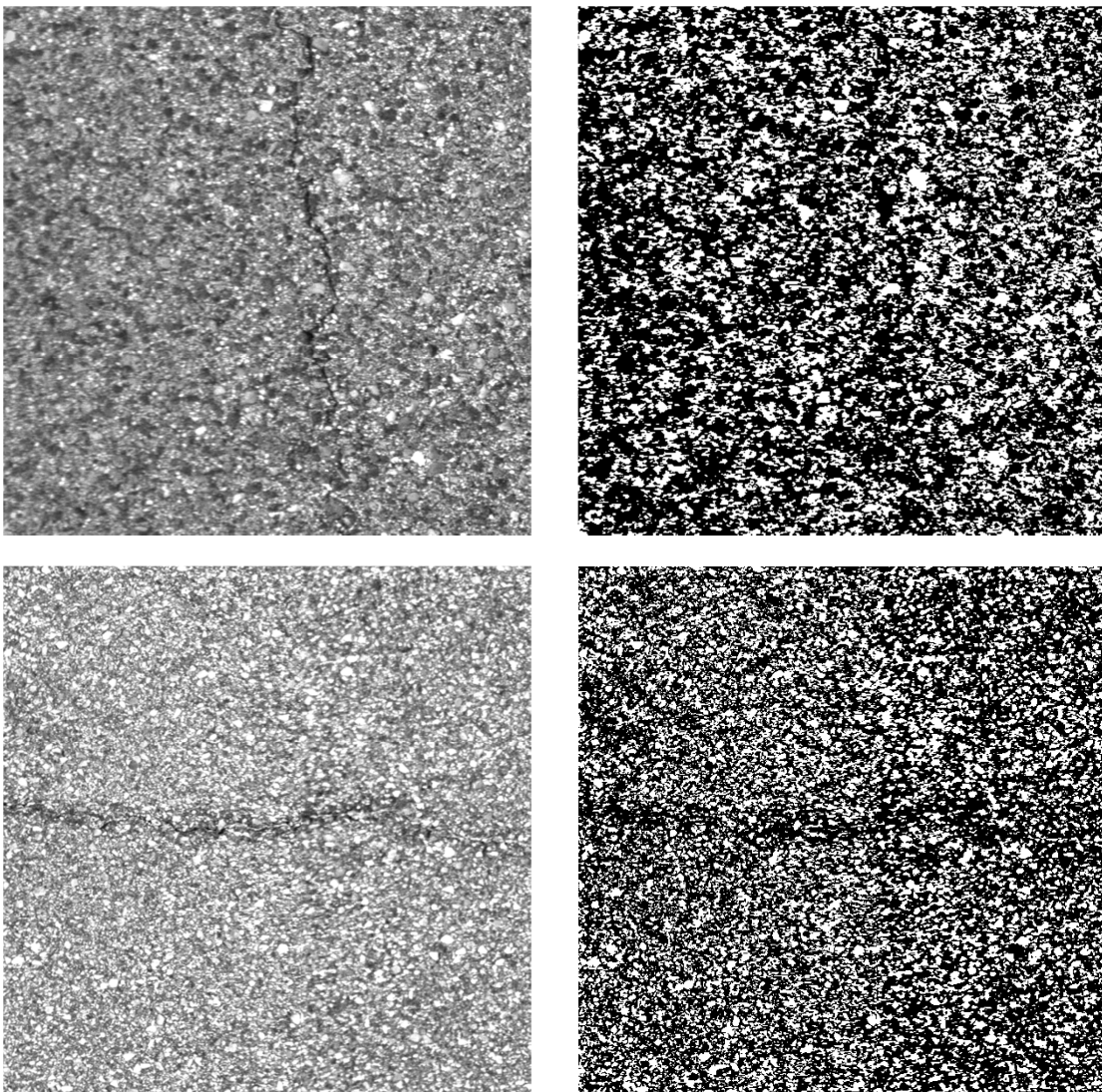


Figura 22. Dos imágenes de 512x512 píxeles con una resolución de 1mm de grietas en carretera y su segmentación mediante el método de Otsu. En la fila superior una grieta longitudinal. En la fila inferior una grieta transversal.

En la **Figura 22** se encuentran dos imágenes de grietas en carretera segmentadas mediante el método clásico de Otsu, [Ots79]. Como se puede ver, el resultado de la segmentación no permite la discriminación de las grietas del asfalto. También se ha aplicado esta misma técnica a imágenes de grietas en túneles de hormigón, cuyo resultado se encuentra en la **Figura 23**, con resultados un poco mejores, pero claramente insuficientes.

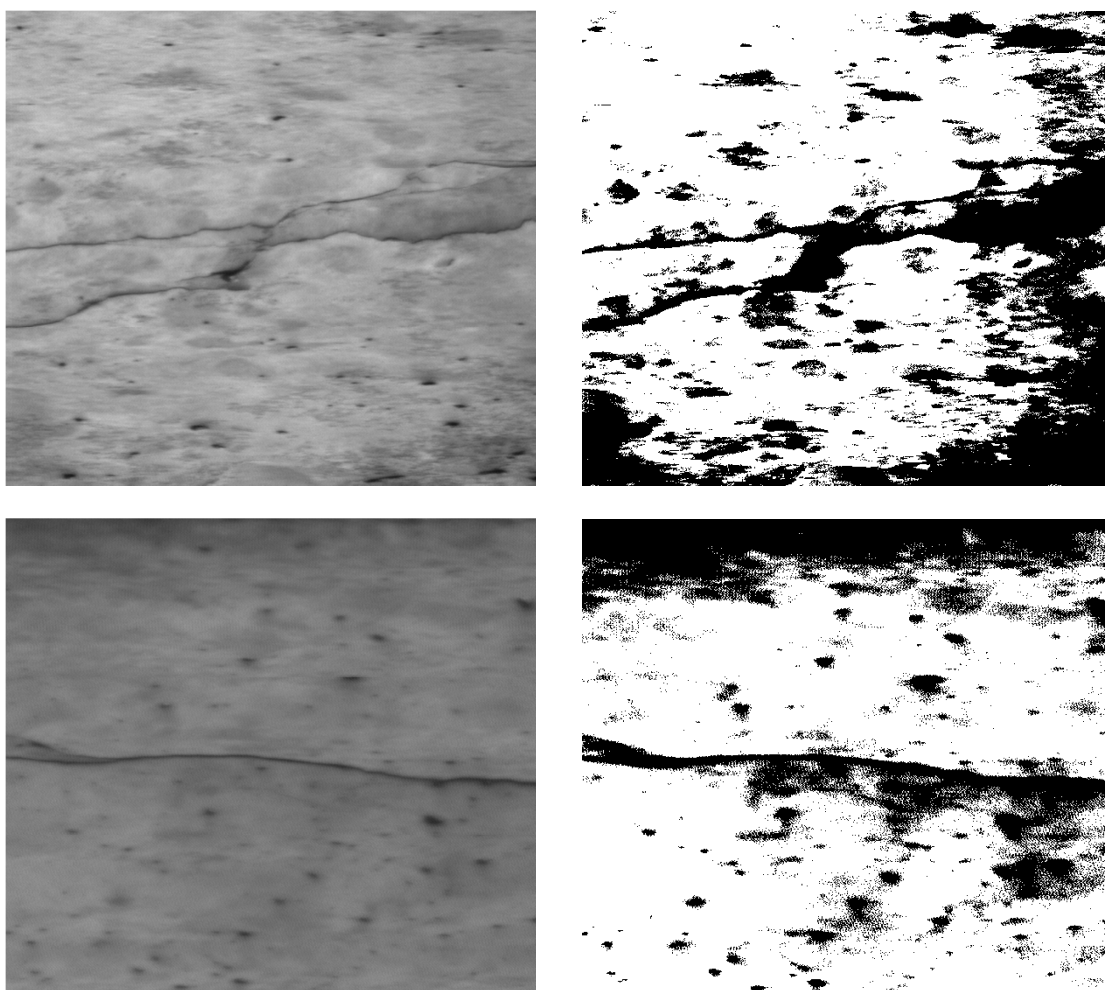


Figura 23. A la izquierda, dos grietas en la superficie de un túnel de hormigón con una apertura aproximada de 1mm; a su derecha, el resultado de la segmentación de la imagen utilizando el método Otsu. Imágenes de 1024x918 píxeles adquiridas con una resolución de 0,2mm/píxel

El método de cálculo del umbral es crítico en este tipo de métodos y algunas operaciones morfológicas aplicadas al resultado de la segmentación pueden llegar a mejorar sensiblemente el resultado final. En la **Figura 24** se muestra a la izquierda el resultado de segmentar la imagen superior de la **Figura 22**, que corresponde a una grieta longitudinal, eligiendo el umbral manualmente. A la derecha se encuentra el resultado de aplicar un filtro morfológico denominado cierre. Aunque el resultado es

sensiblemente mejor al obtenido por el método de Otsu, que calcula el umbral de forma automática, el resultado no es aún adecuado. Además, el mismo umbral fijo no va a funcionar igual en todo tipo de carreteras, ni siquiera en distintos tramos de una misma carretera.

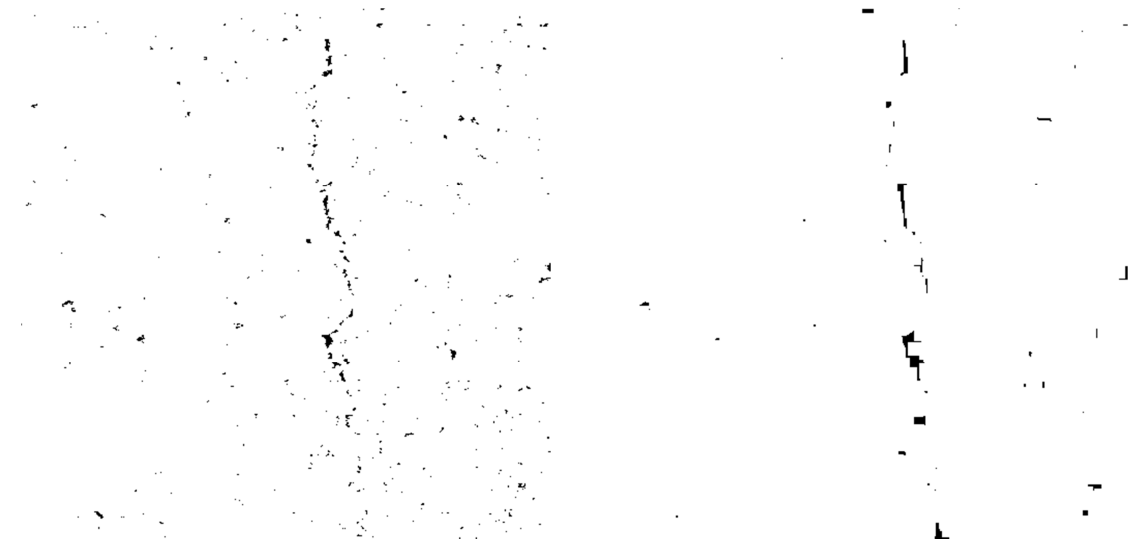


Figura 24. A la izquierda, la grieta longitudinal de la **Figura 22** segmentada eligiendo el umbral manualmente; a la derecha la misma imagen tras la operación morfológica de cierre

Si se aplica la segmentación a imágenes filtradas mediante filtros morfológicos, de detección de bordes, etc. el resultado que se ha obtenido es similar. A continuación se muestran varios ejemplos en los que se aplican diversos filtros y se segmenta la imagen filtrada. En la **Figura 25** se utiliza el filtro de Prewitt, en la **Figura 26** el filtro de Sobel, en la **Figura 27** se han aplicado los de Canny de Roberts, que proporcionan ya la imagen binarizada y en la **Figura 28** el filtro morfológico top-hat con un elemento estructural circular de 7 píxeles de radio.

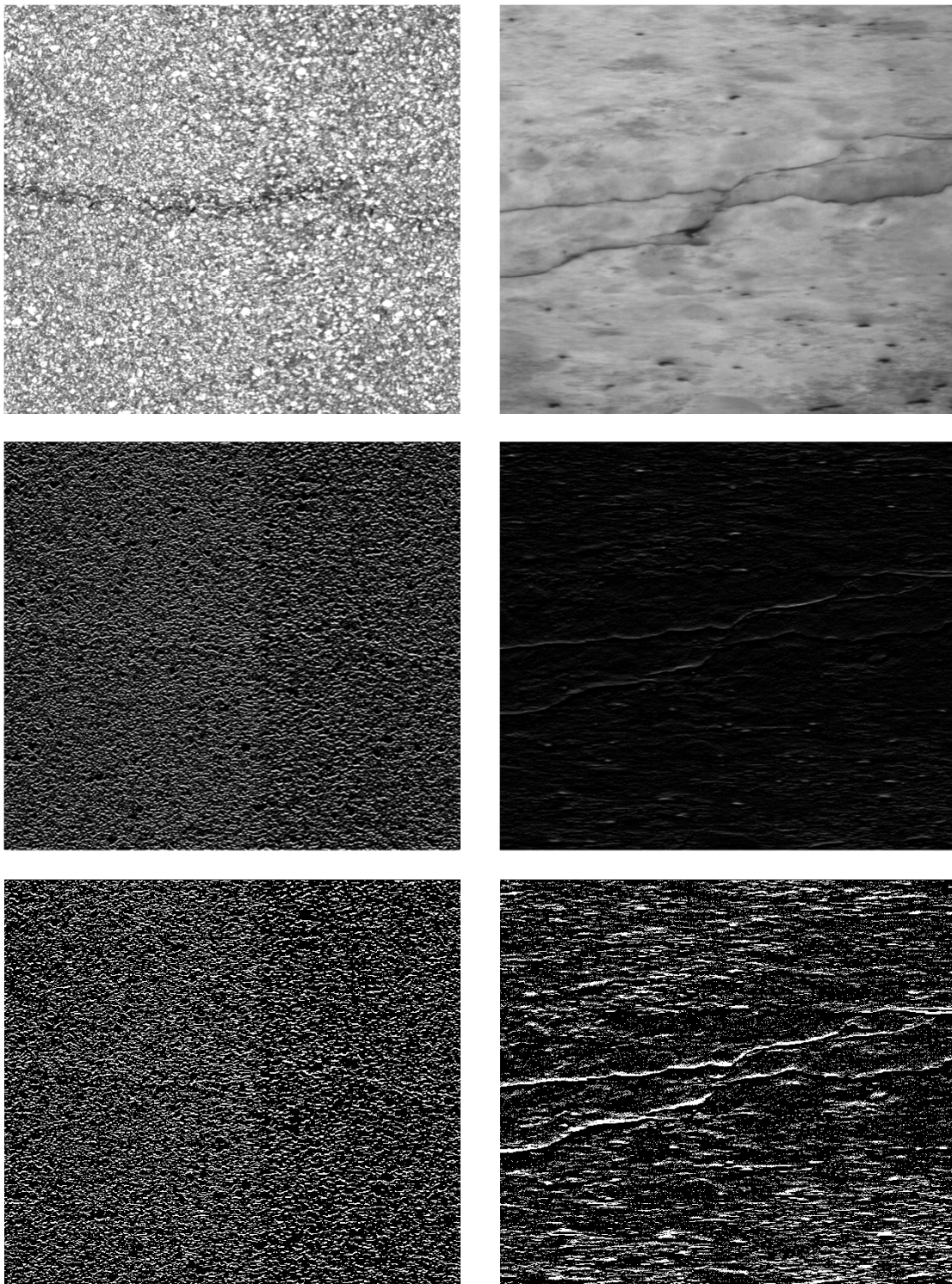


Figura 25. Dos imágenes filtradas por el Filtro Prewitt. La columna de la izquierda es una grieta de carretera, la columna de la derecha es una grieta en un túnel. La fila superior es la imagen original preprocesada, la fila central es el resultado de aplicar el filtro de Prewitt, la fila inferior es la imagen filtrada segmentada por el método Otsu

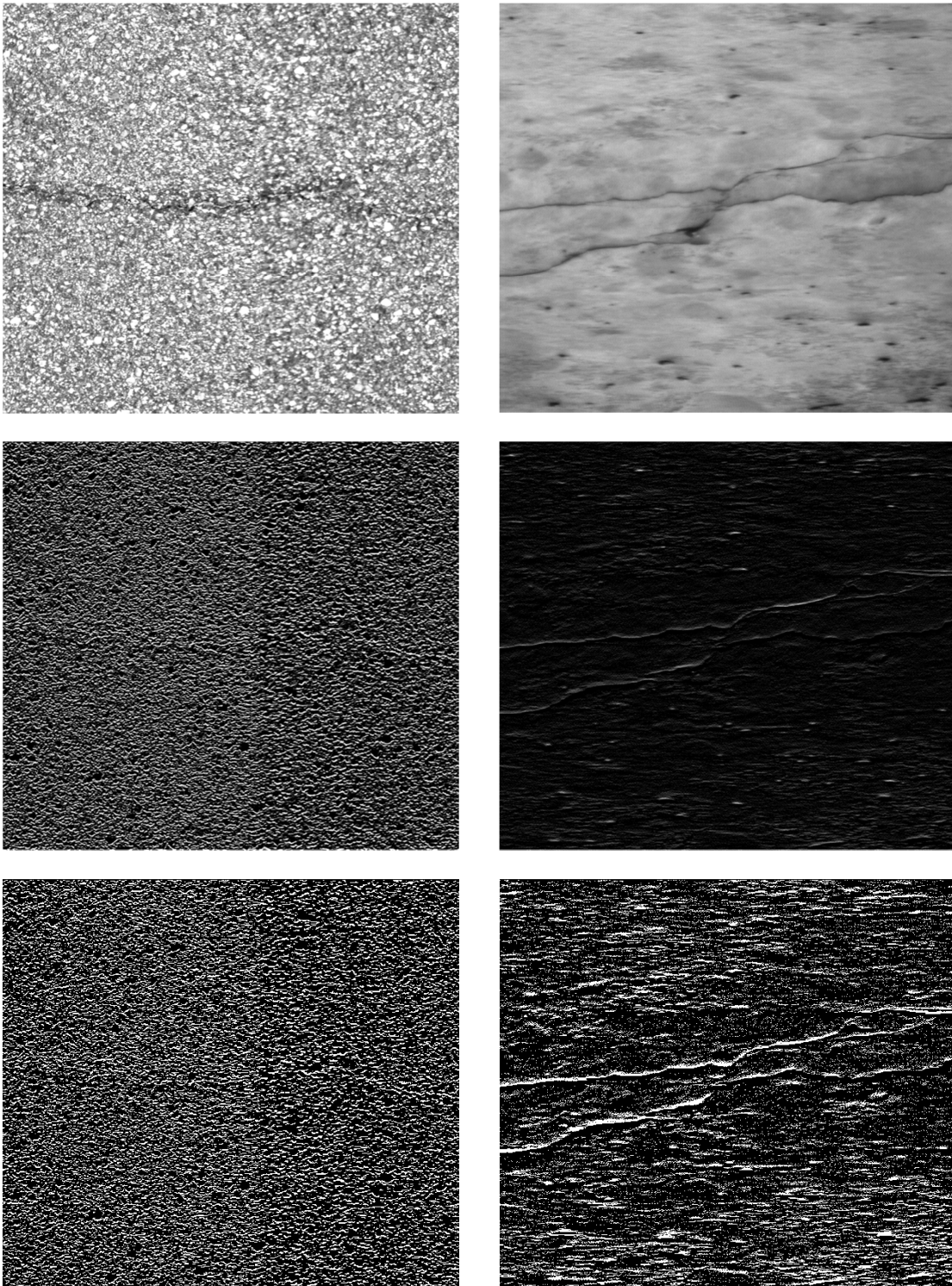


Figura 26. Dos imágenes filtradas por el Filtro Sobel. La columna de la izquierda es una grieta de carretera, la columna de la derecha es una grieta en un túnel. La fila superior es la imagen original preprocesada, la fila central es el resultado de aplicar el filtro de Sobel, la fila inferior es la imagen filtrada segmentada por el método Otsu

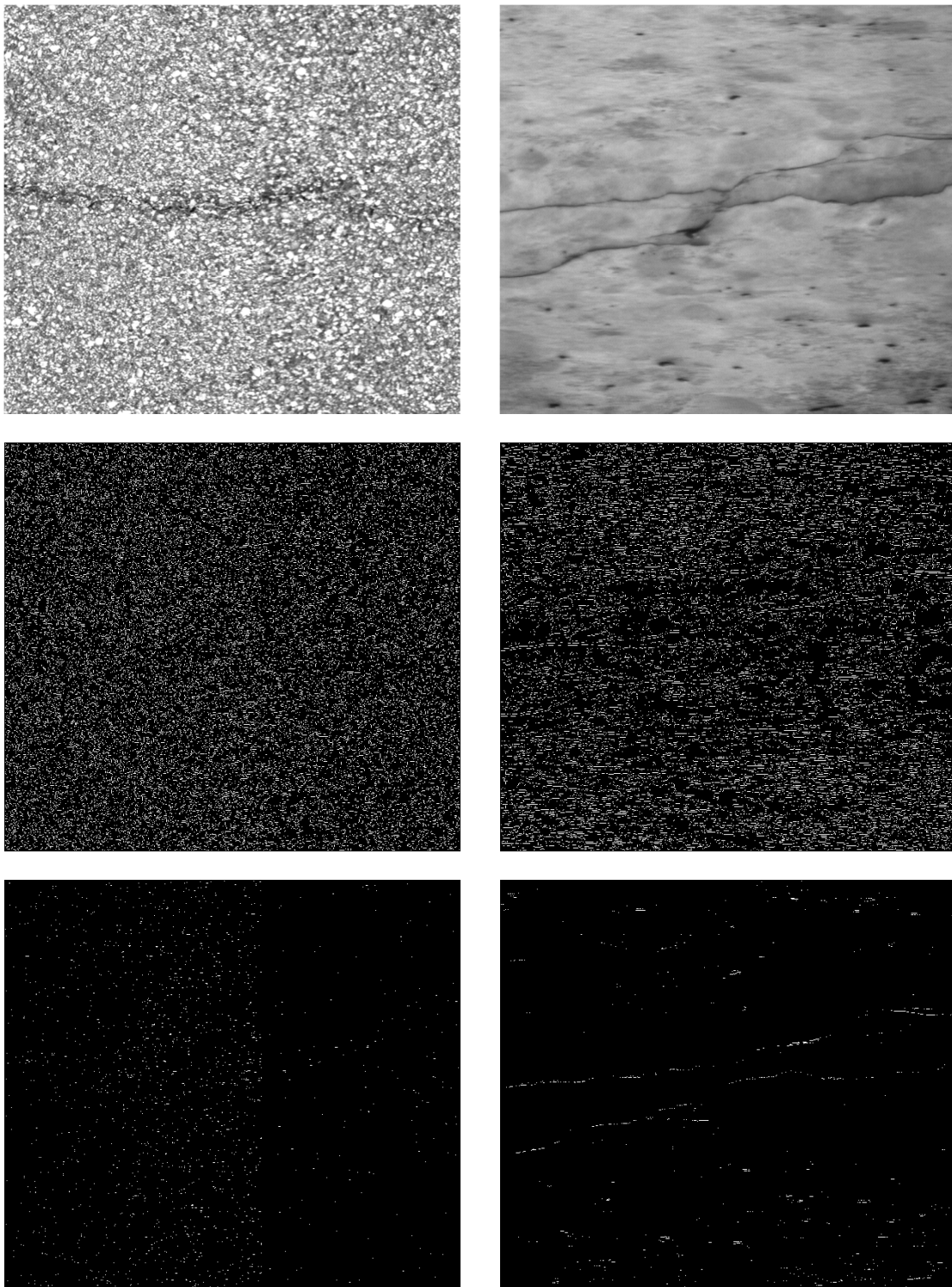


Figura 27. Dos imágenes de grietas en carretera y en túneles a los que se les han aplicado los filtros de Canny y Roberts.

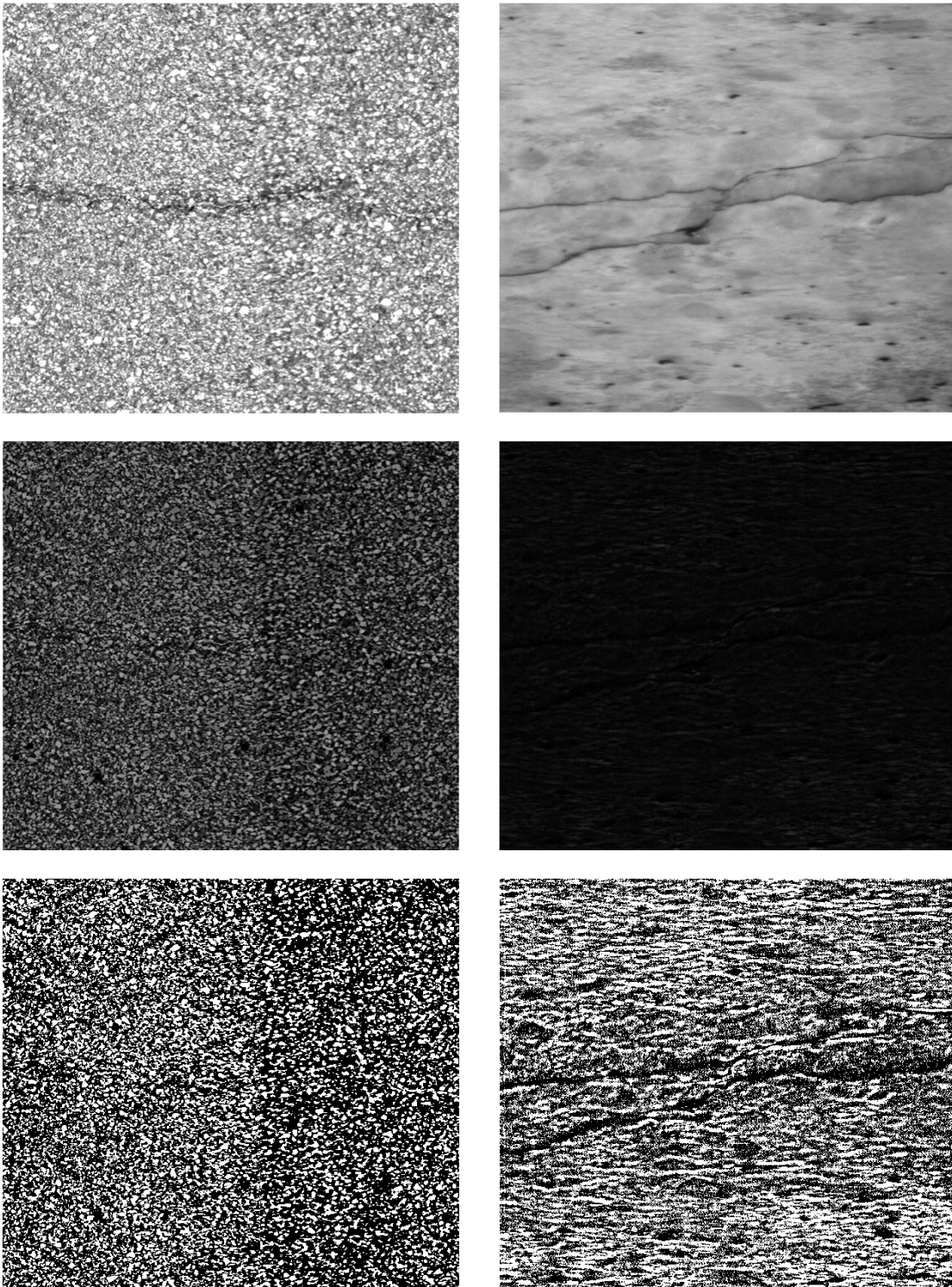


Figura 28. Dos imágenes filtradas por el filtro top-hat con un elemento estructural circular de radio 7. La columna de la izquierda es una grieta de carretera, la columna de la derecha es una grieta en un túnel. La fila superior es la imagen original preprocesada, la fila central es el resultado de aplicar el filtro top-hat, la fila inferior es la imagen filtrada segmentada por el método Otsu

Como puede deducirse al observar las anteriores imágenes a las que se les ha aplicado algunos de los métodos que tradicionalmente se utilizan para la segmentación de imágenes, no son lo suficientemente precisos como para ser aplicados para la segmentación o para la identificación de estos tipos de defectos.

En esta tesis se plantea la posibilidad de realizar el **proceso de segmentación utilizando información espacio-frecuencial localizada de la imagen**.

Por otro lado, la clasificación de defectos consiste en asignar a un tipo concreto de defecto cada una de las agrupaciones de píxeles pertenecientes al grupo de píxeles con defecto extraídas en la fase de detección. Como se puede ver a lo largo de la sección 2, se han aplicado numerosas técnicas para la clasificación de defectos en superficies continuas, como las que son objeto de estudio en esta tesis, entre las que destacan las redes neuronales, [Yaz08], [Pen09], [Mas12], el algoritmo de vecinos cercanos, [Dup97], y las máquinas de soporte vectorial, [Suv12], [Liu02]. En [Yun11] se comparan estos tres métodos aplicándolos a imágenes de chapa laminada, frente al algoritmo AdaBoost, el cual combina clasificadores binarios sencillos para obtener un único clasificador fuerte.

Independientemente del algoritmo de clasificación empleado, es necesario disponer de un conjunto de características que permitan diferenciar un tipo de defecto de otro. Con el objetivo de que todos los defectos presentes en la superficie a inspeccionar pasen a la fase de clasificación, las técnicas de detección también realizan agrupaciones de píxeles que realmente no son defectos o no pertenecen a ninguno de los tipos de defectos que se pretende clasificar. Estas agrupaciones de defectos se suelen llamar pseudodefectos y puede considerarse un tipo más de defecto que debería contemplarse en la fase de clasificación.

La forma habitual de caracterizar estas agrupaciones de defectos es utilizar su forma y su nivel de gris. Los defectos que se pretenden detectar en esta tesis, en muchas ocasiones, no presentan una forma y color característicos, por lo que estas técnicas no lograrían resolver el problema. Es por ello que se plantea la utilización de otro tipo de características que permitan describir mejor estos defectos, como puede ser la información espacio-frecuencial de la imagen. También se plantea la posibilidad de realizar el **proceso de detección y clasificación en único paso**, de tal manera que se aplique una técnica de detección específica para cada tipo de defecto, por lo que el resultado de cada detección solo sea un tipo de defecto, y por tanto ya se encuentre asignado a un grupo concreto de defectos.

En este contexto, el análisis de información espacio-frecuencial de las imágenes se plantea como una posibilidad que permita abordar la detección y clasificación de defectos superficiales desde una perspectiva diferente.

Desde que J. Fourier descubrió en 1.807 que cualquier señal arbitraria podía ser representada por la suma ponderada de funciones de senos y cosenos de diversas frecuencias, en el ámbito del análisis de señales se han propuesto multitud de técnicas basadas en el análisis en frecuencia con resultados satisfactorios, resultando difícil destacar alguna de ellas. Estas técnicas son también aplicables al procesamiento de imágenes, analizando la frecuencia espacial, aunque su implantación no ha sido habitual debido a su relativa lentitud de procesamiento que, hasta hace unos años, hacían imposible su incorporación en procesos productivos con altas cadencias.

El análisis de Fourier se puede entender como una técnica matemática de transformar la señal del dominio del tiempo al dominio de la frecuencia. En el caso de imágenes, esta técnica se extrapola a dos dimensiones espaciales, es decir, en lugar de analizar como varía una señal con el tiempo, se analiza como varía una señal (la imagen) en sus dos direcciones espaciales (el ancho y el alto de la imagen). El filtrado en el dominio de la frecuencia permite la modificación de las componentes de distintas frecuencias, resaltando o eliminando la información que contiene la imagen a unas determinadas frecuencias.

Para muchas señales, el análisis de Fourier es tremendamente útil porque la frecuencia contiene información de gran importancia. Sin embargo, el análisis de Fourier tiene un gran inconveniente, al transformar al dominio de la frecuencia, la información del tiempo o el espacio se pierde, por lo que es imposible determinar cuándo o dónde ocurre un determinado evento.

Si las propiedades de la señal no varían en gran medida a lo largo del tiempo o el espacio, es decir, la señal es estacionaria, este inconveniente no tiene demasiada importancia. Sin embargo numerosas señales son no estacionarias, o requieren el estudio de sus transitorios. Estas características son frecuentemente la parte más importante de las señales, y el análisis de Fourier no permite detectarlas. En el caso de detección de defectos en imágenes, determinar dónde se produce el cambio en frecuencia es el objetivo habitual del procesamiento de la imagen, por lo que el análisis de Fourier no suele aplicarse directamente.

Para corregir estas deficiencias, Denis Gabor en 1946 adaptó la transformada de Fourier para analizar solamente una pequeña porción de la señal (ventana) [Gab46]. Esta adaptación fue llamada *Short-Time Fourier Transform* (STFT), representando la señal mediante una función bidimensional en el tiempo y en la frecuencia.

La STFT representa un compromiso entre la visión de la señal en el tiempo y en la frecuencia. Proporciona información de cuándo y a qué frecuencias ocurre un evento en una señal. Esta información se puede obtener con ciertas limitaciones de precisión, que viene determinada por el tamaño de la ventana utilizada.

A pesar de que la STFT puede ser muy útil, presenta el inconveniente de que, una vez fijado el tamaño de la ventana, esa ventana es la misma para todas las frecuencias. Muchas señales requieren un tamaño de ventana variable para calcular con más precisión tanto el tiempo o la posición en el que un evento ocurre como la frecuencia.

El análisis de *wavelets* (en español ondículas, ondeletas u onditas), utiliza ventanas de tiempo de tamaño variable. Del mismo modo que el análisis de Fourier consiste en descomponer la señal en ondas sinusoidales de distintas frecuencias, el análisis de *wavelets* consiste en descomponer la señal en versiones desplazadas y escaladas de la ondícula original (*mother wavelet*).

Los filtros de Gabor pueden ser considerados detectores adaptados en orientación y en escala, lo que permite su utilización para la extracción de características locales de la imagen. Un banco de filtros de Gabor son un grupo de *wavelets*, de forma que cada una de ellas captura la energía a una determinada frecuencia y a una determinada orientación o dirección.

La forma en que el cerebro de los mamíferos procesa las imágenes que se forman en la retina, [Hub65], [Val82], es similar a la que se puede realizar al analizar las imágenes filtradas por un banco de filtros de Gabor.

La información existente sobre este tema indica que la forma en que se utilizan los filtros de Gabor para el procesamiento de imágenes es muy diferente dependiendo del objetivo buscado. Estos filtros tienen la capacidad de poder ser utilizados para detectar defectos que ocasionan un cambio en el patrón normal de la imagen, para resaltar ciertas características de las imágenes, para clasificar diferentes tipos de texturas, o para extraer características de las imágenes filtradas para usarlas como entradas a un clasificador.

Los avances tecnológicos de los últimos años permiten, cada vez más, utilizar estas técnicas en aplicaciones donde la velocidad de procesamiento es un factor limitante. Por ejemplo, recientemente, en [Mio13] se describe una técnica de procesamiento de imágenes para la clasificación del entorno de la carretera utilizando hardware dedicado que permite el procesamiento de 4 filtros de Gabor de 64 x 64 píxeles en 7ms.

En los siguientes apartados de esta sección se describe con más detalle cómo diseñar y utilizar los filtros de Gabor, prestando especial atención a la creación de bancos de filtros de Gabor y a la utilización de las imágenes filtradas para extraer información de las imágenes. Asimismo, se presenta una nueva metodología para el diseño de bancos de filtros de Gabor.

También se describe cómo utilizar el filtro de Gabor cuando los defectos que se quieren resaltar no tienen una dirección conocida, mediante lo que se ha denominado el filtro de Gabor invariante a la rotación.

Finalmente, también se propone una nueva técnica apropiada para la detección de defectos cuya longitud sea mucho mayor que su anchura, como es el caso de las grietas. Para ello, se combina el algoritmo CTA (*Conditional Texture Anisotropy*) con los filtros de Gabor invariantes a la rotación, proponiendo algunas mejoras en el algoritmo CTA original.

5.2 El filtro de Gabor

Un filtro de Gabor bidimensional consiste en una onda sinusoidal modulada por una función Gaussiana. El filtro realiza un análisis frecuencial de la señal localizada en posición y dirección. La formulación del filtro en el dominio espacial es la siguiente:

$$G_{\sigma,F,\theta}(x,y) = g_{\sigma}(x,y) \exp[j2\pi Fx']$$

donde $g_{\sigma}(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left[-\frac{1}{2}\left(\left(\frac{x'}{\sigma_x}\right)^2 + \left(\frac{y'}{\sigma_y}\right)^2\right)\right]$ **Ec. 19**

siendo $x' = x \cos \theta + y \sin \theta$ $y' = -x \sin \theta + y \cos \theta$

x, y son las coordenadas espaciales, F es la frecuencia central, θ es el ángulo que forma la dirección de la onda sinusoidal y el eje de abscisas. σ_x y σ_y son las desviaciones estándar de la función Gaussiana en la dirección de la onda y perpendicular a ella.

Dicho filtro en el dominio de la frecuencia es:

$$G_{\sigma,F,\theta}(u,v) = \exp\left[\frac{-1}{2}\left(\frac{(u'-F)^2}{\sigma_u^2} + \frac{v'^2}{\sigma_v^2}\right)\right]$$

$$\text{donde } \sigma_u = \frac{1}{2\pi\sigma_x}, \quad \sigma_v = \frac{1}{2\pi\sigma_y}$$

Ec. 20

$$\text{siendo } u' = u \cos \theta + v \sin \theta \quad \text{y} \quad v' = -u \sin \theta + v \cos \theta$$

u, v son las coordenadas espacio frecuenciales del filtro.

En la sección 7.2 se encuentran algunos ejemplos de filtros de Gabor que han sido utilizados en esta tesis en la **Figura 57**.

5.3 Diseño de bancos de filtros de Gabor

Un filtro de Gabor queda definido por 4 parámetros, F, θ, σ_x y σ_y , por lo que se pueden definir infinitos filtros de Gabor distintos entre sí. Un banco de filtros de Gabor es un subconjunto de todos los posibles filtros de Gabor que guardan cierta relación entre ellos, permitiendo analizar una zona del espacio de frecuencias y orientaciones. El diseño de un banco de filtros de Gabor consiste en la selección del conjunto apropiado de filtros que permitan analizar esa zona del espacio de frecuencias y orientaciones.

Es habitual que un banco de filtros de Gabor abarque todo el espacio de frecuencias y orientaciones. Para ello se va a realizar una transformación similar a la descomposición *wavelet*, donde la *wavelet* principal será un filtro de Gabor. Para ello se va a definir una metodología que permita realizar esa selección de forma sistemática. En primer lugar, se considera que las orientaciones, θ , de los filtros de Gabor que forman parte de un banco están separadas de forma uniforme y comprendidas entre 0 y 180°, de tal manera que se analiza todo el espacio de orientaciones. El ratio entre dos frecuencias centrales consecutivas, F , va a ser constante, de tal manera que las frecuencias seleccionadas estén separadas de forma uniforme. Dos filtros de distinta frecuencia son dos filtros idénticos pero escalados por el ratio de frecuencias, por lo que el ratio entre las desviaciones estándar de la función Gaussiana es el inverso al ratio de sus frecuencias. Por tanto, el diseño del banco de filtros se reduce a la selección de los siguientes parámetros:

1. La frecuencia central máxima, F_M , que es la frecuencia central más alta y se mide en píxeles⁻¹.
2. El valor de σ_x y σ_y para el filtro con la máxima frecuencia central.

3. El número total de frecuencias n_F .
4. El número total de orientaciones n_θ .
5. El ratio entre dos frecuencias centrales consecutivas k_F .

Al establecer el número de orientaciones, n_θ , y el número de frecuencias, n_F , se está determinando el número total de filtros que tiene el banco de filtros, que es el resultado de multiplicar ambos valores.

El ancho de banda de la orientación, B_θ , es π/n_θ , expresado en radianes, y es la mínima diferencia entre la orientación de dos filtros. Habitualmente, el ancho de banda de la orientación suele fijarse entre $\pi/8$ y $\pi/4$, es decir, se divide el espacio de orientaciones entre 4 y 8 orientaciones. Hubel y Wiesel, [Hub65], dedujeron que las células del cerebro que procesan la señal ocular son sensibles a determinadas orientaciones con un ancho de banda aproximado de $\pi/6$ radianes, valor que se encuentra comprendido entre los dos anteriores. Posteriores trabajos, como el realizado por De Valois et al, [Val82], establecen que existen células que procesan las imágenes con distintos anchos de banda, las cuales varían entre los 130° y los 26° , es decir, que se pueden llegar a distinguir hasta 7 orientaciones. Aumentar el valor del ancho de banda implica una mayor sensibilidad de los filtros ante la orientación a costa de aumentar el número de filtros total del banco de filtros y por tanto de aumentar el tiempo de procesamiento. Una mayor sensibilidad no implica necesariamente que se vayan a obtener mejores resultados en el procesamiento de imágenes.

En [Val82] también se analiza la respuesta que el cerebro de los mamíferos produce ante imágenes con distintas frecuencias espaciales, analizando la respuesta de distintas regiones del cerebro ante estímulos con distintas frecuencias. Concluyen que existen varias regiones del cerebro que responden ante distintas frecuencias espaciales con distinto grado de selectividad, y, aunque la mayoría tienen anchos de banda entre 1 y 1,5 octavas¹, existen regiones con un ancho de banda entre 0,5 y 2 octavas. Cuando se procesan imágenes mediante filtros de Gabor, la mayoría de los autores, [Cla05], [Ahm04], [Zha06], adoptan el intervalo de una octava para la frecuencia espacial, cuyo resultado es llamado *dyadic decomposition*. Esto implica que las frecuencias se obtienen dividiendo entre potencias de dos a la frecuencia central, es decir, el ratio de dos frecuencias consecutivas es 2, $k_F=2$. Sin embargo, otros autores, [She04], [Bia06], han utilizado el intervalo de media octava con lo que $k_F=2^{1/2}$. Cuanto mayor es este ratio, el banco de filtros realiza una mayor discriminación en frecuencias, pero a costa de aumentar el número de filtros del banco y por lo tanto del tiempo de procesamiento. Al igual que en el caso de la orientación, aumentar el

¹ El ancho de banda de la frecuencia, en octavas, desde la frecuencia 1 a la frecuencia 2 viene dado por $\log_2(f_1/f_2)$

número de frecuencias de un filtro no implica que el procesamiento de las imágenes mejore.

Según el teorema de muestreo de Nyquist-Shannon, [Nyq28] y [Sha49], para poder replicar con exactitud la forma de una señal periódica continua es necesario que la frecuencia de muestreo sea superior al doble de la máxima frecuencia de la señal. Por lo tanto, las imágenes deben haber sido adquiridas de tal manera que el periodo mínimo de la información contenida en éstas sea de al menos dos píxeles, ya que el periodo de muestreo es de un píxel, por lo que resulta que la frecuencia de Nyquist será $0,5 \text{ píxeles}^{-1}$. Por tanto, no se debe intentar extraer información de las imágenes de frecuencias superiores a la frecuencia de Nyquist. Por ello, la frecuencia central máxima, F_M , es seleccionada de tal manera que la curva de nivel de la mitad del valor de pico (valor máximo) del módulo de la transformada de Fourier (**half-peak magnitude iso-curve**), para la frecuencia máxima, tenga 0,5 como mayor valor de la frecuencia, que es la frecuencia de Nyquist. De esta manera ningún filtro tendrá una respuesta superior al 50% de su valor máximo para ninguna frecuencia por encima de la frecuencia de Nyquist.

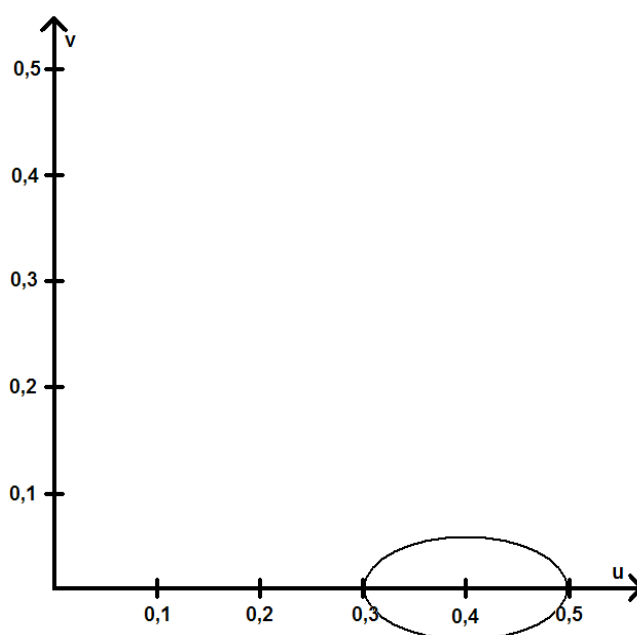


Figura 29. La elipse es la isocurva de valor de medio pico de la salida de un filtro de Gabor en el espacio de la frecuencia. Para cumplir el teorema de Nyquist-Shannon esta isocurva no debe superar el valor de frecuencia de 0,5

Para obtener la isocurva de nivel de medio pico se utiliza la ecuación del filtro de Gabor en el dominio frecuencial, **Ec. 20**. El valor máximo se alcanza en el punto del

espacio de frecuencias $(u', v') = (F, 0)$. En este punto $G(u, v)$ es igual a 1 para cualquier filtro de Gabor, por lo que el 50% será 0,5. La isocurva de nivel de medio pico está formada por todos los puntos (u, v) que cumplan la condición $G(u, v)=1/2$. En la **Figura 29** se puede ver la isocurva de nivel de medio pico de un filtro de Gabor. Se ha representado el filtro con orientación, $\theta=0^\circ$, para poder comprender mejor el problema. La frecuencia máxima de esta isocurva es 0,5, por lo que la salida del filtro por encima de frecuencia de Nyquist no supera nunca el valor de 0,5.

Entonces, si $G(u, v)=1/2$ se cumple

$$\frac{2\pi^2\sigma_x^2}{\ln 2}(u' - F)^2 + \frac{2\pi^2\sigma_y^2}{\ln 2}v'^2 = 1 \quad \text{Ec. 21}$$

que es la ecuación de una elipse cuyos semiejes, a, b , son

$$a = \frac{\sqrt{\ln 2}}{\sqrt{2\pi\sigma_x}} \quad b = \frac{\sqrt{\ln 2}}{\sqrt{2\pi\sigma_y}} \quad \text{Ec. 22}$$

Por lo tanto, la frecuencia máxima que alcanza la isocurva de valor de medio pico para el filtro con la frecuencia central más alta, F_M , es la suma de la frecuencia central más alta y el valor del semieje mayor de la elipse, a , de la ecuación de la elipse anterior. Ese valor se representa por F_{max} ,

$$F_{max} = F_M + \frac{\sqrt{\ln 2}}{\sqrt{2\pi\sigma_x}} \quad \text{Ec. 23}$$

Nuestro objetivo es determinar cuál es la frecuencia central máxima, F_M , que cumple que F_{max} sea igual a 0,5, por lo que sustituyendo en la ecuación anterior F_{max} por $1/2$,

$$F_M = \frac{1}{2} - \frac{\sqrt{\ln 2}}{\sqrt{2\pi\sigma_x}} = \frac{\pi\sigma_x - \sqrt{2\ln 2}}{2\pi\sigma_x} \quad \text{Ec. 24}$$

Es habitual la representación de estas curvas de nivel de la mitad del valor de pico del módulo de la transformada de Fourier de todos los filtros. Muchos autores consideran que estas curvas deben ser tangentes entre sí, [Jai90], [Ahm04], [Bia06], [Zha06]. Mientras otros autores, [Tur86], [Bia06], consideran cierto grado de solapamiento o distanciamiento entre los filtros. Cuanto mayor solapamiento haya

entre los filtros del banco mayor redundancia de información es extraída al aplicar el conjunto de filtros del banco de filtros.

El solapamiento radial es calculado considerando el solapamiento de las curvas de nivel de valor 0,5 del filtro de la más alta frecuencia y con una orientación de 0° y el filtro de la segunda más alta frecuencia y la misma orientación, como se representa en la **Figura 30 (a)**. Se denomina P_1 al menor valor radial de la curva de nivel de medio pico a la más alta frecuencia, y P_2 al mayor valor radial de la curva de nivel de medio pico de la segunda frecuencia más alta. El solapamiento radial es $S_R = P_2 - P_1$. Por tanto, cuando las curvas de nivel son tangentes entre sí en la dirección radial, el valor de R_θ es cero, siendo mayor de cero cuando no se solapen y menor de cero cuando haya solapamiento.

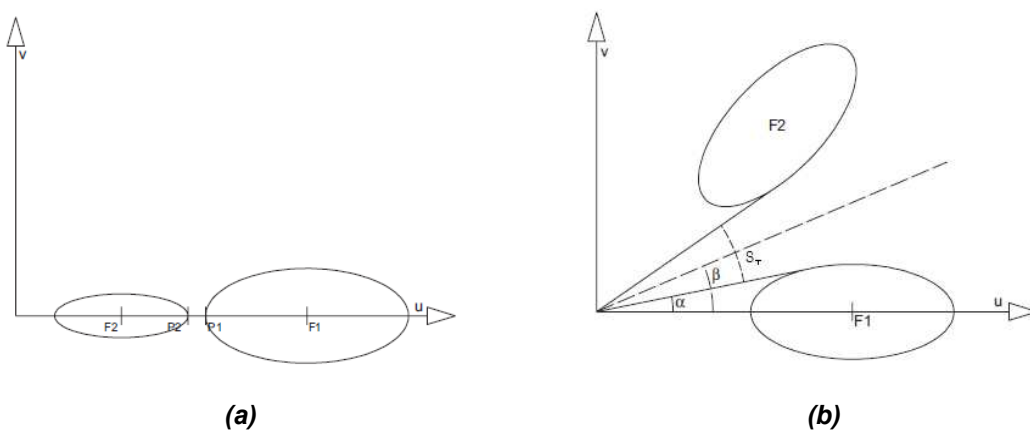


Figura 30. (a) Solapamiento radial. (b) Solapamiento tangencial.

El solapamiento tangencial se calcula considerando la superposición entre la curva de nivel del filtro con la más alta frecuencia y con orientación de 0° , que corresponde al Filtro 1, F_1 , y la del filtro con la más alta frecuencia y la segunda orientación π/n_θ , que corresponde al Filtro 2, F_2 , como se puede ver en la **Figura 30 (b)**. El ángulo α es el ángulo que forma la tangente con la curva de nivel de F_1 y el eje u . β se define como la mitad del ángulo entre dos orientaciones consecutivas, $\beta = \pi/(2n_\theta)$. El solapamiento tangencial será $S_T = 2(\beta - \alpha)$. Por tanto, cuando el valor de S_T es cero, los filtros se solapan en la dirección tangencial, no solapándose cuando este valor es menor de 0 y produciéndose solapamiento cuando es mayor de 0.

Una vez fijados los valores del ancho de banda de la frecuencia, B_F , en octavos, y el ancho de banda de la orientación, B_θ , en radianes, se pueden calcular los valores σ_x y σ_y , que consiguen que las isocurvas de nivel de medio pico sean tangentes entre sí. Recordemos que para un filtro de Gabor dado, el ratio entre las desviaciones estándar de la función gaussiana es la inversa de la relación entre sus frecuencias.

Para obtener el valor de σ_x se buscan los valores que hacen que los puntos P_1 y P_2 , que se pueden ver en la **Figura 30 (a)**, sean el mismo punto, es decir, $F_1 - a_1 = F_2 - a_2$, donde el subíndice 1 hace referencia al filtro de mayor frecuencia y el 2 al de la siguiente, F es la frecuencia central de cada uno de los filtros y a es el semieje mayor de la elipse que describe la isocurva de valor de medio pico calculado en la Ec. 22. Para calcular el valor de σ_y que hace que las isocurvas de valor de medio pico sean tangentes entre sí en la dirección tangencial, se debe hacer que la pendiente de la recta que pasa por el origen de coordenadas y que es tangente a la isocurva de nivel de medio pico sea la $\tan(\beta) = \tan(\pi/(2n_\theta))$, ya que, como se pudo ver en la **Figura 30 (b)**, este ángulo es la mitad del ángulo que hay entre dos filtros consecutivos. Resolviendo las ecuaciones planteadas se obtienen los siguientes resultados

$$\sigma_x = \frac{\sqrt{\ln 2} (2^{B_f} + 1)}{\sqrt{2\pi F} (2^{B_f} - 1)} \quad \sigma_y = \frac{\sqrt{\ln 2}}{\sqrt{2\pi F} \tan(B_g/2)} \quad \text{Ec. 25}$$

Por tanto, utilizando estos valores para calcular las desviaciones estándar de la función gaussiana que modula cada filtro de Gabor, las curvas de nivel de medio pico de los filtros de un banco de filtros de Gabor serían tangentes entre sí tanto en la dirección radial como en la tangencial, tal y como se puede ver en la **Figura 31**.

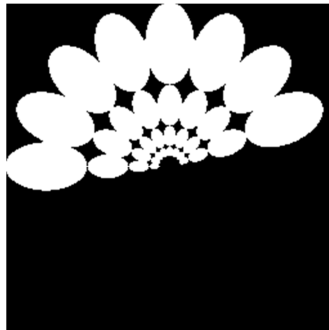


Figura 31. Representación de las curvas de nivel de medio pico. Las curvas de nivel de medio pico son tangentes entre sí tanto en la dirección radial como en la tangencial.

En la Ec. 25 los valores de las desviaciones estándar de la función gaussiana dependen de la frecuencia central, concretamente es inversamente proporcional al valor de frecuencia. Entonces, si se busca un mayor solapamiento entre los filtros se debe disminuir el valor obtenido en las ecuaciones anteriores o aumentarlo en caso contrario. Por tanto, se ha introducido una constante a cada una de estas ecuaciones, K_x y K_y , que regula el mayor o menor solapamiento de los filtros del banco de filtros de Gabor. Si el valor de estas constantes es menor de uno hay solapamiento y si es mayor de uno hay un hueco.

$$\sigma_x = K_x \frac{\sqrt{\ln 2} (2^{B_F} + 1)}{\sqrt{2\pi F} (2^{B_F} - 1)} \quad \sigma_y = K_y \frac{\sqrt{\ln 2}}{\sqrt{2\pi F} \tan(B_g/2)} \quad \text{Ec. 26}$$

Por otro lado, en la Ec. 24, el valor de la frecuencia central máxima, F_M , depende del valor de la desviación estándar de la función gaussiana en el eje de abscisas, σ_x , y el valor de desviación estándar para que las isocurvas sean tangentes entre sí dependen de la frecuencia central. Fijado el valor de K_x se puede sustituir el valor de σ_x de la Ec. 26 en la Ec. 24, y obtener el valor de la frecuencia central que hace que la curva de nivel de valor 0,5 del módulo de la transformada de Fourier del filtro de Gabor a la más alta frecuencia, tenga como frecuencia máxima 0,5.

$$F_M = \frac{K_x (2^{B_F} + 1)}{2(2^{B_F} (K_x + 1) + (K_x - 1))} \quad \text{Ec. 27}$$

Si K_x vale 1, que es el caso en el que las curvas de nivel son tangentes entre sí, la ecuación se simplifica a la expresión:

$$F_M = \frac{2^{B_F} + 1}{2^{B_F + 2}} \quad \text{Ec. 28}$$

Se ha definido una metodología para la generación de bancos de filtros de Gabor que permite la variación sistemática de los distintos parámetros que intervienen en su creación. Los parámetros que pueden variar son los siguientes:

1. La frecuencia central máxima, F_M , que es la frecuencia más alta y se mide en píxeles⁻¹.
2. El número total de frecuencias n_F .
3. El número total de orientaciones n_θ .
4. El ratio entre dos frecuencias consecutivas k_F .
5. Las constantes de solapamiento K_x y K_y .

Para la generación de un banco de filtros de Gabor se van a seguir los siguientes pasos:

1. Se fija el tamaño del filtro, $M \times N$. El número de filas y columnas debe coincidir para poder aplicar la transformada de Fourier al filtro de Gabor.
2. Se fija el número de orientaciones, n_θ , y el ratio entre dos frecuencias consecutivas, k_F . También se fijan las constantes de solapamiento, K_x y K_y , que indican el grado de solapamiento entre los distintos filtros en las dos direcciones, radial y tangencial.
3. El ancho de banda de las orientaciones, B_θ , se obtiene de dividir π radianes (180°) entre el número de orientaciones, n_θ , de esta manera todas las orientaciones son abarcadas. El ancho de banda en frecuencia, B_F , se define como $\log_2(k_F)$.
4. Se calcula el valor de la frecuencia central máxima, F_M , de tal manera que la curva de nivel de la mitad de valor de pico del módulo de la transformada de Fourier tenga 0,5 como mayor valor de la frecuencia, aplicando la Ec. 27.
5. Conocido el valor de la frecuencia central máxima, F_M , se calculan cada una de las frecuencias centrales, F , que van a ser utilizadas en el banco de filtros dividiendo por el ratio de frecuencias, k_F , tantas veces como el número de frecuencias que se desee que tenga el banco de filtros. La frecuencia más baja no debería ser inferior al inverso del tamaño de la imagen, N^{-1} , ya que no se pueden analizar frecuencias tan bajas con ese tamaño de imagen. Por tanto, el número de frecuencias, n_F , no debe ser mayor que el que permita alcanzar ese valor de frecuencia mínimo.
6. Con cada valor de la frecuencia central de cada filtro del banco de filtros de Gabor se calculan sus correspondientes valores de la desviación estándar de la función gaussiana, σ_x y σ_y .
7. Se genera cada uno de los filtros de Gabor con los valores obtenidos en los pasos anteriores.

El proceso se muestra esquematizado en la **Figura 32**. Se puede ver que en primer lugar se fijan los valores de $M=N$, n_θ , k_F , K_x y K_y , y se calcula el ancho de banda en orientación, $B_\theta = \pi/n_\theta$ y el ancho de banda en frecuencia $B_F = \log_2(k_F)$. A continuación se calcula la frecuencia central, F_M , cumpliendo las especificaciones impuestas, tanto de solapamiento entre filtros como de máxima frecuencia de los de las isocurvas de medio pico. El número de filtros, n_F , es fijado teniendo en cuenta que el filtro de más baja frecuencia no supere el valor de N^{-1} . Con estos datos se puede calcular cada uno de los valores de la frecuencia central, orientación y dispersiones necesarios para generar cada uno de los filtros.

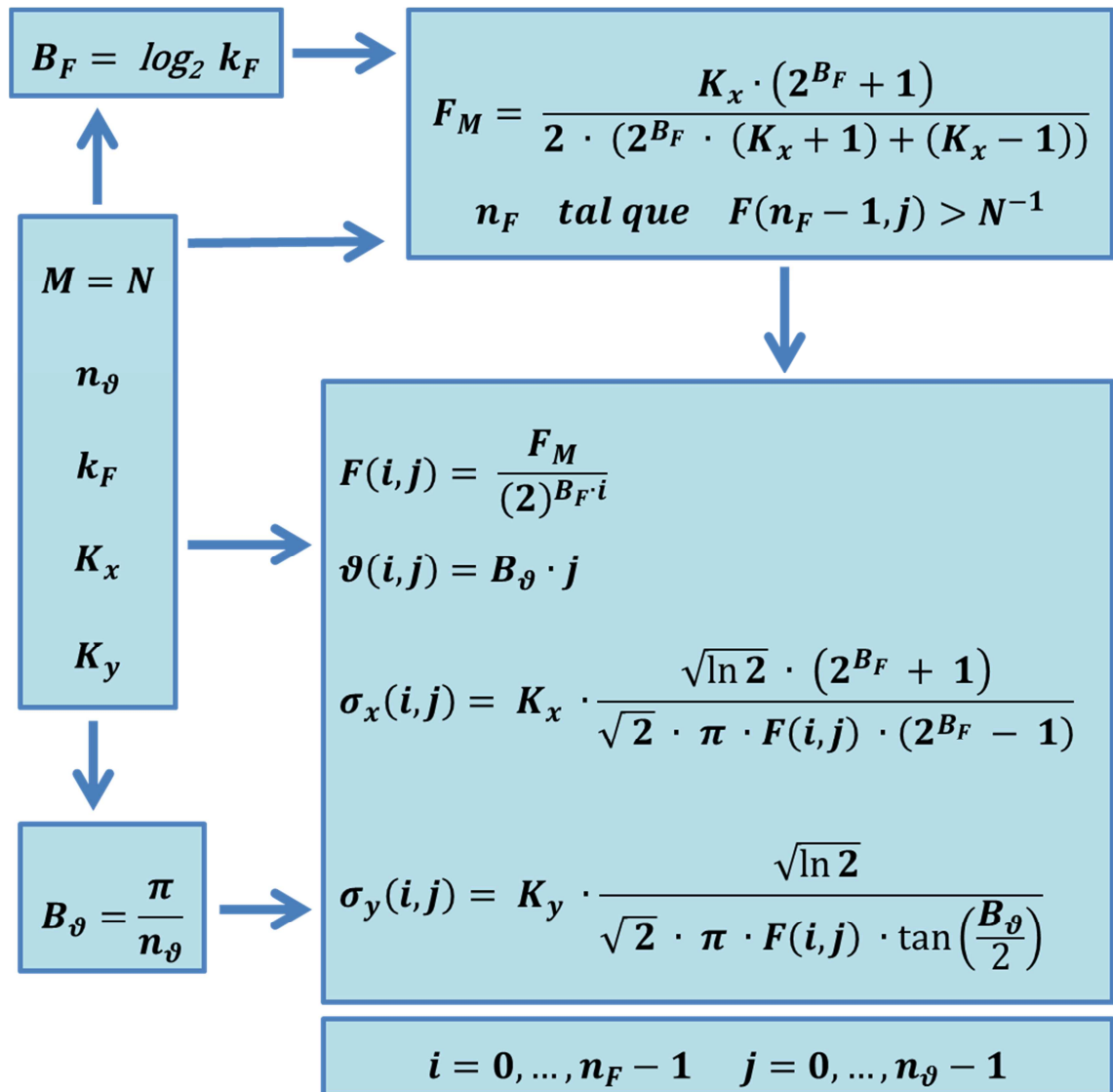


Figura 32. Esquema de la metodología definida para el cálculo de los parámetros de cada filtro de Gabor de un banco de filtros

De esta manera, se generan los $n_F \times n_\vartheta$ filtros de Gabor que forman el banco de filtros de Gabor sustituyendo cada uno de los valores $F(i, j)$, $\theta(i, j)$, $\sigma_x(i, j)$ y $\sigma_y(i, j)$ en la Ec. 19, que es la ecuación de un filtro de Gabor.

5.4 Análisis de las imágenes filtradas

Una vez diseñado el banco de filtros de Gabor se dispone de un conjunto de filtros de Gabor que se debe aplicar a toda la imagen, de tal manera que se obtengan tantas

imágenes filtradas como filtros de Gabor. Para conseguir altas velocidades de procesamiento se utiliza el teorema de la convolución.

$$I'(x, y) = I(x, y) * G(x, y) = \mathcal{F}^{-1}(I(u, v)G(u, v)) \quad \text{Ec. 29}$$

donde **I** representa la imagen, **G** el filtro, **I'** la imagen filtrada y \mathcal{F}^{-1} es la transformada inversa de Fourier.

Para poder aplicar el teorema de convolución, se deben dividir las imágenes que se van a filtrar en sub-imágenes del mismo tamaño que los filtros de Gabor. A cada sub-imagen se le aplica el teorema de convolución para cada uno de los filtros del banco de filtros de Gabor. Para ello se realiza la transformada de Fourier tanto de la sub-imagen como del filtro, se multiplican las matrices, y al resultado de esta multiplicación se le aplica la transformada inversa de Fourier. El resultado es la imagen filtrada en el dominio espacial. En la sección 7, donde se muestran los resultados obtenidos en esta tesis, se muestran en la **Figura 53** y en la **Figura 66** algunos ejemplos del resultado de aplicar determinados filtros de Gabor a imágenes obtenidas en el marco de esta tesis.

Cada una de las imágenes filtradas, **I'**, es una matriz de números complejos, por lo que esta imagen filtrada ofrece un amplio abanico de posibilidades a la hora de su análisis. Por un lado, tanto su parte real como su parte imaginaria pueden ser analizados aisladamente. Por otro lado, el módulo y la fase pueden ser calculados y procesados.

A cada una de estas cuatro imágenes, (parte real, parte imaginaria, módulo y fase), se les pueden aplicar diferentes operaciones, tal y como se haría con la imagen original no filtrada. Por un lado, se pueden utilizar algoritmos para la segmentación de imágenes, así como extraer características de las imágenes, como su valor máximo, mínimo o medio, o estadísticos como la desviación estándar, momentos centrales, etc., para ser utilizadas como entradas de un clasificador.

La utilización de todos los filtros de un banco de filtros, tal y como se ha definido en esta sección, permite extraer toda la información frecuencial contenida en la imagen, que puede ser útil para reconstruir la imagen, o comprimirla. Sin embargo, para otros objetivos, como la segmentación o clasificación de imágenes, no tiene por qué ser necesario la utilización de todos los filtros, sino solo aquellos que proporcionen información relevante que permita resolver el problema objetivo.

Tanto el diseño de los bancos de filtros, la selección de los filtros y la forma en que se analizan las imágenes filtradas son procesos interrelacionados, que no pueden

abordarse de forma independiente, y que dependen tanto de las imágenes que se vayan a analizar como de cuál sea el objetivo de ese análisis.

5.5 Filtro de Gabor invariante a la rotación

Un filtro de Gabor analiza la imagen en frecuencia y orientación. Esto implica que si se rota una imagen un determinado ángulo no se obtiene la misma respuesta del filtro que el obtenido con la imagen original. Esta propiedad del filtro de Gabor, que puede ser utilizada para discriminar imágenes, también puede ser desfavorable para ciertos tipos de análisis de imágenes. En multitud de ocasiones el objetivo del procesamiento de imágenes es detectar un objeto o anomalía en una superficie con independencia de su orientación.

Algunos autores ya han tratado ese problema. En [Xie08] se aplica un banco de filtros de Gabor a la imagen completa. Ésta se divide en sub-imágenes, y para cada sub-imagen se calcula la media y la desviación del módulo de la imagen filtrada. Los resultados se ponen en dos matrices, una para la media y otra para la desviación, en las que en cada columna están los resultados de los filtros con la misma orientación, y en cada fila los de la misma frecuencia. A continuación se busca el mayor valor de la media, y se desplazan las filas hasta situar ese valor en la primera columna. Esta matriz de resultados modificados es la que se utiliza para compararla con una base de datos y clasificar texturas de Brodatz.

J. Zhang propuso en [Zha02] una nueva versión del filtro de Gabor para convertirlo en filtro con simetría circular. Para ello modificó las ecuaciones del filtro tal y como se muestra en la siguiente ecuación:

$$G_{\sigma,w}(x,y) = g_{\sigma}(x,y) \exp\left[j2\pi F\left(\sqrt{x^2 + y^2}\right)\right]$$
$$\text{donde } g_{\sigma}(x,y) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{x^2 + y^2}{2\sigma^2}\right] \quad y \quad j = \sqrt{-1} \quad \text{Ec. 30}$$

donde $\sigma = \sigma_x = \sigma_y$. El resultado de la convolución de este filtro con una imagen es invariante a rotaciones en la imagen

En [Ria13] también se describe un método para aplicar filtros de Gabor invariantes a la orientación y a la frecuencia. El método consiste en utilizar un banco de filtros de Gabor con varias orientaciones y frecuencias. Se aplica todo el banco de filtros a cada imagen y se calcula la media y la desviación estándar del módulo de cada imagen

filtrada. Los resultados se colocan en dos matrices, una para la media y otra para la desviación estándar, y se les aplica la transformada discreta de Fourier de dos dimensiones (2D DFT). Los coeficientes de la transformada discreta de Fourier son invariantes respecto a desplazamientos, por lo que un cambio de orientación o de frecuencia, que supone un desplazamiento en los resultados de las matrices, no supone un cambio en los coeficientes de la DFT. Estos coeficientes son los utilizados para realizar la clasificación, aplicando este método a imágenes de la base de datos de texturas de Brodatz, a distintos escenarios de rotación y escala.

En la presente tesis se propone un nuevo método para abordar el problema. La principal aplicación del método que se va a proponer a continuación es la segmentación de defectos superficiales sin un conocimiento previo de la orientación de los defectos que se pretende detectar. En la sección 7.3 se muestra en la **Figura 69** dos imágenes obtenidas al aplicar este método a la detección de grietas en la superficie de túneles de hormigón.

Tal y como se describe en la sección 5.2, un filtro de Gabor queda definido por 4 parámetros, F , θ , σ_x y σ_y . El objetivo del método que se propone en esta sección es buscar los valores de F , σ_x y σ_y que provocan que la imagen filtrada tenga en alguna de las orientaciones un valor elevado en el caso de que haya un defecto en la superficie y un valor bajo en todas ellas en el caso de que no lo haya.

Para ello se fija inicialmente el número de orientaciones, n_θ , se selecciona el valor de F , σ_x y σ_y y se calcula la imagen filtrada para las n_θ orientaciones. A partir de estas imágenes filtradas se forma una nueva imagen filtrada en el que a cada píxel se le asigna el valor máximo de las n_θ imágenes filtradas inicialmente. Es a esta imagen a la que se le aplicará el apropiado algoritmo de selección del umbral más apropiado para su segmentación.

La elección de los parámetros F , σ_x y σ_y de los filtros de Gabor es fundamental para obtener unos buenos resultados. Para ello se ha elegido el método de optimización denominado **Evolución Diferencial**, introducido por primera vez por Storn y Price, [Sto97], un algoritmo genético que permite alcanzar buenos resultados de convergencia, evitando en gran medida caer en mínimos locales. Existen varias variantes de este mismo algoritmo para mejorar la velocidad de convergencia y evitar caer en mínimos locales. Para aplicar este método es necesario disponer de una **base de datos de imágenes segmentadas manualmente**, seleccionar un **método para el cálculo del umbral en la segmentación**, como los que se describen en la sección 4.3, y seleccionar un **método de evaluación de los resultados obtenidos**, como los que se describen en la sección 4.2.1.

A continuación se describe el algoritmo genético **Evolución Diferencial** que se emplea en este trabajo. Se trata de una modificación respecto al trabajo presentado por Storn y Price, en el que la fase de mutación se realiza de forma diferente:

- 1. Inicialización:** Generar N_p vectores de forma aleatoria dentro de unos valores acotados. Cada vector está compuesto por las variables de la función que se quiere minimizar, es decir, los parámetros del método de segmentación de imágenes que se quieren optimizar. Se aplica el método de segmentación a las imágenes de la base de datos y se evalúa el resultado para cada uno de esos vectores.
- 2. Mutación:** Se crean N_p nuevos vectores aplicando la siguiente formula:

$$v_{i,g} = x_{best,g} + F(x_{r1,g} - x_{r2,g}) \quad \text{Ec. 31}$$

donde i indica el número de vector, g el número de generación, v es el vector mutado, x es el vector sin mutar, el subíndice *best* se refiere al vector que mejor resultado obtuvo, $r1$ y $r2$ son valores aleatorios enteros entre 0 y $N_p - 1$ y distintos entre sí, y F es un vector de escala de la mutación. El algoritmo original de Storn y Price utilizaba otro vector aleatorio en lugar del mejor y no se permitía que ninguno de los tres coincidiera con el vector número i .

- 3. Recombinación:** En este paso se recombinan el vector mutado con el vector sin mutar aplicando la siguiente formula:

$$u_{ji,g} = \begin{cases} v_{ji,g} & \text{si } \text{rand}_j[0,1] \leq CR \\ x_{ji,g} & \text{en otro caso} \end{cases} \quad \text{Ec. 32}$$

donde j es el número de variable, u es el vector recombinado y CR es la probabilidad de que mute una variable. Se debe garantizar que al menos una variable sea mutada por lo que se muta una inicialmente y se aplica la probabilidad al resto.

- 4. Selección:** Se evalúan los vectores recombinados y se asigna a la siguiente generación el mejor entre el vector original y el recombinado.

$$x_{i,g+1} = \begin{cases} u_{i,g} & \text{si } f(u_{i,g}) < f(x_{i,g}) \\ x_{i,g} & \text{en otro caso} \end{cases} \quad \text{Ec. 33}$$

donde la función f es el método de evaluación de los resultados seleccionado. A continuación se repite el proceso desde el punto 2 hasta que se alcance un criterio de parada.

Los resultados del algoritmo mejoran al modificar la fase de inicialización (punto 1) del proceso anteriormente descrito de tal forma que en vez de partir de un grupo de vectores seleccionados al azar, se utiliza un banco de filtros de Gabor, calculado de la forma que se describe en la sección 5.3. De esta manera se consigue reducir la posibilidad de caer en mínimos locales y se reduce el número de generaciones necesarias para alcanzar el criterio de parada.

5.6 La textura de anisotropía condicional en imágenes filtradas por filtros de Gabor

El concepto de textura de anisotropía condicional (*Conditional Texture Anisotropy*, CTA) fue introducido por primera vez por F. Roli, [Rol96]. La idea era buscar una medida que tuviera un alto valor en una dirección y un valor muy bajo en el resto. Esta característica fue utilizada posteriormente por Nguyen et al. para detectar grietas en el pavimento de carreteras, [Ngu09].

En esta tesis se analiza la implementación realizada por Nguyen et al. y se propone una variación de este algoritmo para la detección de grietas. En la sección 7.3 se muestran los resultados obtenidos al aplicar las técnicas propuestas a grietas en túneles. Las principales innovaciones que se introducen en el algoritmo original consisten en proponer una nueva función de probabilidad en el cálculo de la CTA y aplicar los cálculos a las imágenes filtradas por un filtro de Gabor invariante a la rotación en lugar de aplicarla directamente sobre la imagen capturada.

El algoritmo CTA propone el cálculo de un vector de características a lo largo de diferentes direcciones para evaluar la probabilidad de que el defecto buscado se presente en esas direcciones. Para el cálculo del valor CTA de un determinado píxel se utiliza la siguiente ecuación:

$$CTA(X^l) = \frac{\max_j \{p(x_j^l / l \in w_1)\} - \min_j \{p(x_j^l / l \in w_1)\}}{\max_j \{p(x_j^l / l \in w_1)\}} \quad \text{Ec. 34}$$

Siendo X^l un vector de m características evaluado en k direcciones a lo largo de una línea de longitud l centrada en cada píxel. W_1 se utiliza para referirse a la

superficie que no contiene grietas, por lo que $p(x_j^l / l \in w_1)$ es la probabilidad de que los píxeles de la línea con dirección j no sean grietas.

El valor de CTA tendrá un valor próximo a 0 en las regiones sin grietas, y próximo a 1 cuando haya grietas. Para poder aplicar el algoritmo anterior es necesario conocer la probabilidad de que un pixel de la imagen no sea un defecto. En Nguyen et al. se propone que dicha probabilidad sea calculada aplicando la siguiente ecuación:

$$p(x_j^l / l \in w_1) = \frac{D - \Delta_j}{D} \quad \text{Ec. 35}$$

siendo $D = \|x\|_2$ y $\Delta_j = \|x - x_j^l\|_2$

siendo x el valor medio de las características seleccionadas en toda la imagen. El algoritmo de Nguyen et al. propone la utilización de la media y la desviación estándar del nivel de gris de la imagen como vector de características. Este algoritmo se utiliza en esta tesis para la detección de grietas en túneles de hormigón. Los resultados obtenidos se encuentran en la sección 0, donde se analiza el comportamiento del método y se muestran algunas imágenes procesadas en la **Figura 70**.

La aplicación de dicho algoritmo es un proceso lento debido a la necesidad de calcular la desviación estándar de los valores de nivel de gris de los píxeles de la imagen de cada línea. Con el objetivo de reducir el tiempo de procesamiento y mejorar los resultados obtenidos se han propuesto modificaciones sobre el algoritmo original de Nguyen et al. En primer lugar, se propone el cálculo de las características en un entorno circular de la imagen lo suficientemente grande, que se denotaría por x^c . El tamaño de la circunferencia en las que se calcula el valor medio se establece en 255 píxeles. Se ha excluido la característica de la desviación estándar del vector de características, manteniendo solo la media, que se denota por m . Al utilizar una única variable, la formulación se simplifica significativamente, de tal forma que para cada píxel $D_j = m^c$, el superíndice c indica que se aplica en un entorno circular de cada píxel, y $\Delta_j = |m^c - m_j^l|$, es decir, para cada dirección j , Δ_j es el valor absoluto de la diferencia entre la media de la imagen filtrada en un entorno circular de cada píxel y ese mismo valor en una línea. Sustituyendo en la Ec. 34, se obtiene el siguiente resultado:

$$CTA(X^l) = \frac{\max_j \left\{ \frac{m^c - |m^c - m_j^l|}{m^c} \right\} - \min_j \left\{ \frac{m^c - |m^c - m_j^l|}{m^c} \right\}}{\max_j \left\{ \frac{m^c - |m^c - m_j^l|}{m^c} \right\}} = \frac{\max_j \{m^c - m_j^l\} - \min_j \{m^c - m_j^l\}}{m^c - \min_j \{m^c - m_j^l\}} \quad \text{Ec. 36}$$

El resultado final de aplicar esta técnica es una imagen en la que para cada píxel se obtiene el valor CTA calculado en la Ec. 36. Esta imagen es segmentada para identificar qué píxeles corresponden a las grietas que píxeles corresponden a las zonas sin defecto. Para calcular el umbral de segmentación se debe aplicar una técnica de segmentación de imágenes como las que se describen en la sección 4.3.

La aplicación de esta modificación reduce el tiempo de procesamiento a menos de la sexta parte, obteniendo resultados similares, tal y como se detalla en la sección 0, donde también se muestran algunas imágenes con el resultado de aplicar el filtro en la **Figura 71**. En este caso es el tiempo de cálculo de la media en un entorno circular la operación que consume mayores recursos. Por otro lado, se ha detectado que el cálculo de la probabilidad propuesto de Nguyen et al. presenta algunos inconvenientes. Por ello, se propone un nuevo enfoque para el cálculo de la probabilidad de que una zona de la imagen esté libre de defectos y eliminando la necesidad de calcular el valor medio en un entorno circular de la imagen.

La ecuación para el cálculo de la probabilidad propuesta por Nguyen et al. presenta algunos inconvenientes que desaconsejan su uso. Se puede observar que cuando el valor de Δ_j es mayor que el valor de D , la probabilidad, $p(x_j^l / l \in w_1)$, es menor que cero, lo cual no tiene sentido físico. Esto ocurre cuando el valor medio en una línea es superior al doble del valor medio en toda la imagen. Por ello, se propone en este trabajo utilizar una nueva ecuación de probabilidad cuyos valores estén restringidos al entorno (0, 1). Entre las distintas alternativas que cumplen esta condición, la función sigmoidea es la que mejor se adapta a nuestros objetivos:

$$p(x_j^l / l \in w_1) = \frac{1}{1 + e^{-t(m_j^l, m, \sigma)}} \quad \text{Ec. 37}$$

siendo $t(m_j^l, m, \sigma)$ la variable que se quiere evaluar y que debe ser función de x_j^l , que en este caso es el valor medio del nivel de gris, m_j^l y del valor medio del nivel de gris la imagen, m , y su desviación estándar, σ . La función sigmoidea se caracteriza por tener valores próximos a cero cuando t tiene valores negativos y valores próximos a 1 cuando tiene valores positivos, mientras que en torno a cero existe una zona de aceleración intermedia.

Para la definición de la función de t se debe tener en cuenta qué características tienen los defectos que se quieren detectar. El nivel de gris de una grieta se caracteriza por tener un valor por debajo de la media de la imagen en la dirección de la grieta, mientras que en el resto de las direcciones su valor debe ser similar a la media o superior. Se considera que el valor en torno al cual se debe producir la aceleración de la curva del sigmoide debe ser el valor medio de la imagen menos dos

veces y media el valor de la desviación estándar, y que la ecuación debe tener un factor de escala del 25% del valor medio del nivel de gris de la imagen, de tal forma que la función t queda definida de la siguiente forma:

$$t(m_j^l, m, \sigma) = \frac{m_j^l - (m - 2.5 * \sigma)}{m / 4} \quad \text{Ec. 38}$$

donde m_j^l es el valor medio del nivel de gris a lo largo de una línea de longitud l en la dirección j , m el valor medio del nivel de gris en toda la imagen, y σ la desviación estándar del nivel de gris en toda la imagen.

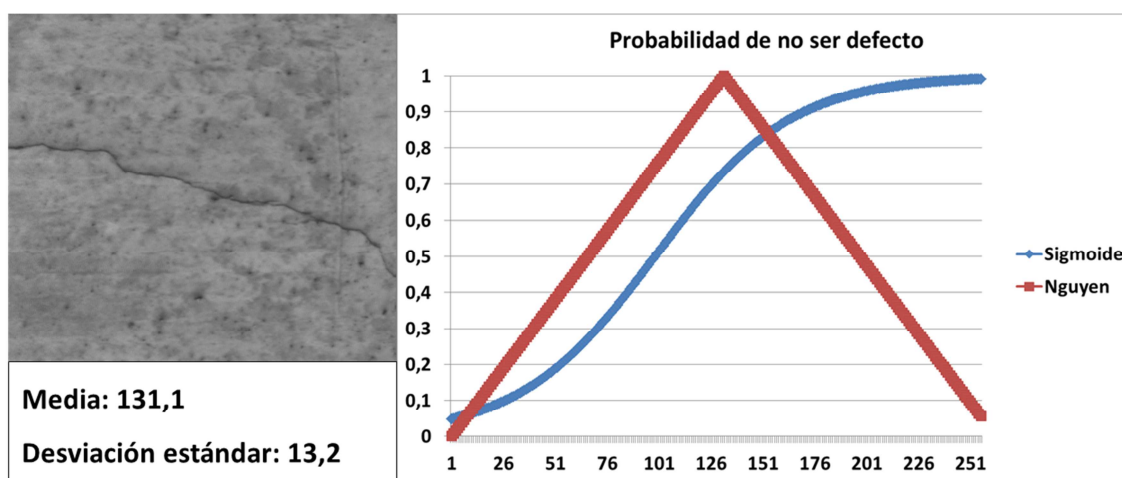


Figura 33. Probabilidad de que una línea de la imagen sea defecto dependiendo del nivel de gris medio de la imagen. En rojo la ecuación propuesta por Nguyen et al. en [Ngu09], en azul la ecuación propuesta en este trabajo

En la **Figura 33** se representa la función de probabilidad de que un pixel pertenezca a una grieta según la ecuación de Nguyen et al. frente a la propuesta en este trabajo. La función propuesta por Nguyen et al. presenta un máximo en el valor de la media y desciende de forma simétrica a partir de este valor mientras que en la función propuesta en este trabajo es creciente, de tal forma que la probabilidad de no ser grieta aumenta con el valor del nivel de gris, tal y como ocurre en realidad en las imágenes disponibles. Se ofrecen más detalles en la sección 0 junto a algunas imágenes en la **Figura 72**, donde se aplica este método.

Otra de las ventajas del algoritmo propuesto es el aumento de la velocidad de procesamiento, siendo aproximadamente 60 veces más rápida su ejecución frente al algoritmo anterior.

La siguiente modificación propuesta consiste en aplicar el algoritmo propuesto anteriormente sobre la salida de la imagen filtrada por el filtro de Gabor invariante a la rotación, descrito en la sección 5.5. Para su aplicación, se requiere una adaptación que se describe a continuación.

El valor del módulo de Gabor siempre es mayor o igual que 0. Se busca que al procesar una imagen sin defectos la salida del filtro de Gabor en cualquier dirección este próxima a 0 y que cuando exista una grieta la imagen filtrada alcance el mayor valor posible en la dirección de la grieta. Por tanto, para que la probabilidad, $p(x_j^l / l \in w_l)$, cuando una imagen no contenga defectos tienda a cero, se debe modificar la función sigmoide para que presente un valor alto cuando la salida del filtro de Gabor sea próxima a cero y un valor bajo cuando la salida del filtro de Gabor sea elevada. Por tanto, la función de probabilidad se expresa de la siguiente forma:

$$p(x_j^l / l \in w_l) = \frac{1}{1 + e^{+t(m_j^l, m, \sigma)}} \quad \text{Ec. 39}$$

Ecuación análoga a la Ec. 37 con el signo de la exponencial positivo. El valor de $t(m_j^l, m, \sigma)$ se calcula utilizando la Ec. 38, teniendo en cuenta que los valores de las medias y la desviación estándar son los de la imagen filtrada por el filtro de Gabor invariante a la rotación.

Aplicando esta nueva definición de probabilidad en la ecuación de la CTA, Ec. 34, en las zonas donde haya una grieta en la imagen, la probabilidad de que no haya una grieta en alguna de las direcciones es baja, mientras que en el resto de las direcciones la probabilidad de que no haya grieta es elevada, y por tanto el valor de CTA se encontrará próximo a 1. Sin embargo, en las zonas de la imagen que no haya grietas, las probabilidades de que no haya grietas en cualquiera de las direcciones serán altas y próximas entre sí, por lo que el valor de CTA será próximo a cero.

El cálculo de los parámetros del filtro de Gabor se realiza utilizando la misma técnica empleada en la sección 5.5, con la única diferencia que la función f , que evalúa los resultados obtenidos, se calcula aplicando a las imágenes empleadas para el entrenamiento el algoritmo descrito en esta sección.

La aplicación de este algoritmo supone un incremento significativo del tiempo de procesamiento ya que requiere del cálculo de filtros de Gabor y la aplicación de un algoritmo iterativo para ajustar los parámetros del filtro de Gabor, aunque esta operación solo se aplica en la fase de entrenamiento. En la sección 7.3.3 se ofrecen los resultados de aplicar esta técnica para la detección de grietas en túneles de hormigón, mostrándose algunos ejemplos en la **Figura 74**.

5.7 Conclusiones

En este capítulo se propone la evaluación de imágenes filtradas por filtros de Gabor como alternativa al uso de características relacionadas con la distribución del nivel gris, las cuales no son válidas para el procesamiento de algunos de los defectos que aparecen sobre las superficies propuestas en esta tesis. Estos filtros realzan las **anomalías de la imagen en el dominio espacio-frecuencial** de forma localizada, tratando de emular la forma que el cerebro de los mamíferos procesa las imágenes.

Para utilizar estos filtros es necesario ajustar un conjunto de parámetros que se adecúen al problema específico que se quiere resolver. El **diseño de bancos de filtros** de Gabor permite analizar una zona completa del espacio de orientaciones y frecuencias con un conjunto de filtros de Gabor. Se ha propuesto una metodología para crear estos bancos de filtros de forma eficiente y para la extracción de características utilizando estos filtros. Estas características son utilizadas para aplicar las técnicas de detección y clasificación de defectos descritas en el Capítulo 4.

También se han propuesto algunas variantes al filtro de Gabor para la detección de defectos cuya orientación no es conocida a priori. También se han propuesto mejoras respecto a otros trabajos que afectan tanto al tiempo de procesamiento como a su eficacia.

En el Capítulo 7 se muestran los resultados de aplicar las técnicas de procesamiento de imágenes propuestas para resolver las dificultades encontradas en la inspección de tres tipos de superficies en entornos reales de funcionamiento, cuya implementación se describe en el Capítulo 6.

6 SISTEMAS DE INSPECCIÓN IMPLEMENTADOS

6.1 Introducción

En esta sección se describen los sistemas de inspección implementados en esta tesis. Se ha implementado un sistema de inspección superficial de bobinas de chapa laminada, un sistema de inspección del estado de la superficie de la carretera y un sistema de inspección de grietas y fisuras en túneles.

El primero de ellos se encuentra instalado en un entorno industrial, en una empresa dedicada al corte de acero laminado. Ofrece resultados en tiempo real interactuando con la línea de producción.

El segundo es un sistema de inspección del estado de la superficie de la carretera que utiliza para la adquisición de las imágenes el equipo comercial LRIS. La información proporcionada por este equipo se complementa con una serie de dispositivos cuya información se captura de forma sincronizada. Se ha desarrollado el *software* de adquisición, procesamiento y visualización, así como la electrónica apropiada para la captura sincronizada de todos los dispositivos. La captura de datos se realiza a la velocidad de la vía, almacenándola para su posterior procesamiento.

Por último, se ha implementado un prototipo para la inspección de la superficie de túneles con el objetivo de detectar grietas y fisuras. Se trata de una plataforma móvil cuyo fin es verificar la viabilidad de la detección, capturando una pequeña porción del túnel. Se ha realizado el diseño y se ha desarrollado el *software* necesario para poder realizar la inspección completa de distintas tipologías de túnel y resoluciones de inspección requeridas.

6.2 Inspección superficial de bobinas de chapa laminada

Para desarrollar el sistema de inspección de chapa laminada se ha seleccionado el sistema de adquisición de imágenes, las fuentes de iluminación, los diferentes interfaces de comunicación y los equipos de procesamiento de imágenes. Así mismo se han establecido los procedimientos de adquisición de imágenes, de transmisión de la información y las estrategias de sincronización de los dispositivos. A continuación se

detallan los aspectos más importantes que han sido tenidos en cuenta para desarrollar la arquitectura global del sistema.

6.2.1 Arquitectura jerárquica

La arquitectura del sistema de inspección de chapa laminada ha sido diseñada para poder llevar a cabo la adquisición, el almacenamiento y el procesamiento de las imágenes en paralelo, y poder así utilizar los resultados de forma inmediata. El sistema de inspección visual automática está organizado mediante una estructura jerárquica que incluye tres niveles organizados de abajo a arriba, como se puede ver en el esquema de la **Figura 34**.

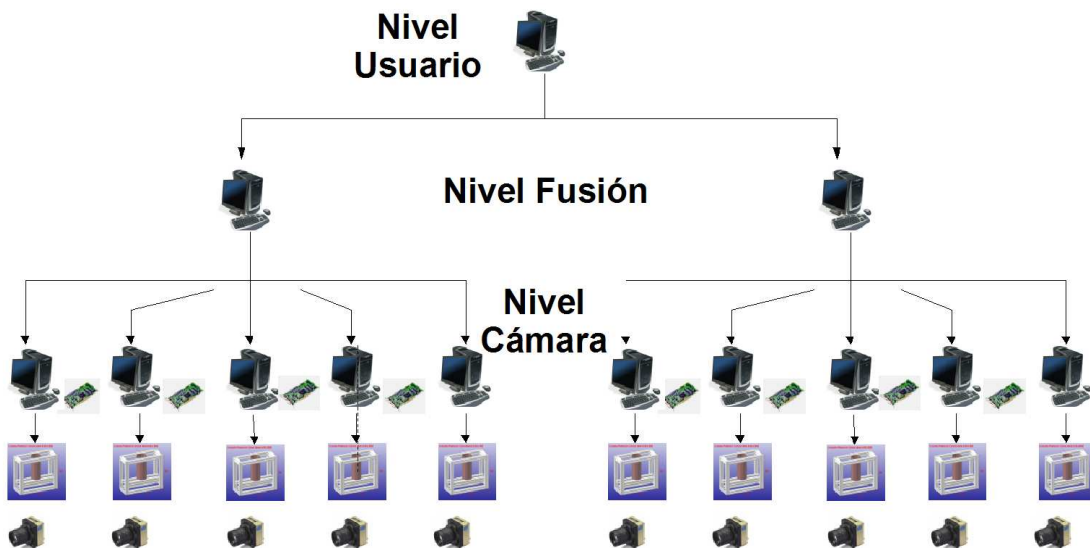


Figura 34. Arquitectura jerárquica del sistema de inspección visual de chapa laminada

Estos niveles tienen las siguientes funcionalidades:

- 1. Nivel Cámara:** Su principal función es adquirir las imágenes, almacenarlas localmente, preprocesarlas, segmentarlas y, a continuación, enviarlas al nivel superior.
- 2. Nivel Fusión:** En este nivel, la información proveniente de todas las cámaras es combinada y procesada. Las imágenes segmentadas y preprocesadas son las entradas de este nivel, mientras que la información sobre la calidad de la bobina es su salida.
- 3. Nivel Usuario:** Este nivel es el que interactúa con el operario. La información de la calidad de la bobina es presentada al usuario de forma apropiada. El usuario puede interactuar con el sistema, y el sistema

interactúa con el usuario y con la línea de producción, llegando incluso a detener la línea de procesado si los parámetros de calidad se encuentran fuera de los rangos admitidos.

6.2.2 Adquisición de imágenes

El proceso de corte de chapa laminada se realiza a altas velocidades, alcanzado una velocidad de hasta 200m/min, por lo que la utilización de cámaras de alta velocidad es imprescindible. Estas cámaras deben ofrecer tanto altas velocidades de adquisición como de transferencia de datos, altas ganancias y una relación señal-ruido bajo, ya que los tiempos de exposición deben ser necesariamente cortos.

Como se discute en el Capítulo 3, aunque la utilización de cámaras matriciales de alta velocidad podría llegar a solventar el problema, la solución más apropiada para este caso es la utilización de cámaras lineales. Estas cámaras alcanzan las velocidades establecidas, simplifican el proceso de iluminación, ya que es más sencillo iluminar homogéneamente una línea que una superficie, y además, el espacio requerido para instalar el sistema es mucho menor, lo cual facilita su integración en una línea de producción industrial.

Como los defectos objetivo de esta instalación pueden llegar a tener un tamaño de 1mm^2 , la resolución de la imagen necesaria que se ha establecido es de $0,5 \times 0,5\text{mm}$. Esto supone una velocidad de adquisición cercana a los 7kHz, es decir, 7.000 líneas por segundo. Las bobinas procesadas en estas líneas de producción pueden alcanzar 2,1m de ancho, por lo que más de 4.000 píxeles en la dirección normal al avance deben ser adquiridos. Para esta instalación se ha seleccionado la cámara lineal de Atmel AVIIVA 1010 SM2. Es una cámara con sensor de nivel de gris CCD, de 1.024 píxeles y comunicación Camera Link. Para cada lado de la bobina son necesarias 5 cámaras, de tal manera que se adquiere el ancho completo de la bobina incluyendo un pequeño solapamiento. El sistema instalado para la cara superior se puede ver en la **Figura 35**.

La etapa de adquisición también debe tener en cuenta la selección de las condiciones de iluminación, un problema especialmente clave al iluminar superficies metálicas con alta reflectividad. Debido a esta alta reflectividad el espectro electromagnético de la fuente de iluminación no tiene demasiada influencia, ya que la mayor parte de luz es reflejada, independientemente de su longitud de onda. Por tanto, la opción más razonable es utilizar la longitud de onda en la que la sensibilidad de la cámara sea mayor.

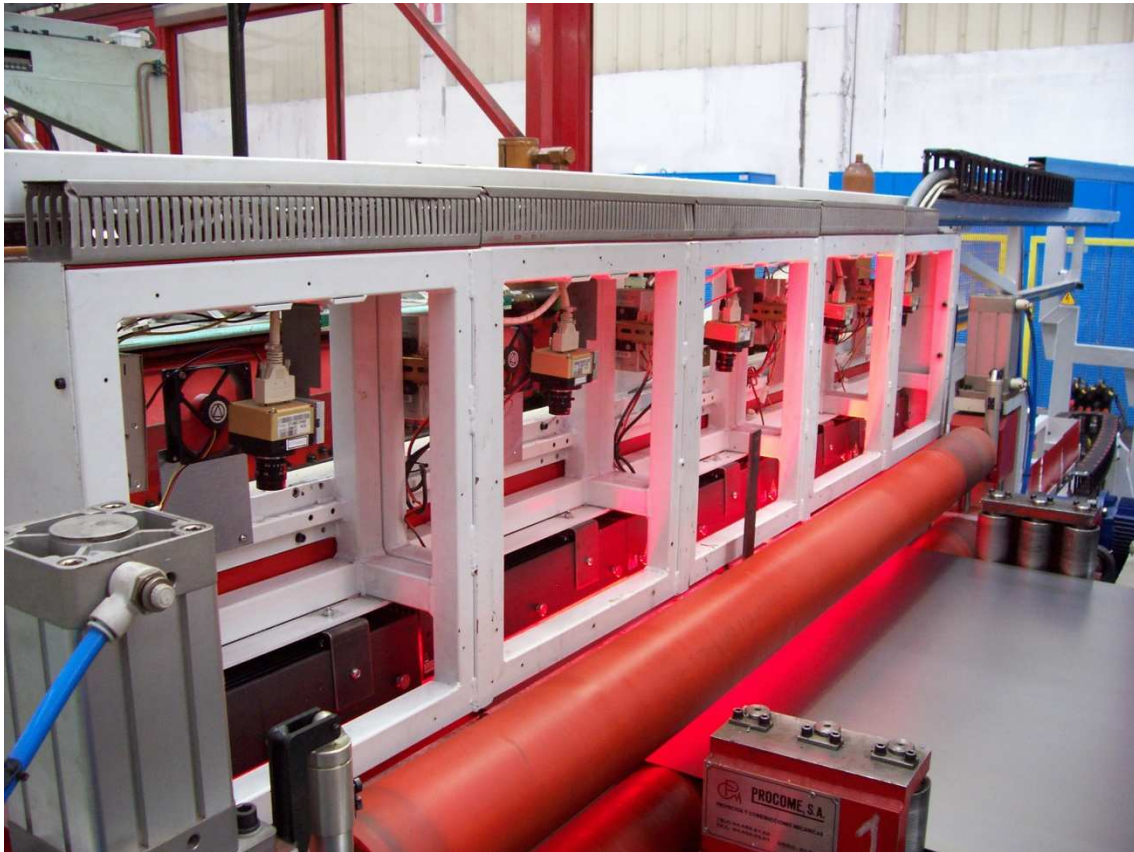


Figura 35. Módulos de la cara superior del sistema de inspección de chapa laminada. Cada módulo dispone de cámara e iluminación. Las cubiertas para aislar el sistema de la iluminación ambiente han sido retiradas.

Sin embargo, en un entorno industrial real, la bobina se lubrica para evitar que se dañe durante el proceso de corte. El aceite lubricante tiene una influencia negativa en el proceso de adquisición de las imágenes, ya que enmascara los defectos y modifica la apariencia visual de la bobina. Para reducir su efecto, se ha seleccionado una fuente de iluminación cuya longitud de onda atraviese mejor el aceite. Mediante un análisis espectral se comprobó que el infrarrojo cercano convertía el aceite prácticamente transparente para el sistema de visión artificial. La fuente utilizada es una barra lineal de leds de alta potencia, ECL480R de i2S, con una lente focalizadora que suministra 54.000 lux con una longitud de onda de 625nm. Esta fuente de iluminación permite iluminar en la zona transparente del aceite lubricante, y al mismo tiempo trabajar a una frecuencia donde la cámara tiene elevada ganancia.

6.2.3 Transmisión de la información y sincronización

Una gran cantidad de información se genera desde las cámaras del sistema que debe ser adquirida de forma sincronizada. Cinco cámaras adquieren la información de cada cara de la bobina y deben adquirir cada línea de forma simultánea. Para ello se realiza una sincronización *hardware*, mediante un codificador incremental acoplado a uno de los dos cilindros entre los que pasa la bobina. El codificador incremental genera un pulso cada vez que la bobina avanza medio milímetro, el cual es utilizado para enviar la señal de adquisición de todas las cámaras al mismo tiempo.

El estándar IEEE 802.3ab para la red de área local gigabit Ethernet es utilizado para comunicar los ordenadores. El sistema está formado por tres redes. Dos de ellas comunican los ordenadores que controlan las cámaras y realizan el primer procesamiento con el ordenador que se encarga de fusionar las imágenes y finalizar el procesamiento conjunto. La otra red conecta los ordenadores encargados de fusionar la información con el encargado de hacer de interfaz con el usuario.

Todas las cámaras se inicializan cuando reciben la orden de comenzar la adquisición. Cuando cada cámara se encuentra preparada para capturar las imágenes, envía una señal al ordenador encargado de fusionar las imágenes. Una vez que todas las cámaras están listas, el ordenador encargado de fusionar las imágenes envía la señal de que todas cámaras están listas al ordenador interfaz. Cuando los dos ordenadores del nivel de fusión han enviado su señal al ordenador de usuario, el codificador incremental se activa. De esta manera, todas las cámaras comienzan la adquisición de forma simultánea.

La secuencia de detener la adquisición sigue el proceso contrario. Primero se desactiva la señal del codificador incremental y a continuación se envía la señal de detener la adquisición a todas las cámaras. De esta manera se consigue que el último pulso de codificador incremental sea el mismo para todas las cámaras.

El diagrama de flujo para todos los ordenadores y las comunicaciones entre todos ellos y los dispositivos se encuentra en la **Figura 36**. Cada cara de la bobina se inspecciona con 5 cámaras que son controladas por 5 ordenadores que se comunican con el encargado de fusionar la información de esa cara. Otro ordenador se comunica con los encargados de fusionar la información de cada una de las dos caras e interactúa con el usuario y la línea de producción. En total se han dispuesto 13 ordenadores en cada línea. La correcta sincronización de todos los componentes del sistema es fundamental para adquirir la información.

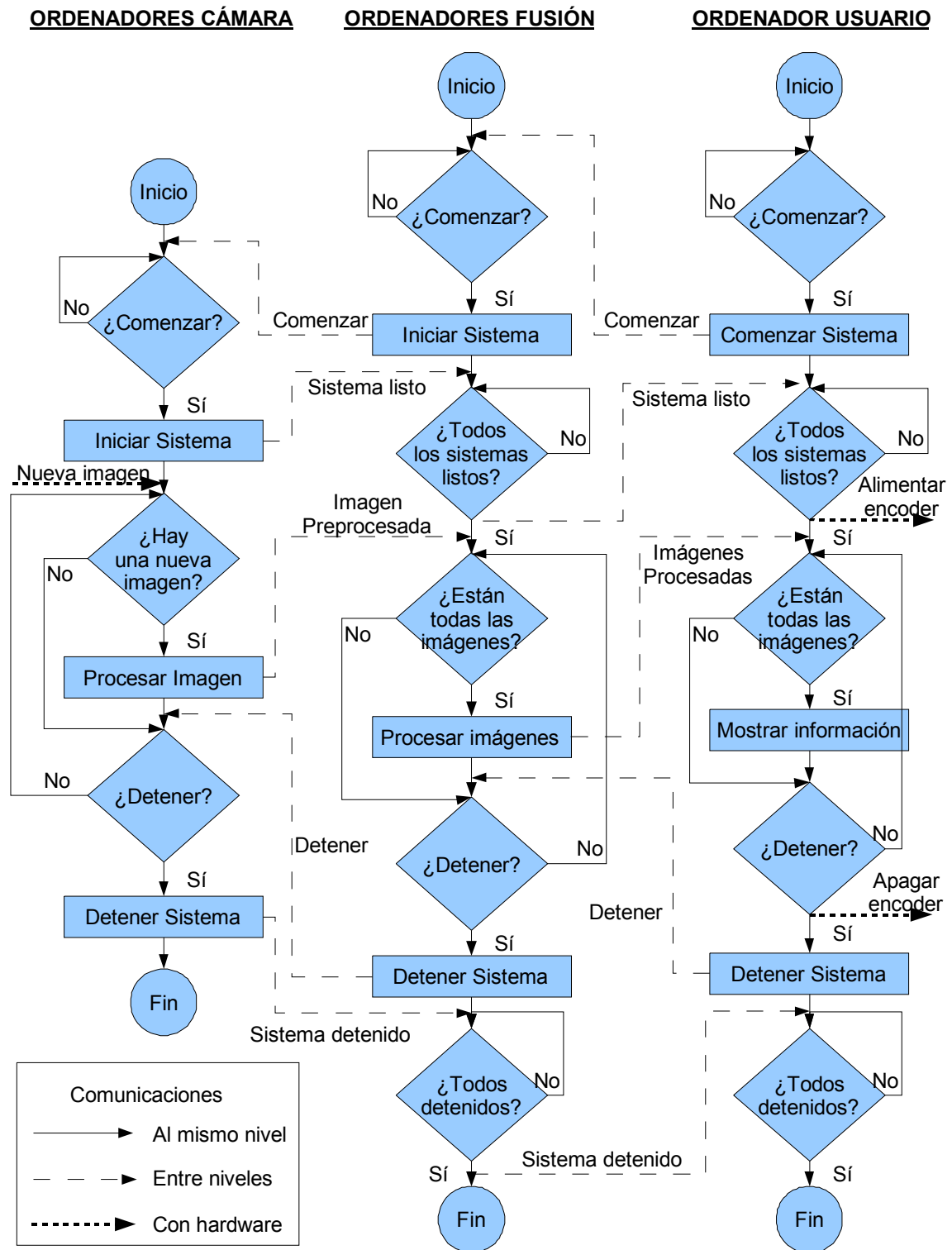


Figura 36. Diagrama de flujo del proceso de adquisición y procesamiento de imágenes.

Cada cámara envía la información en bloques de 1.000 líneas. Cada ordenador conectado a cada una de las cámaras se encarga de preprocesar y segmentar la imagen y enviar la información necesaria al ordenador encargado de fusionar la información. Existe un pequeño solapamiento horizontal así como un pequeño desfase vertical entre dos cámaras adyacentes que es corregido por el ordenador encargado de fusionar la información. Asimismo este ordenador se encarga de realizar la clasificación de los defectos. Los resultados son mostrados y almacenados en el ordenador usuario.



Figura 37. Puesto de control del sistema de inspección

Desde el puesto de inspección se pueden visualizar los resultados obtenidos del procesamiento en cualquier momento posterior al proceso de corte de la bobina. Se pueden visualizar tanto los datos brutos almacenados como los datos procesados. Para ello se ha desarrollado un *software* específico, del que se muestran algunas capturas de pantalla en la **Figura 38**. Esta herramienta permite la navegación por la bobina, ver un mapa completo de los defectos que aparecen en la bobina, visualizar los defectos ordenados por tipo, etc.



Figura 38. Capturas de pantalla del software de visualización de resultados

6.3 Inspección del estado de la superficie de la carretera

La inspección de la superficie de la carretera requiere que la captura de datos se realice a la velocidad de la vía para no interrumpir el tráfico. Para el desarrollo del sistema de inspección se han seleccionado los equipos que permiten obtener la información necesaria a altas velocidades, se ha desarrollado tanto el *software* como la electrónica para la captura sincronizada de toda la información y un *software* para que el gestor de la carretera pueda utilizar la información para tomar las medidas oportunas de mantenimiento y rehabilitación de la calzada.

6.3.1 Equipos para la captura de datos

Para el desarrollo del vehículo de auscultación de carreteras se han incorporado un conjunto de dispositivos de adquisición a un vehículo comercial, una furgoneta *Citroën Jumper*. Entre estos dispositivos destaca el equipo de adquisición de imágenes LRIS, cuyas imágenes son las que van a ser evaluadas en este trabajo. Además, el vehículo dispone de un conjunto de equipos que complementan esta información y permiten su correcta adquisición. Dos perfilómetros láser miden el perfil de la carretera en sus dos rodadas, el cual también se utiliza para calcular el Índice de Rugosidad Internacional (IRI, *International Roughness Index*), unidad de medida de la regularidad superficial de la carretera. Una cámara adicional captura imágenes del entorno de la carretera, lo cual facilita localizar mejor los defectos detectados cuando van a ser reparados. Un GPS diferencial permite georreferenciar con precisión los datos adquiridos con el resto de los dispositivos. La señal de disparo de la toma de datos se realiza utilizando un codificador incremental. Asimismo, este permite sincronizar los distintos dispositivos de forma muy precisa y ayuda a la localización exacta de los datos adquiridos. Dos

Inspección visual automática de superficies continuas caracterizando anomalías locales en el dominio espacio-frecuencial

ordenadores, *Pentium Core 2 Duo*, son utilizados para controlar los dispositivos y almacenar toda la información. En la **Figura 39** se muestra el vehículo utilizado y los equipos que en él se encuentran instalados.



(a)



(b)

Figura 39. Vehículo utilizado para la inspección de carreteras. (a) Vista trasera (b) Vista frontal

6.3.2 Configuración y sincronización de la adquisición de datos

El sistema de adquisición de imágenes embarcado debe ser lo suficientemente rápido para adquirir y transferir las imágenes a una velocidad que permita conducir el vehículo a la velocidad normal de circulación de la vía. De esta manera la inspección de la carretera no afectará al tráfico. Para capturar las imágenes a esta velocidad, los tiempos de exposición de las cámaras deben ser pequeños, necesariamente inferiores al tiempo que transcurre entre dos pulsos del codificador incremental, siendo las imágenes más nítidas cuanto menor sea el tiempo de exposición. Para que la cantidad de luz que llegue al sensor de la cámara en un pequeño intervalo de tiempo sea suficiente, es necesario utilizar un potente sistema de iluminación. Aunque podrían utilizarse cámaras matriciales, es preferible usar cámaras lineales porque su velocidad de adquisición es mayor y la iluminación de una línea es más sencilla y consume menos energía que la iluminación de una porción mayor de la carretera.

El sistema de adquisición LRIS, fabricado por INO, gracias a la utilización de iluminación láser y cámaras lineales de alta resolución es el sistema comercial que mejor se ajusta a los objetivos establecidos, ya que permite obtener imágenes de alta calidad a altas velocidades. Este sistema dispone de dos cámaras lineales que pueden adquirir a una velocidad máxima de 28.000 líneas por segundo, formando una única imagen de 4.096 píxeles de ancho. Por tanto, el vehículo puede capturar la información mientras circula a una velocidad máxima de 100km/h con una resolución de las imágenes de 1 x 1 mm por píxel. Este sistema resultó elegido porque cumple los requerimientos de velocidad de adquisición, resolución de la imagen y consumo de energía.

El vehículo también cuenta con dos perfilómetros láser para la medida del perfil longitudinal de la carretera en sus dos rodadas. Los equipos instalados son dos *LaserProf* de Greenwood Engineering, que es el perfilómetro láser más utilizado en la actualidad. Cada equipo dispone de un láser y un sensor inercial. Estos sensores, así como la señal de un codificador incremental acoplado a la rueda del vehículo, se conectan a un sistema de control que integra toda la información. Alcanza una frecuencia de muestreo de 16kHz, de tal manera que se dispone de un punto del perfil de la carretera cada 1.39mm a 80km/h. Los perfiles longitudinales proporcionan información adicional del estado de la carretera y sirven para calcular el IRI.

En el vehículo de inspección también se encuentra instalado un GPS diferencial para obtener su posición de forma precisa, así como una cámara para grabar el entorno de la zona de inspección. Cada 1,7 metros de avance del vehículo, se lanza una señal para almacenar una imagen y una posición GPS. De esta forma la frecuencia de adquisición es independiente de la velocidad del vehículo. Estos dos dispositivos proporcionan una localización precisa de los defectos detectados por el sistema de inspección.

Todos los equipos del vehículo de inspección están sincronizados gracias a un codificador incremental que genera 5.000 pulsos por vuelta de rueda, lo que supone un pulso cada vez que el vehículo avanza 0,33mm. Cada dispositivo debe tomar los datos a distintas frecuencias. La señal normal del codificador incremental es utilizada por el *software* de control de los perfilómetros láser para ofrecer una medida del perfil cada 1.39mm. La frecuencia de muestreo del resto de los dispositivos se consigue mediante una electrónica de control diseñada para adaptarse a las características de cada dispositivo. El sistema LRIS adquiere la imagen de una línea transversal de la carretera cada tres pulsos del codificador incremental, mientras que el GPS diferencial proporciona un dato de posición y la cámara una imagen cada 5.000 pulsos.

Además, el codificador incremental se utiliza para sincronizar el comienzo y el fin de la grabación. Para garantizar la correcta sincronización de los datos de todos los dispositivos, el codificador incremental se alimenta cuando todos los dispositivos están listos para adquirir datos, de tal manera que el primer pulso de codificador incremental es el mismo para todos los dispositivos. De igual manera, el codificador incremental se apaga cuando los dispositivos aún están adquiriendo información, de tal manera que el último pulso de codificador incremental es el mismo para todos ellos.

Un *software* de control, cuya captura de pantalla se muestra en la **Figura 40**, se encarga de gestionar las comunicaciones con los dispositivos, su configuración y la captura y el almacenamiento de los datos.



Figura 40. Captura de pantalla del software de adquisición de datos

La información adquirida es procesada después de su adquisición. En primer lugar, se fusiona toda la información obtenida y a continuación se aplica a las imágenes obtenidas con el LRIS las técnicas de procesamiento de imágenes para la detección y clasificación de grietas, que se describen en la sección 7.2.

6.3.3 Visualización de la información obtenida

Se ha desarrollado un *software* a medida para analizar todos los datos adquiridos y procesados y que sus resultados puedan ser utilizados por un gestor de mantenimiento de carreteras. En la **Figura 41** se muestran dos capturas de pantalla de la interfaz del *software* utilizado para mostrar los resultados. La posición exacta de las grietas se muestra en el mapa de la esquina superior derecha de las dos capturas de pantalla. Esta ventana puede ser utilizada para navegar por toda la ruta donde se han adquirido datos. Para ello solo se debe hacer "*click*" en la posición donde se desea obtener la información. Asimismo, en la parte de abajo a la derecha de ambas capturas de pantalla se muestra la imagen obtenida por la cámara utilizada para ver el entorno. A la izquierda de estas dos ventanas se encuentran los perfiles longitudinales de las dos rodadas de la carretera.

La ventana principal de ambas capturas de pantalla se muestra a su izquierda. En la **Figura 41 (a)**, se encuentra la ventana utilizada para mostrar la imagen adquirida por el sistema de adquisición de imágenes LRIS. Una imagen continua de toda la carretera puede ser reproducida. Las funciones de pausa, parada, adelante y atrás también están disponibles. También es posible desplazarse por los datos adquiridos introduciendo el punto kilométrico donde se encuentran los datos que se desean ver. Así mismo, es posible realizar zoom para observar con más detalle las imágenes adquiridas con el LRIS.

En la **Figura 41 (b)**, la ventana principal muestra información de los defectos detectados por el *software* de procesamiento de imágenes. Se muestra un mapa de defectos para informar de las zonas con presencia de defectos, de su tipo y cantidad. Cuando el usuario selecciona un defecto, la imagen y las características de ese defecto también se muestran por pantalla. Este programa también permite realizar informes sobre el estado de la carretera, ofreciéndonos información sobre el número total de defectos clasificados por tipo, el porcentaje de superficie defectuosa, la longitud de las grietas y fotografías de los defectos, tanto del total de datos inspeccionados como de los tramos seleccionados por el usuario.

En esta sección se van a estudiar las distintas tipologías de túneles y las posibilidades de adaptación de una plataforma de inspección según sea el tamaño mínimo del ancho de grieta que se quiera detectar. A continuación se describe el prototipo de la plataforma de inspección de túneles desarrollado para capturar imágenes en un túnel y poder comprobar la viabilidad del sistema.

6.4.1 Adaptación del sistema de inspección al tipo de túnel y a la resolución de la inspección

Tres túneles tipo han sido analizados con el fin de conocer las distintas tipologías de túneles que se pueden encontrar.

- **Un túnel del AVE.** El túnel del AVE que une la estación de Chamartín con la estación de Atocha es actualmente el túnel ferroviario de mayor diámetro de España. Se trata de un túnel de doble vía, con un diámetro de excavación de 11,50m y un diámetro interior de 10,40m. El perímetro de hormigón visible en este túnel es de 20m.
- **Un túnel de metro:** El túnel de metro de la línea 1 de Sevilla tiene un diámetro de excavación de 6,05m y un diámetro interior del túnel de 5,30m. El perímetro de hormigón visible del túnel es de 13m.
- **Un túnel para el transporte de energía eléctrica:** El túnel de Inelfe para la interconexión eléctrica entre España y Francia tiene un diámetro de excavación de 4,50m y un diámetro interior de 3,50m, con un perímetro a inspeccionar de 8,50m.

Para conseguir una resolución uniforme de todo el perímetro del túnel se deben disponer las cámaras de forma radial, de tal forma que las cámaras estén dispuestas en una circunferencia cuyo centro coincida con el del túnel y orientadas de forma que sus ejes ópticos corten el centro del túnel.

Debido a las características del túnel, existe otro factor restrictivo que se debe cumplir a la hora de decidir la disposición de las cámaras sobre la plataforma de inspección. Solo se pueden colocar los elementos de inspección en la zona del túnel libre de obstáculos.

En la **Figura 42** se pueden ver las secciones tipo de los túneles propuestos así como las zonas de trabajo donde es posible situar las cámaras. La zona de trabajo es una circunferencia cuyo centro coincide con el del túnel y cuyo radio se denomina radio de la zona de trabajo, R_{ZT} .

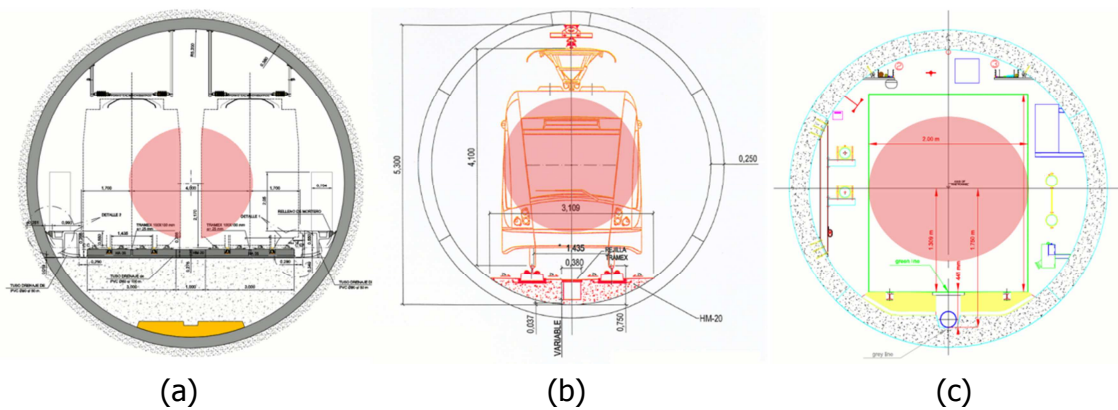


Figura 42. Secciones tipo de túnel ferroviario. En rojo las zonas de trabajo donde es posible situar las cámaras (a) Túnel del AVE Chamartín-Atocha; (b) Línea 1 del metro de Sevilla; (c) Túnel de interconexión eléctrica de Inelfe

Dependiendo del tamaño mínimo de grieta que se tenga por objetivo detectar, se debe seleccionar un tipo de cámara y una óptica apropiada, tal y como se explicará más adelante. Estos parámetros determinarán la distancia de trabajo, DT , es decir, la distancia entre el centro óptico de la cámara y la superficie del túnel. Esta distancia de trabajo debe ser menor que el radio del túnel, R_T , más el radio de la zona de trabajo y mayor que el radio del túnel menos el radio de la zona de trabajo.

$$R_T - R_{ZT} < DT < R_T + R_{ZT}$$

Ec. 40

Cuando la distancia de trabajo, DT , sea mayor que el radio del túnel, R_T , solo se podrán colocar cámaras en la mitad de la circunferencia, ya que si no unas cámaras obstaculizarían la visión de las opuestas. Ante esta situación se dispone de dos posibilidades, bien realizar dos inspecciones en las que se inspeccionaría la mitad del perímetro del túnel cada vez, o bien situar las cámaras en dos planos diferentes, de tal forma que haya un desfase entre las cámaras. En el caso de un túnel de doble vía, como el túnel del AVE entre las estaciones de Atocha y Chamartín descritos anteriormente, siempre nos encontraremos ante esta situación y la única solución posible para capturar la totalidad de la superficie del túnel es realizar dos inspecciones, una desde cada vía.

Otra circunstancia a tener en cuenta, es que la distancia de trabajo no debe coincidir con el radio de túnel, ya que en esta situación todas las cámaras deberían situarse en el mismo punto, lo que evidentemente es físicamente imposible. Por ello, los valores de distancia de trabajo en un entorno de $\pm 25\text{cm}$ del radio del túnel tampoco estarán permitidos.

Tabla 4. Máxima y mínima distancia de trabajo (DT) para cada uno de los tres tipos de túneles propuestos. En un entorno de 25cm al radio del túnel (R_T) tampoco será posible situar las cámaras

Tipo Túnel	Diámetro	Mínima DT	DT = R_T	Máxima DT
Transporte de energía	3,50 m	0,75 m	1,75 m	2,75 m
Metro	5,30 m	1,15 m	2,65 m	4,15 m
AVE	10,40 m	3,20 m	4,90 m	7,20 m

Por otro lado, se dispone de diversos tipos de cámaras en el mercado con distintas resoluciones que pueden permitir abarcar un gran número de posibilidades. Se han considerado para el estudio 5 tipos de cámaras lineales con resoluciones que van desde los 1.024 píxeles hasta los 12.288, tal y como se puede ver en la **Tabla 5**.

Tabla 5. Cinco cámaras lineales tipo con diferentes resoluciones

Cámara	Tamaño del sensor (píxeles)	Tamaño de píxel (μm)	Tamaño del sensor (μm)	Velocidad máxima (kHz)
e2v AviiVA SM2 1010	1.024	10	10.240	54.4
e2v AviiVA SM2 4010	4.096	10	40.960	14.8
Dalsa P2-22-06K40	6.144	7	43.008	12
Dalsa P2-22-08K40	8.194	7	57.358	9
e2v AviiVA UM8	12.288	5	61.440	25

Se puede observar que el tamaño del sensor de las dos últimas cámaras es superior a los 43,3mm, que es el máximo tamaño de sensor válido para ópticas de 35mm (*full frame* o tamaño completo, que coincide con el tamaño de las películas de 35mm). Por encima de este tamaño la variedad de ópticas que se encuentran en el mercado se reduce significativamente, de la misma forma que aumenta su precio.

El último aspecto que se debe plantear es cuál es el tamaño de grieta mínimo que se quiere detectar. La resolución necesaria que permite a la cámara poder medir este tipo de grietas requiere el doble de resolución de la imagen, es decir, para detectar grietas de 1mm de ancho se deben adquirir las imágenes con una resolución de al menos 2 píxeles por milímetro.

Se han considerado 4 tamaños de grieta mínimo: 0,1mm, 0,4mm, 1mm y 2mm. Los valores 0,1mm, 0,4mm y 2mm se han seleccionado porque son los límites entre los distintos tipos de grietas considerados en la sección 2.5.2, mientras que el límite de 1mm se ha añadido por ser un valor intermedio entre los dos últimos que se considera que es un valor límite de detección razonable.

Para calcular el número de cámaras mínimo necesario para poder adquirir todo el túnel en una sola inspección, se aplica la siguiente ecuación:

$$N_C > \frac{2P_T}{T_G^{\min} T_S} \quad \text{Ec. 41}$$

Donde N_C es el número de cámaras, P_T es el perímetro del túnel que se va a inspeccionar, T_G^{\min} es el tamaño de grieta mínimo que se desea detectar y T_S es el tamaño del sensor en píxeles. Se debe tener en cuenta que es necesario un pequeño solapamiento entre las imágenes de las distintas cámaras, por lo que el valor obtenido corresponderá al número mínimo de cámaras necesarias.

Finalmente, para configurar el sistema se debe seleccionar una óptica apropiada para la cámara seleccionada que permita capturar la imagen en los límites de la distancia de trabajo establecidos anteriormente y que sea capaz de enfocar a dicha distancia. Dentro de las opciones que se pueden encontrar en el mercado para los distintos tamaños de sensor se pueden considerar las siguientes posibilidades:

- Para inspeccionar un túnel tipo para el transporte de energía de **3,5m de diámetro**:
 - Si se requiere una **precisión de 0,1mm**, la solución más recomendable es utilizar 7 cámaras de 12.288 píxeles a una distancia de trabajo de 0,825m y con una óptica de 80mm (cubriendo así la mitad del túnel y necesitando por tanto dar dos pasadas). Con 14 cámaras se podría inspeccionar el túnel completo.
 - Si se requiere una **precisión de 0,4mm**, la solución más recomendable es utilizar 4 cámaras de 6.144 píxeles o 3 cámaras de 8.194 píxeles a una distancia de trabajo de unos 1,5m y con una óptica de 50mm (cubriendo en ambos casos la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas). Se podría inspeccionar con una única pasada con 7 de cámaras de 6.144 píxeles o 6 cámaras de 8.194 píxeles
 - Si se requiere una **precisión de 1mm**, la solución más recomendable es utilizar 2 cámaras de 6144 píxeles a una distancia de trabajo de unos 2,4m y con una óptica de 28mm o bien 3 cámaras de 4.096 píxeles y

- una óptica de 50mm (cubriendo en ambos casos la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas).
- Por último, si se requiere una **precisión de 2mm**, la solución más recomendable es utilizar 5 cámaras de 1.024 píxeles a una distancia de trabajo de 1,25m y con una óptica de 12mm (cubriendo así la mitad del túnel y necesitando por tanto dar dos pasadas). Con 9 cámaras se podría adquirir el túnel completo en una única pasada.
 - Túnel tipo Metro de **5,3m de diámetro**:
 - Si se requiere una **precisión de 0,1mm**, no existe una óptica apropiada para alcanzar esa resolución que enfoque a la distancia permitida. Sería necesario estar a 1,05m, pero las cámaras no se pueden situar a menos de 1,15m de la superficie del túnel. Con una óptica de 105 mm se podría alcanzar una precisión muy cercana a la requerida. Se necesitan 11 cámaras de 12.288 píxeles (cubriendo la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas) o 22 cámaras para cubrir el túnel completo en una única pasada.
 - Si se requiere una **precisión de 0,4mm**, la solución más recomendable es utilizar 6 cámaras de 6.144 píxeles o 4 cámaras de 8.194 píxeles a una distancia de trabajo de unos 1,5m y con una óptica de 50mm (cubriendo en ambos casos la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas). Con 11 cámaras de 6.144 píxeles u 8 cámaras de 8.194 píxeles se podría cubrir toda la superficie del túnel de una única pasada.
 - Si se requiere una **precisión de 1mm**, la solución más recomendable es utilizar 3 cámaras de 6.144 píxeles a una distancia de trabajo de unos 2,35m y con una óptica de 28mm (cubriendo la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas). También se podrían utilizar 5 cámaras para cubrir el túnel con una única pasada.
 - Por último, si se requiere una **precisión de 2mm**, la solución más recomendable es utilizar 2 cámaras de 4.096 píxeles a una distancia de trabajo de 3,3m y con una óptica de 28mm (cubriendo así la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas).
 - Túnel tipo AVE de **10,4m de diámetro**:
 - Si se requiere una **precisión de 0,1mm**, se necesitaría utilizar 17 cámaras de 12.288 píxeles para inspeccionar el túnel en dos pasadas. A una distancia de trabajo de 1,8m y con una óptica de 180mm se podría obtener la resolución requerida, pero estas ópticas no pueden enfocar a menos de 4m de distancia y además se situaría fuera de la zona de trabajo. Esta opción no es viable dentro del abanico de ópticas que se pueden encontrar en el mercado.
 - Si se requiere una **precisión de 0,4mm**, la solución más recomendable es utilizar 9 cámaras de 6.144 píxeles o 7 cámaras de 8.194 píxeles a una distancia de trabajo de unos 3,90m y con una óptica de 135mm

(cubriendo en ambos casos la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas).

- Si se requiere una **precisión de 1mm** la solución más recomendable es utilizar 4 cámaras de 6.144 píxeles o 3 cámaras de 8.194 píxeles a una distancia de trabajo de 3,85m y con una óptica de 50mm (cubriendo en ambos casos la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas).
- Por último, si se requiere una **precisión de 2mm**, la solución más recomendable es utilizar 3 cámaras de 4.096 píxeles a una distancia de trabajo de 3,20m y con una óptica de 28mm (cubriendo así la mitad del túnel y necesitando por tanto realizar la inspección en dos pasadas).

Como se puede comprobar existen diversas soluciones que permiten la inspección en la mayoría de los casos planteados, siendo únicamente problemático la detección de grietas de 0,1mm de ancho en los túneles más grandes. En la **Tabla 6** se muestra un resumen de los datos expuestos anteriormente:

Tabla 6. Resolución y número de cámaras para los distintos tipos de túneles estudiados

Precisión	Tamaño cámara (píxeles)	Túnel electricidad			Túnel Metro			Túnel AVE		
		Nº camaras		Óptica (mm)	Nº camaras		Óptica (mm)	Nº camaras		Óptica (mm)
		1 vez	2 veces		1 vez	2 veces		1 vez	2 veces	
0,1 mm	12.288	14	7	80	22	11	105	-	-	-
0,4 mm	8.194	6	3	50	8	4	50	-	7	135
	6.144	7	4		11	6		-	9	
1 mm	6.144	-	2	28	-	3	28	-	4	50
2 mm	4.096	-	-		-	2	28	-	3	28
	1.024	9	5	12	-	-	-	-	-	-

Para la detección de grietas entre 0,4mm y 1mm de ancho, las cámaras de 6.144 píxeles son válidas para los distintos tipos de túneles utilizando ópticas convencionales.

Por último, la información capturada por las cámaras debe ser almacenada para su análisis posterior en el laboratorio. En el caso más desfavorable planteado, el túnel del AVE de 10,4m de diámetro en el que se buscan grietas de 0,1mm de ancho, se

necesitan cerca de 8Tbytes para almacenar las imágenes de un 1km de túnel. Si se reduce el tamaño de grieta buscado a 0,4mm, se podría almacenar en 0,5Tbytes. Para el resto de las resoluciones los requerimientos de almacenamiento son mucho más manejables.

6.4.2 Descripción del prototipo desarrollado

Para probar la viabilidad del sistema propuesto se ha desarrollado un pequeño prototipo con una cámara lineal de 1.024 píxeles de la empresa e2v (llamada Atmel anteriormente) modelo Aviiva SM2 CL 1010. Se trata del mismo modelo de cámara utilizado para la inspección de bobinas de chapa laminada. Este prototipo permite la captura de imágenes de la superficie del túnel para poder después procesarlas y verificar los algoritmos de procesamiento de imágenes para la detección de grietas. El diseño del prototipo donde se sitúan los equipos se encuentra en la **Figura 43**.

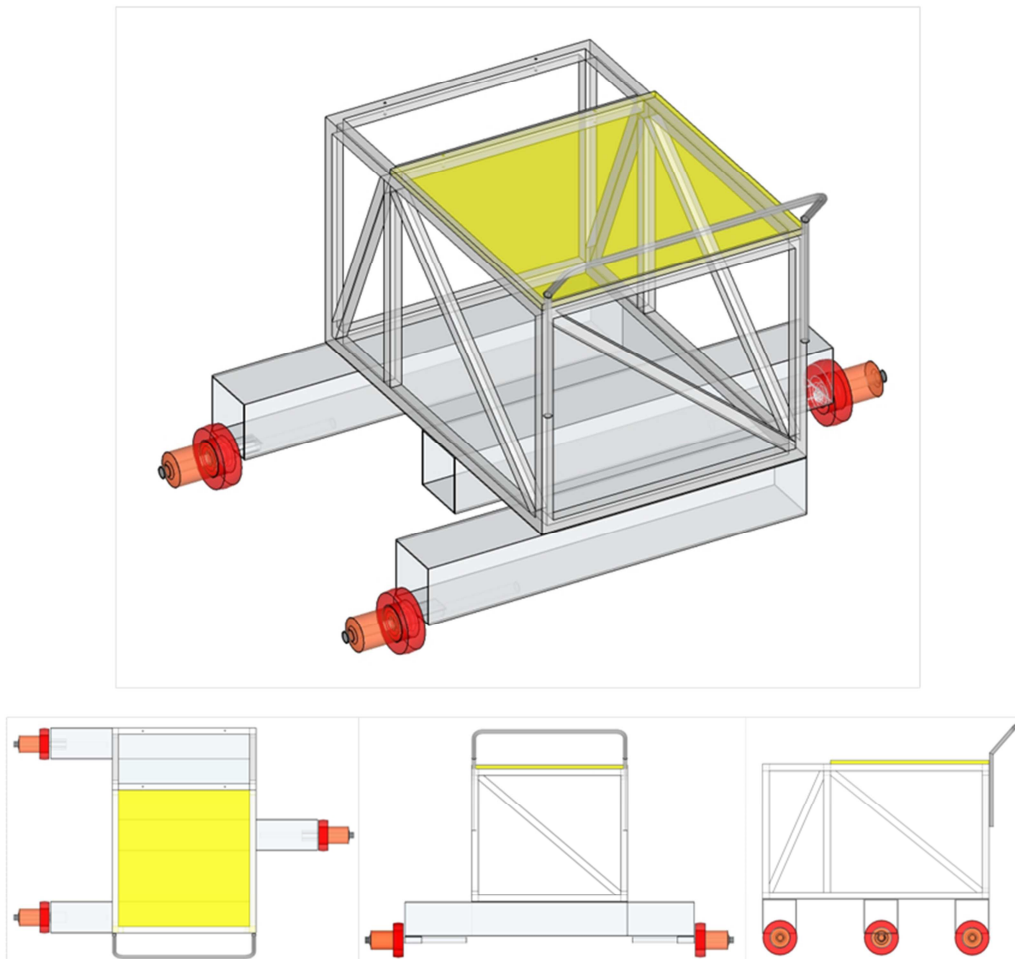


Figura 43. Diseño del prototipo de inspección de túneles ferroviarios

La plataforma de inspección consta de una cámara, una fuente de iluminación, un codificador incremental, un ordenador con una tarjeta de adquisición de datos y las fuentes de alimentación. Dicha plataforma puede circular sobre vías ferroviarias de diferentes anchos. Se ha diseñado para circular en ancho de vía internacional (1.435mm) aunque permite la modificación de la separación de las ruedas para circular por otros anchos de vía, como el ancho ibérico (1.668mm). En la **Figura 44** se puede ver la plataforma de inspección situada en los tres tipos de túneles descritos anteriormente.

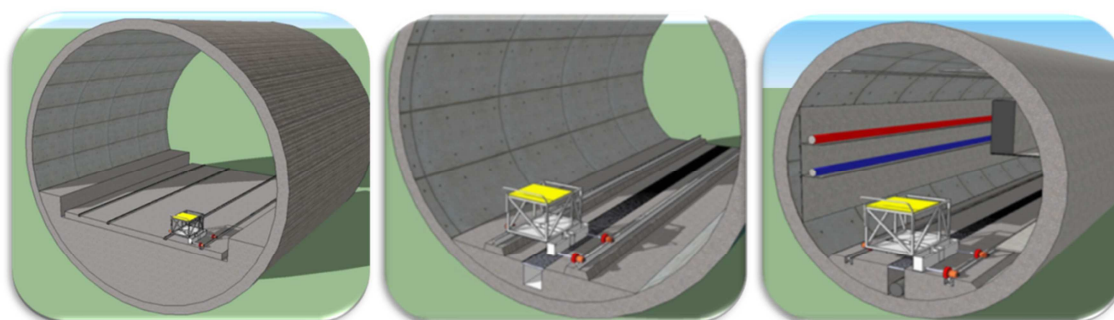


Figura 44. Plataforma de inspección en los túneles tipo descritos

La plataforma ha sido desarrollada según el diseño anterior y se le han incorporado los dispositivos necesarios para la adquisición de imágenes, tal y como se puede ver en la **Figura 45** y que serán explicados con más detalle a continuación.



Figura 45. Prototipo de la plataforma de inspección de túneles

La plataforma de inspección se apoya sobre tres ruedas para conseguir la máxima estabilidad posible. Se ha optado por utilizar tres ruedas ya que de esta forma las tres ruedas siempre se apoyan sobre un plano consiguiendo la máxima estabilidad posible. Estas ruedas permiten circular por las vías del tren así como desplazarse por terrenos llanos, lo que facilita su transporte hasta la infraestructura a inspeccionar. Así mismo se ha incluido un sistema de guiado para evitar el cabeceo durante la inspección. A una de las ruedas se le ha acoplado un codificador incremental de 5.000 pulsos por vuelta, para la sincronización de la toma de datos. En la **Figura 46** se muestra tanto la rueda apoyada sobre la vía y el sistema de guiado como la rueda a la que se le ha acoplado el codificador incremental.



Figura 46. Ruedas para el apoyo de la plataforma. (a) Rueda sobre vía de tren con sistema de guiado. (b) Rueda con el codificador incremental acoplado

El sistema cámara-iluminación está compuesto por la cámara e2v Aviiva SM2 CL 1010, la óptica Tamron AF 70-300mm F/4-5.6 Di LD Macro 1:2 y un foco halógeno de 100W de potencia con un reflector AR111 para la concentración de la luz en 8° . La cámara lineal tiene una resolución de 1.024 píxeles y puede adquirir hasta 54.400 líneas por segundo. La óptica seleccionada es de zoom variable entre 70 y 300mm, lo que permite su utilización en túneles de distintos diámetros y con distintas soluciones. El sistema de iluminación permite la concentración de la iluminación en una circunferencia de 8° . Tanto la cámara como la fuente de iluminación pueden ser orientadas en múltiples posiciones lo que permite al prototipo adaptarse a gran número de situaciones, permitiendo la captura de distintas zonas del túnel con diferentes resoluciones. En la **Figura 47** se puede ver un detalle de estos elementos del prototipo.

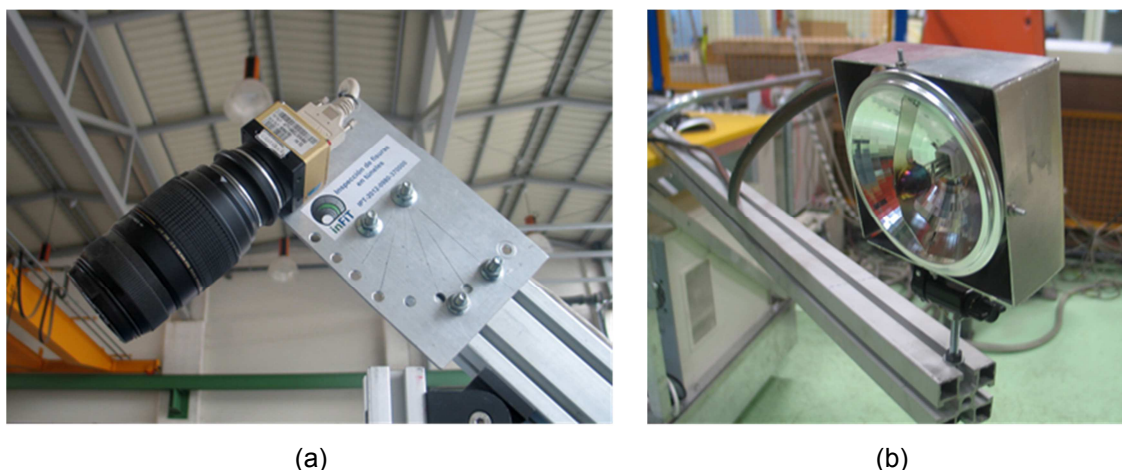


Figura 47. Sistema cámara-iluminación. (a) Cámara e2v Aviiva SM2 CL 1010 con óptica Tamron AF 70-300mm F/4-5.6 Di LD Macro 1:2. (b) Iluminación halógena de 100W con reflector AR111 para concentrar la luz con un ángulo de 8°

Finalmente, el prototipo dispone de un ordenador con una tarjeta de adquisición de imágenes mediante el protocolo *Camera Link*, para la adquisición, el almacenamiento de datos y la configuración del sistema. El *software* de adquisición de imágenes ha sido diseñado de forma modular, de tal forma que por un lado se ha desarrollado un *software* para realizar las funciones de interfaz de usuario que permite la configuración y el control de hasta 255 cámaras distintas conectadas a distintos ordenadores mediante comunicación Ethernet, mientras que por otro lado se ha desarrollado un *software* de adquisición para cada cámara que se comunica bidireccionalmente con el interfaz de usuario, tal y como se puede ver en el esquema de la **Figura 48**.

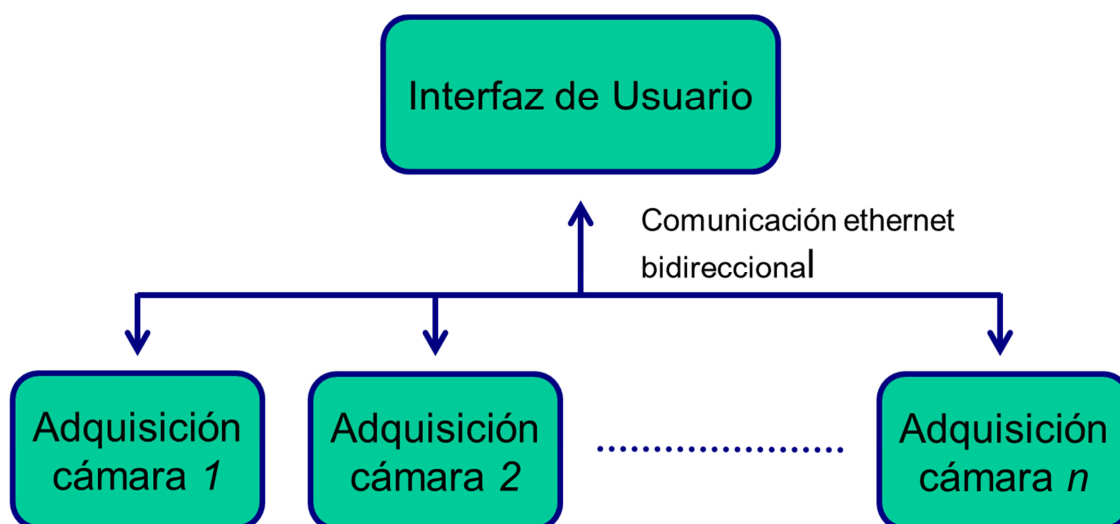


Figura 48. Esquema del software de adquisición y control

El interfaz de usuario, cuya captura de pantalla se encuentra en la **Figura 49**, permite configurar los siguientes valores:

- **El identificador del túnel:** Para nombrar al fichero e identificarlo posteriormente
- **El tipo de túnel:** Lo que permite configurar el equipo tanto para túneles ferroviarios como carreteros. Es utilizado para conocer el tamaño de la rueda actual del vehículo de inspección.
- **El número de cámaras:** Permite seleccionar el número de cámaras que se van a utilizar en la inspección
- **El puerto de comunicaciones:** Información necesaria para conectar los distintos ordenadores.
- **La resolución de las grietas:** Lo que determina la resolución con que son adquiridas las imágenes en el eje de avance. La resolución en la dirección perpendicular al avance viene determinada por la distancia de la cámara al túnel y por la óptica utilizada.
- **El modo de ajuste del tiempo de exposición:** El tiempo de exposición puede ajustarse de forma manual, siendo posible establecer un valor diferente para cada una de las cámaras, o de forma automática, de tal manera que el sistema captura una serie de imágenes al comienzo de la adquisición hasta conseguir, mediante un algoritmo iterativo, que el valor medio de la imagen sea 128.
- **El modo de adquisición:** Puede ser "*free run*" o "*encoder*". En el primer caso se capturan las imágenes a la máxima velocidad posible, mientras que en el segundo caso se captura una línea cada vez que el vehículo avanza una distancia preestablecida, definida por el tamaño de grieta que se busca detectar y el diámetro de la rueda del vehículo. El modo "*free run*" se utiliza para comprobar que las cámaras funcionan y están situadas correctamente, mientras que el modo "*encoder*" es el que se utiliza para realizar la inspección.

El sistema también permite la grabación y carga de configuraciones predefinidas y la selección de las cámaras que van a ser utilizadas en cada momento. Este puede ser apagado completamente seleccionando la opción de "Apagar Sistema", mediante la que se envía la orden de apagado a todos los ordenadores que conformen el sistema.

En todo momento se muestra por pantalla el estado de las distintas cámaras, la velocidad máxima a la que puede transitar el vehículo y la velocidad actual a la que se está desplazando el vehículo de inspección, con el objetivo de no sobrepasar esa velocidad, lo que provocaría la pérdida de información.



Figura 49. Captura de pantalla del software de interfaz de usuario

El módulo de adquisición, cuya ventana se muestra en la **Figura 50**, puede funcionar en modo remoto, controlado por el interfaz de usuario, o bien en modo local, especialmente diseñado para realizar pruebas de funcionamiento de los equipos. Además de configurar los distintos parámetros de adquisición, al igual que el módulo interfaz, permite la visualización de las imágenes en tiempo real.

Por último se ha desarrollado un *software* que permite la visualización de la información almacenada y de la información procesada, cuya captura de pantalla se muestra en la **Figura 51**. Este *software* permite la navegación por toda la información almacenada del túnel y visualizar el resultado de la detección de las grietas y fisuras del túnel.

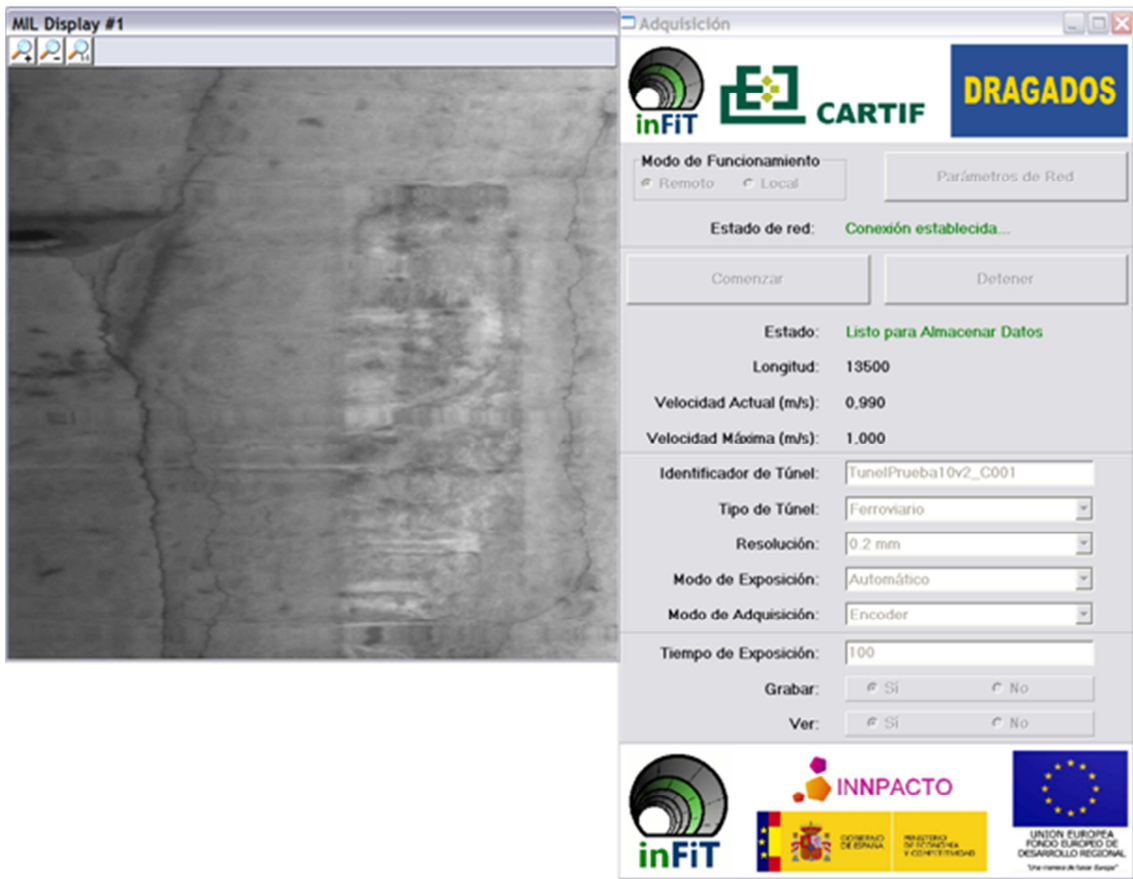
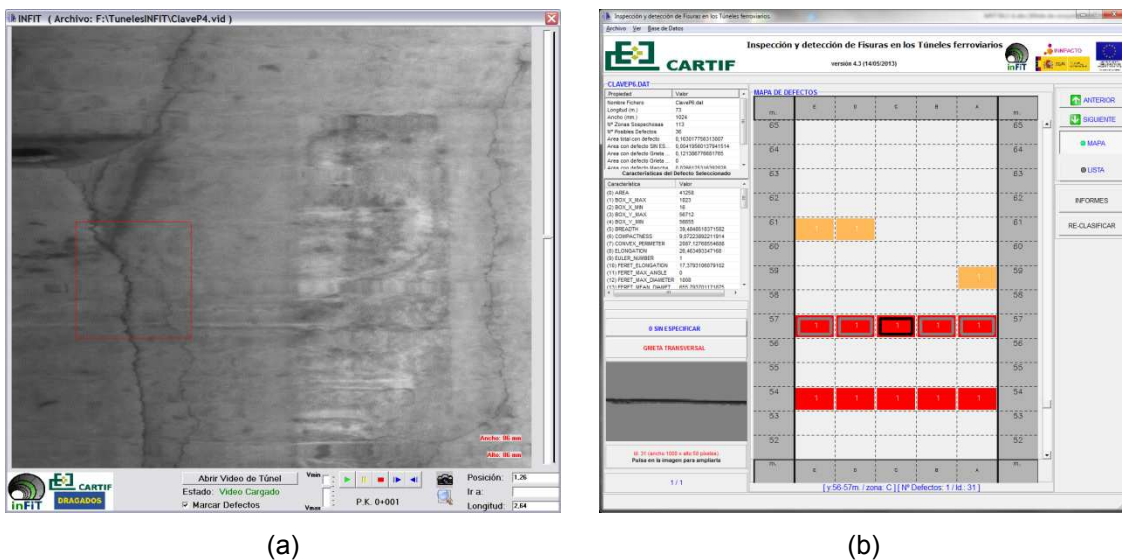


Figura 50. Captura de pantalla del módulo de adquisición de imágenes



(a)

(b)

Figura 51. Captura de pantalla del módulo de visualización de datos. (a) Datos brutos (b) Mapa de defectos

6.5 Conclusiones

En este capítulo se describen los sistemas que han sido implementados en esta tesis. Se trata de **tres sistemas de inspección visual**: una aplicación industrial para la inspección de la superficie de chapa laminada, un vehículo para la inspección de carreteras y un prototipo para la inspección del revestimiento de túneles.

Se han detallado las características de la **arquitectura** de los sistemas implementados, describiendo las principales etapas para la adquisición de las imágenes y el manejo de la información obtenida. El sistema de inspección de chapa laminada procesa la información en tiempo real, mientras que los otros dos sistemas procesan las imágenes almacenadas posteriormente.

También se han descrito los equipos utilizados para la adquisición de imágenes. En los tres casos se utilizan **cámaras lineales** para capturar las imágenes, un codificador incremental para generar la señal de captura de cada línea de la imagen y una iluminación que concentra su potencia lumínica en la zona de inspección de la cámara. Se utiliza iluminación LED y láser, ambos en el infrarrojo cercano, así como iluminación halógena en el espectro visible, adaptándose a las necesidades específicas de cada aplicación. El vehículo que inspecciona las carreteras también dispone de varios dispositivos auxiliares cuya información es capturada de forma sincronizada gracias a una electrónica de control que permite capturar los datos a distintas frecuencias de muestreo.

La información almacenada y procesada puede ser visualizada mediante un *software* que tiene un núcleo común a todas ellas y las funcionalidades específicas requeridas por cada aplicación, ofreciendo la información extraída conforme a las necesidades de cada tipo de usuario.

A continuación, en el Capítulo 7, se van a describir las técnicas de procesamiento de imágenes que han sido implementadas en cada aplicación y los resultados obtenidos para cada una de ellas.

7 RESULTADOS

En este capítulo se encuentran los resultados obtenidos en el procesamiento de las imágenes adquiridas con los sistemas de inspección implementados que se han descrito en la sección 6. Los resultados obtenidos para cada una de las aplicaciones son fruto de la utilización de las técnicas de procesamiento de imágenes descritas a lo largo de esta tesis. Para cada aplicación concreta se detallan las técnicas de procesamiento empleadas y el objetivo buscado.

El procesamiento utilizado en las tres aplicaciones propuestas tiene como principal elemento innovador la utilización de filtros de Gabor. El diseño de los filtros de Gabor se aborda de forma específica para cada aplicación, así como la forma de utilizar las imágenes filtradas para alcanzar los objetivos propuestos en cada caso.

En primer lugar, se muestran los resultados obtenidos en la detección y clasificación de defectos superficiales en chapa laminada. La principal novedad de las técnicas de procesamiento empleadas en esta aplicación es la utilización de características extraídas de imágenes filtradas por filtros de Gabor para la clasificación en línea de 6 tipos de defectos. En esta aplicación la detección de defectos se realiza de forma conservadora, de tal manera que gran parte de los defectos detectados inicialmente son finalmente descartados en la fase de clasificación. Una de las principales aportaciones en este trabajo es la forma en que se utilizan estos filtros, que permite su aplicación en tiempo real. Para la clasificación se han combinado distintos tipos de clasificadores que mejoran el resultado que se obtendría con un único clasificador.

La segunda aplicación abordada en este trabajo es la detección y clasificación de grietas longitudinales y transversales en carreteras. Se propone la utilización de dos clasificadores que determinan qué zonas de la imagen contienen cada tipo de grieta, de manera que se consigue en un solo paso la detección y la clasificación de los defectos. Para ello, la imagen se divide en pequeñas porciones y se determina si el defecto buscado se encuentra en cada porción empleando clasificadores binarios que utilizan como características el mayor valor del módulo de la imagen filtrada por un filtro de Gabor. La selección del umbral y el filtro de Gabor más apropiado para cada clasificador se abordan en esta aplicación. El algoritmo AdaBoost, adaptado a las características de esta aplicación, se utiliza para combinar varios clasificadores binarios y conseguir un único clasificador fuerte que obtiene mejores resultados que un único clasificador.

Finalmente, se muestran los resultados obtenidos en la detección de grietas en la superficie de túneles con revestimiento de hormigón. Para esta aplicación se comparan tres técnicas introducidas en esta tesis. La primera de ellas es el filtro de Gabor invariante a la rotación, un método para filtrar imágenes con filtros de Gabor que resalta defectos con una dirección predominante pero no conocida de antemano. La segunda, es el método de la textura de anisotropía condicional, en el que se proponen varias modificaciones que mejoran tanto los resultados como los tiempos de procesamiento empleados. Finalmente se propone un método para la combinación de los dos anteriores que alcanza mejores resultados que cualquiera de ellos por separado.

7.1 Detección y clasificación de defectos superficiales en chapa laminada

Para la detección de defectos superficiales en chapa laminada se ha desarrollado un sistema modular para la inspección automática de la superficie de bobinas de acero laminado que se describe con detalle en la sección 6.2. Cada módulo dispone de una cámara lineal, una iluminación infrarroja de alta potencia y su correspondiente unidad de procesamiento de imágenes. Gracias a la modularidad del sistema, su integración en el proceso industrial es más sencilla. El binomio cámara – iluminación lineales, que trabaja en un lugar intermedio entre el campo oscuro y el campo brillante, permite el realce de los defectos objetivo. Las imágenes adquiridas son preprocesadas, segmentadas, y clasificadas mediante la utilización de características extraídas a partir de las imágenes filtradas mediante filtros de Gabor.

7.1.1 Técnicas de procesamiento de imágenes utilizadas

Para la detección de defectos superficiales en chapa laminada se sigue un proceso de tres fases. En una primera fase de preprocesamiento de la imagen, ésta se homogeniza y a continuación se segmentan y agrupan los píxeles candidatos a ser defectos. En una segunda fase se realiza la extracción de características de los grupos de píxeles anteriormente seleccionados. Finalmente se realiza la clasificación de los defectos.

7.1.1.1 Preprocesamiento

El primer problema que se debe abordar es la falta de homogeneidad en los valores del brillo a lo largo de una línea de la imagen. El primer paso del preprocesamiento persigue obtener una imagen con un brillo homogéneo que permita segmentar la imagen de forma más precisa.

Al principio del proceso de inspección se ajustan automáticamente la ganancia y el tiempo de exposición de cada una de las cámaras para que la zona sin defectos tenga un valor intermedio de brillo (en este caso 128, que es el valor intermedio del rango de las cámaras utilizadas que se encuentra entre 0 y 255). La elección de este valor tiene por objetivo que tanto los defectos que se muestran más brillantes en la imagen como los más oscuros sean apreciados correctamente. A pesar de ello, el brillo de la imagen en la misma línea no es constante debido a que el ángulo de incidencia de la luz con respecto a la posición de la cámara no es exactamente el mismo para todos los puntos de la imagen. Por tanto, a cada imagen adquirida por cada una de las cámaras se le aplica la siguiente corrección:

$$I'_{ij} = 128 \frac{I_{ij}}{A_i^k} \quad i = 1, 2, \dots, M \quad j = 1, 2, \dots, N$$

$$A_i^k = \alpha A_i^{k-1} + (1 - \alpha) a_i^k \quad i = 1, 2, \dots, M$$

Ec. 42

donde $\alpha = \beta \frac{k-1}{k}$

I_{ij} es el valor de luminosidad del pixel de la imagen original de la columna i y de la fila j . I'_{ij} es el valor de luminosidad corregido. M es el número de columnas, N es el número de filas y k es el número de imagen. Además A_i^k es el valor ponderado de luminosidad de la columna i después de k imágenes y a_i^k es el valor medio de la luminosidad para la columna i de la imagen k . α es una variable de ponderación y β es una constante para controlar el peso de medidas cercanas en el tiempo que está fijada en 0,75. Se puede comprobar que α se aproxima al valor de β cuando el valor de k tiende a infinito.

Una vez aplicada esta operación sobre la imagen se lleva a cabo la segmentación de la misma. La segmentación de imágenes es utilizada para separar las zonas de la imagen que contienen defectos de las que no los contienen. Los defectos que aparecen sobre la chapa laminada pueden reflejar la luz en mayor o menor cantidad que la chapa laminada sin defectos, por lo que se utilizan dos umbrales para determinar que píxeles son candidatos a ser defectos. Por un lado se obtiene una imagen binarizada con los píxeles de la imagen que son más brillantes que el valor medio, mientras que por otro lado se obtiene una imagen binarizada con los píxeles que son menos brillantes que el valor medio. Cada una de estas imágenes es erosionada por un elemento estructural circular de tamaño 5 x 5 con el objetivo de eliminar píxeles aislados. Los píxeles adyacentes seleccionados en cada imagen se agrupan formando un único defecto para el que se calcula su nivel de gris medio y su área. Los defectos

suelen tener un área elevada o se desvían en gran nivel de gris medio. Para ponderar estas características se utiliza la siguiente ecuación:

$$\sqrt{(128 - G)^2 (K_A + A)} > K_L \quad \text{Ec. 43}$$

En esta ecuación, G es el nivel de gris medio, A es el área y K_A y K_L son dos parámetros de diseño. Para nuestro caso se ha fijado el valor K_A a cero y el valor de K_L a 180. Aumentando el valor K_A se consigue detectar defectos más pequeños a pesar de que su nivel de gris no se distancie demasiado del nivel medio. Aumentando el valor K_L se consigue aumentar el grado de detección y por tanto se detectarán más candidatos que no son defectos.

En la **Figura 52** se muestra un ejemplo del proceso de segmentación explicado anteriormente. Las imágenes (a) y (e) son las imágenes originales. En la imagen (b), los píxeles en blanco son los que tienen un valor inferior a 128, mientras que en la imagen (f) los píxeles en blanco son los que tienen un valor superior a 128. Las imágenes (b) y (f) son erosionadas mediante el elemento estructural circular de tamaño 5 x 5 dando como resultado las imágenes (c) y (g) respectivamente. Los grupos de píxeles de pequeño tamaño son eliminados obteniendo finalmente (d) y (h).

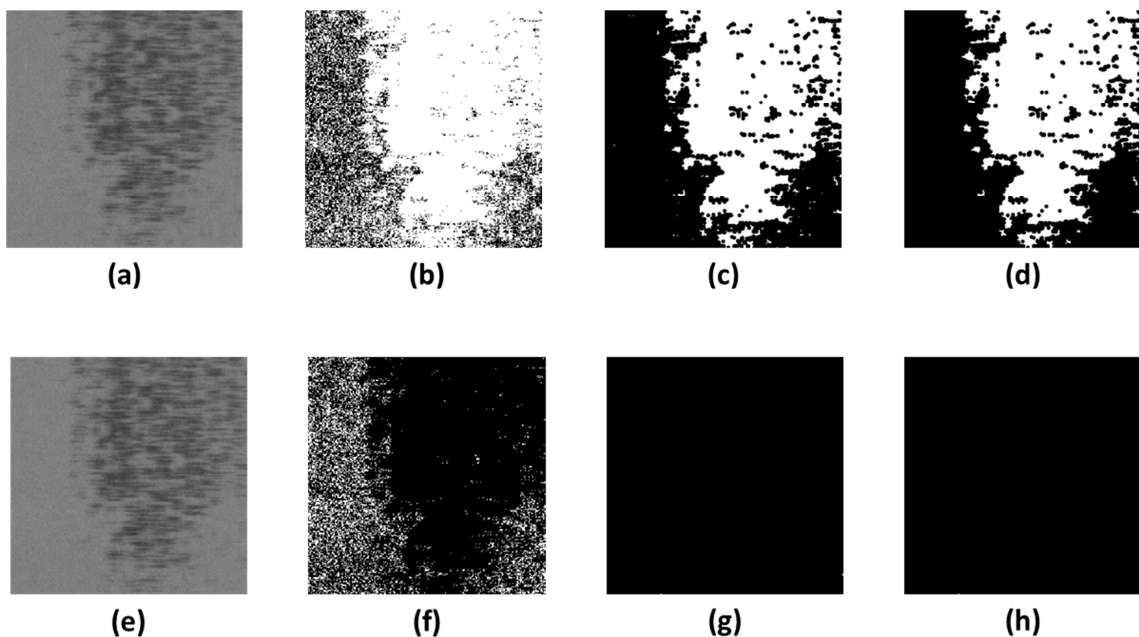


Figura 52. Proceso de segmentación

A continuación se mezclan las imágenes (d) y (h) para formar una única imagen binarizada. En este ejemplo la imagen (h) no contiene píxeles candidatos a ser

defectos por lo que la imagen final coincidirá con (d). Finalmente se agruparán los píxeles próximos, que serán etiquetados como posibles defectos y de los que se extraerán características en la próxima sección.

7.1.1.2 Extracción de características

Habitualmente se utilizan características relacionadas con la forma y el nivel de gris para realizar una clasificación automática. Sin embargo, debido a la gran variabilidad visual de los defectos que se pretende clasificar, no es posible obtener buenos resultados en la clasificación utilizando solo estas características. Con el objetivo de reducir al máximo la cantidad de defectos reales no detectados, durante la segmentación un gran número de posibles defectos es marcado como tal sin serlo. La distribución de los valores de las características consideradas tradicionales no presenta una clara diferenciación entre los distintos defectos ni, sobre todo, respecto al grupo de posibles defectos que realmente no son defectos. Esto hace muy difícil separar los defectos reales de los que no son defectos.

Es por ello por lo que se han utilizado las características extraídas de las imágenes filtradas por los filtros de Gabor, las cuales enriquecen la descripción de los defectos y facilitan la diferenciación entre los defectos de los no defectos, así como los distintos tipos de defectos entre sí.

Los filtros de Gabor que se van a utilizar en este trabajo han sido extraídos de un banco de filtros de Gabor cuya elaboración sigue los principios establecidos en la sección 5.3 de este trabajo, siguiendo el proceso resumido en la **Figura 32**, y que se explica a continuación:

1. En primer lugar, se ha fijado un **tamaño del filtro**, $M \times N$, de **256 x 256 píxeles**, ya que es una imagen lo suficientemente grande como para caracterizar el defecto y no demasiado grande como para no permitir implementar el procesamiento en la línea de producción.
2. El **número de orientaciones**, n_θ , se ha establecido en **8**, por lo que el **ancho de banda en frecuencia**, B_θ , es de **22,5°**
3. El **ancho de banda en frecuencia**, B_F , es **1 octava**.
4. No se considera solapamiento entre los filtros en ninguna de las direcciones por lo que las **constantes de solapamiento**, K_x y K_y , valen **1**.
5. De estos parámetros se deduce la **frecuencia central más alta**, F_M , **3/8**, que ha sido calculada de tal manera que la curva de nivel de la mitad de valor de pico del módulo de la transformada de Fourier tenga 0,5 como mayor valor de la frecuencia.
6. El **número de frecuencias del banco de filtros**, n_F , es **4**, siendo por tanto la frecuencia central más baja $3/64$, que es mayor que $N^{-1}=1/256$.

7. Resultados

7. Para cada uno de los 4 valores de frecuencia se han calculado los valores de dispersión, σ_x y σ_y .

El resultado es un banco de filtros de Gabor compuesto por 32 filtros de Gabor obtenidos al sustituir cada uno de los valores $F(i, j)$, $\theta(i, j)$, $\sigma_x(i, j)$ y $\sigma_y(i, j)$ en la Ec. 19, que es la ecuación de un filtro de Gabor.

En la **Tabla 7** se muestran los parámetros del filtro de Gabor utilizado.

Tabla 7. Parámetros del banco de filtros de Gabor para la clasificación de defectos en chapa laminada

Parámetro	Unidades	Valor
Tamaño (MxN)	Píxeles x Píxeles	256 x 256
Número de Frecuencias (n_f)	-	4
Número de Orientaciones (n_θ)	-	8
Ancho de Banda en Frecuencia (B_f)	Octavas	1
Ancho de Banda en Frecuencia (B_θ)	Grados	22,5
Constante de solapamiento en x (K_x)	-	1
Constante de solapamiento en y (K_y)	-	1
Frecuencia central máxima (F_M)	Píxeles ⁻¹	0,375

Cuando el objetivo es segmentar una imagen, los filtros de Gabor necesarios son aplicados a la totalidad de la imagen, dividiéndola en trozos de tamaño fijo y con un cierto solapamiento. A cada uno de estos trozos de imagen se le aplica el teorema de la convolución, tal y como se describe en la sección 5.4. Aplicar los 32 filtros de Gabor a la imagen completa es una tarea inabordable en línea, debido al elevado número de convoluciones a realizar. Para su reducción se han seguido las siguientes dos estrategias.

- **Reducir el número de filtros del banco de filtros:** Este número ha sido reducido a 5, utilizando solamente los filtros que mayor respuesta tenían ante los defectos habituales.

- **Aplicar cada filtro solo una vez para cada posible defecto:** Como nuestro objetivo es describir mediante características a cada uno de los posibles defectos detectados en la fase de segmentación, los filtros solo son aplicados a las zonas de la imagen marcadas como posibles defectos. Para reducir aún más los tiempos de procesamiento, se han aplicado los filtros una única vez por posible defecto detectado, centrándolo en una zona del defecto. Se ha observado que habitualmente la mayor información frecuencial dentro de un defecto se encuentra en los alrededores del píxel cuyo nivel de gris dista más del valor medio buscado (128), por lo que es ese píxel (o tan cerca como sea posible cuando ese píxel se encuentra cerca del borde de la imagen) donde se ha situado el centro del trozo de la imagen de 256 x 256 al que se aplica cada filtro de Gabor.

La característica de la imagen filtrada que se va a evaluar es el **valor medio del módulo**. Como dentro del trozo de imagen seleccionada para aplicar los filtros de Gabor hay píxeles que corresponden a defectos y otros que no, esa media se va a calcular solamente con los píxeles que han sido marcados como posibles defectos en la fase de segmentación, de tal manera que las características que se van a utilizar en la clasificación se ven menos afectadas por los píxeles que se encuentran alrededor de un defecto, y que no pertenecen al defecto.

Además de las características extraídas de la aplicación de los filtros de Gabor, se ha considerado también combinarlas con las características relacionadas con el brillo y la forma del defecto. Concretamente se han utilizado cuatro características de este tipo: nivel de gris medio, ángulo principal, área y el mayor diámetro de Feret.

Los 5 filtros de Gabor seleccionados han sido seleccionados de tal manera que sean los que ofrecen una respuesta más alta ante los 6 defectos estudiados. Los valores de sus parámetros pueden verse en la **Tabla 8**.

Tabla 8. *Parámetros de los filtros de Gabor utilizados*

Filtro	Frecuencia (píxeles ⁻¹)	Orientación (°)	Dispersión en X (píxeles)	Dispersión en Y (píxeles)
1	0,094	0	6,00	10,05
2	0,094	45	6,00	10,05
3	0,094	112,5	6,00	10,05
4	0,047	90	11,99	20,10
5	0,375	0	1,50	2,51

El comportamiento de estos filtros puede verse en la **Figura 53**. En esta figura, se muestra un ejemplo de cada tipo de defecto y de la imagen filtrada por uno de los anteriores filtros. Con objeto de poder visualizarlas correctamente, las imágenes filtradas están normalizadas de tal manera que el valor más bajo de la imagen filtrada toma el valor de 0 y el valor más alto toma el valor de 255. Puede observarse que las zonas defectuosas son las que toman los valores más altos, y se encuentran habitualmente en el interior de las zonas seleccionadas en la fase de segmentación.

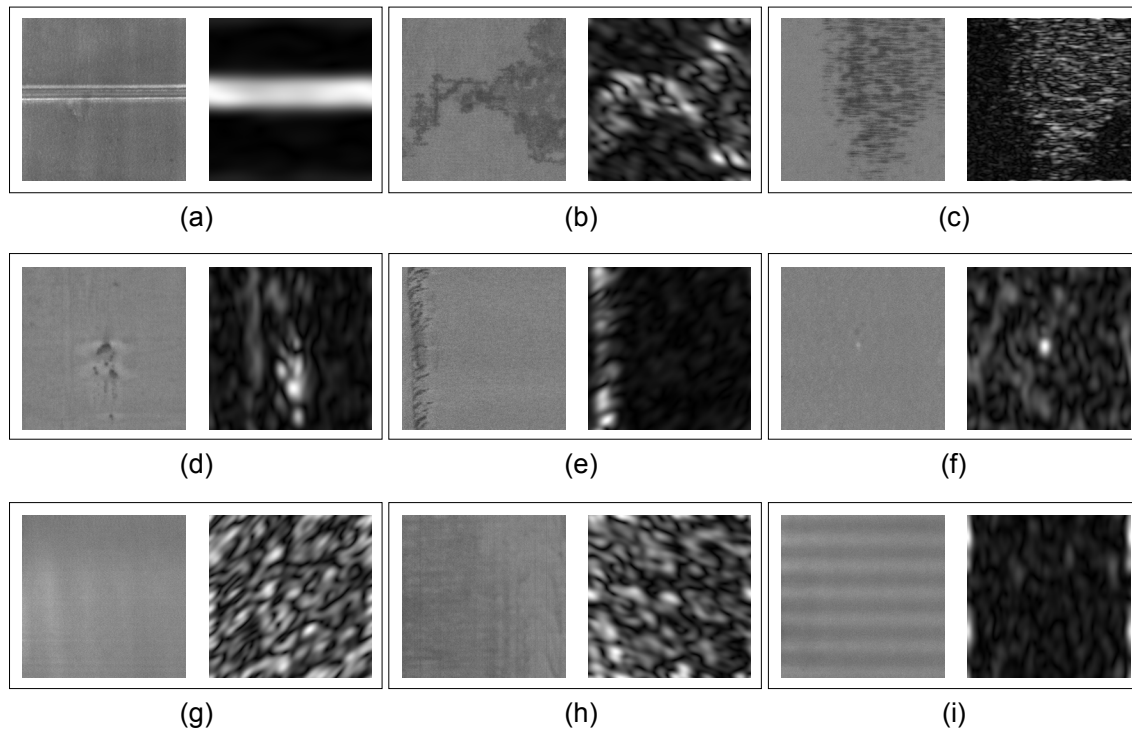


Figura 53. Imágenes de los defectos y sus correspondientes imágenes filtradas por uno de los filtros seleccionados. (a) soldadura, (b) moho blanco, (c) marcas de asentamiento, (d) calvas, (e) grumos de cinc, (f) impresiones, (g), (h), (i) son posibles defectos que no son defectos. El Filtro 1 de la **Tabla 8** se aplica a (d), (c), (i). El Filtro 2 a (e), (h). El Filtro 3 a (b), (g). El Filtro 4 a (a). El Filtro 5 a (c)

Por tanto, al utilizar solo los píxeles marcados como posibles defectos en la fase anterior aumenta el valor obtenido como media del módulo de la imagen filtrado, lo que mejora los resultados de la clasificación que se lleva a cabo en la siguiente fase. Sin embargo, en las imágenes que no son defectos, las de la última fila de la **Figura 53**, las zonas más brillantes de la imagen se encuentran distribuidas por toda ella, pero con valores muy bajos, próximos a cero. Se debe tener en cuenta que, para los ejemplos mostrados, el valor más alto de las imágenes filtradas que contienen defecto es entre 2 y 10 veces mayor que para las imágenes que no contienen defectos, aunque en la representación gráfica de la figura se muestran con el mismo valor, 255. Esto

ocurre en la gran mayoría de las ocasiones para alguno de los filtros, lo que posibilita diferenciar los defectos de los no defectos en la fase de clasificación.

7.1.1.3 Clasificación

Las nueve características extraídas de las imágenes con posibles defectos descritas en el apartado anterior se utilizan para alimentar tres diferentes tipos de clasificadores: una red neuronal, el algoritmo de los vecinos cercanos y un clasificador bayesiano. Así mismo se ha realizado la combinación de estos tres clasificadores para mejorar los resultados obtenidos.

La arquitectura de la red neuronal es 9/12/6, de tal manera que tiene 9 entradas y 6 salidas. Las entradas corresponden a las nueve características descritas en el apartado anterior, mientras que cada una de las salidas corresponde a cada uno de los seis tipos de defectos seleccionados en esta tesis. Cuando ninguna de las salidas alcanza el valor mínimo fijado, este candidato es asignado al grupo sin defecto. La red ha sido entrenada usando el algoritmo *backpropagation* tal y como se describe en la sección 4.5.1.

El clasificador de vecinos cercanos utilizado en este trabajo comprueba las tres muestras de la base de datos cuyas características se encuentran más próximas a la muestra que se va a clasificar. Se utiliza la variante con rechazo, es decir, si dos o más de esas muestras corresponden al mismo defecto, se le asigna ese grupo, si no, no es clasificado y es asignado al grupo de sin defecto. Para incrementar la velocidad de procesamiento se utiliza el algoritmo de búsqueda *X-tree*, que reduce drásticamente el número de comparaciones, [Ver96].

Por último, también se ha implementado el clasificador *naïve Bayes* para clasificar los defectos tal y como se describe en la sección 4.5.3. Para cada uno de los seis tipos de defectos objeto de estudio se ha realizado una clusterización de los vectores de características utilizando el algoritmo *K-means*, dada la gran variación existente en los valores de las características entre defectos del mismo tipo. Cuando la probabilidad de un nuevo caso no alcanza una probabilidad mínima de pertenecer a uno de los clusters se considera que pertenece a la clase sin defecto.

La combinación de clasificadores siempre mejora el comportamiento de un clasificador aislado, [Kon07]. Esto es debido a que cada clasificador produce errores en una región diferente del espacio de patrones de entrada. El principal problema radica en determinar cuál es la mejor estrategia para combinar clasificadores simples. En este caso se ha optado por la estrategia de *voto por mayoría*. Esta estrategia es apropiada para combinar clasificadores cuya salida son etiquetas, como es el caso de los tres clasificadores propuestos.

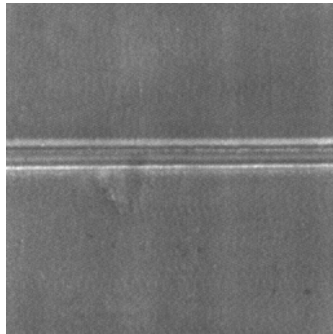
7.1.2 Resultados obtenidos

El sistema ha sido instalado, entrenado y verificado en dos líneas de corte de una de las empresas del grupo Gonvarri situada en Burgos. La información de cada una de las caras de las bobinas se adquiere con 5 cámaras líneas AVIIVA 1010 SM2 de la marca Atmel, que tienen una resolución de 1.024 píxeles. Cada bobina procesada tiene una longitud media de 2 km por lo que se generan alrededor de 10GB de información por cada cara. La información de cada cara es procesada mediante 6 ordenadores, que envían la información a un ordenador que hace de interfaz con el usuario, tal y como se describe en la sección 6.2. El procesamiento de las imágenes se produce a la misma velocidad de la línea, de tal manera que los resultados son mostrados inmediatamente en el puesto de control, generando alarmas cuando aparece el defecto de soldadura y cuando el número de defectos de un tipo se encuentra por encima de los límites especificados.

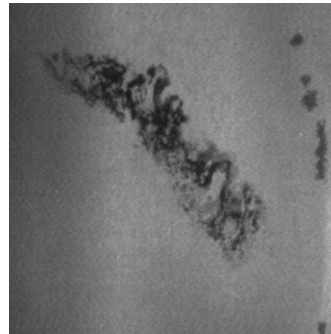
El objetivo buscado es la detección y clasificación de 6 tipos de defectos, los cuales son los que aparecen más frecuentemente en las líneas de producción donde se encuentra instalado el sistema y de los cuales se muestra un ejemplo en la **Figura 54** que ha sido adquirido con el sistema de inspección desarrollado en esta tesis:

- a) **Soldadura:** es la zona de unión entre dos bobinas para conseguir una bobina de una longitud mayor. Es un defecto que debe ser detectado inmediatamente. En la línea de corte longitudinal se marca su posición y en la línea de corte transversal es retirada la pieza que contiene la soldadura, ya que en caso de que una soldadura llegase a una prensa dañaría el útil.
- b) **Moho blanco:** es la oxidación de la superficie de la bobina que se produce al haber estado la bobina en contacto con el agua o haberse producido condensación. A pesar de su nombre, su aspecto suele ser oscuro.
- c) **Marcas de transporte:** se producen durante el transporte de las bobinas, normalmente en ferrocarril, debido al rozamiento continuado entre las caras de la bobina. Se producen en la zona donde descansa la bobina por lo que son cíclicas, con una frecuencia de aparición del perímetro de la bobina.
- d) **Calvas:** son pequeñas zonas de la bobina que carecen del recubrimiento que las protege de la oxidación, bien porque lo han perdido o porque este no se aplicó correctamente.
- e) **Grumos de cinc:** son abultamientos de cinc, debido a que el galvanizado no se aplicó homogéneamente.
- f) **Impresiones:** son pequeñas marcas provocadas por el paso de las bobinas entre los rodillos en alguno de sus procesos. La presencia de una partícula de suciedad provoca que se produzcan estas impresiones que se muestran normalmente de forma cíclica. El periodo de aparición de las impresiones coincide con el perímetro del rodillo que las provocó. También aparecen porque una partícula quedó atrapada en la superficie de la bobina cuando se

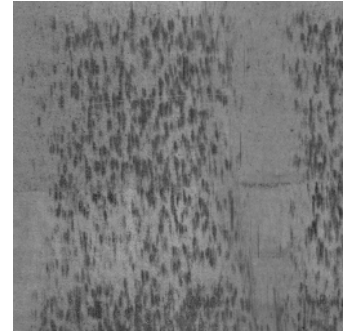
enrollo, provocando que aparezca la impresión en varias vueltas de la bobina. En este caso el periodo de aparición coincide con el perímetro de la bobina.



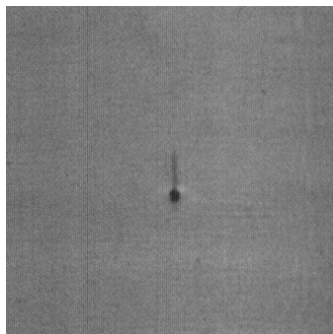
a) Soldadura



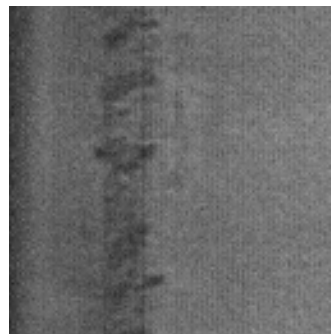
b) Moho blanco



c) Marcas de transporte



d) Calvas



e) Grumos de cinc



f) Impresiones

Figura 54. Muestras de los defectos objeto de estudio adquiridos por el sistema durante su funcionamiento

Para entrenar el sistema final es necesario disponer de un número de significativo de defectos. El número de defectos presentes en las bobinas que habitualmente son procesadas en estas líneas de corte suele ser muy bajo, por lo que utilizar estas bobinas para el entrenamiento supondría un periodo de entrenamiento muy elevado. Afortunadamente para nuestro propósito, ocasionalmente aparecen bobinas con un elevado número de defectos. Estas bobinas no son válidas para ningún uso final y se pueden considerar desechos de producción. Los responsables de calidad de la empresa almacenan las bobinas desechadas y son precisamente éstas las que van a ser utilizadas para entrenar nuestro sistema de inspección, lo que ha reducido el tiempo necesario para crear una base de datos de defectos lo suficientemente amplia para entrenar el sistema.

Las imágenes adquiridas en un entorno industrial, como es una línea de corte de bobinas, presentan una serie de problemas que no aparecen en un entorno de laboratorio. Algunos de los inconvenientes encontrados en la línea de producción son:

- Las bobinas son **lubricadas** con diferentes tipos de aceite para que no se dañen durante el proceso. Esta lubricación puede provenir de procesos anteriores o puede aplicarse en la propia línea de corte, aguas arriba del proceso de inspección visual. La distribución del aceite no siempre es homogéneo.
- Es frecuente la presencia de **suciedad** sobre la bobina, bien porque la haya adquirido durante el transporte, o bien porque durante el proceso de desenrollado la adquiera debido al contacto con los distintos elementos de la línea de producción.
- El ángulo que forman la cámara y la iluminación con la normal a la superficie de la chapa laminada debe ser fijo para que la imagen tenga siempre las mismas condiciones de iluminación. Sin embargo, una bobina de chapa laminada moviéndose a elevada velocidad produce **vibraciones** impredecibles que suponen cambios en la forma de iluminar la chapa laminada. A ambos lados de la zona de inspección se han instalado 2 pares de rodillos que mediante presión neumática tratan de mantener la bobina lo más estable posible. Aunque se reducen considerablemente las vibraciones, éstas siguen apareciendo, especialmente en los comienzos y finales de las bobinas.

El proceso de detección de defectos consiste en determinar qué zonas de la imagen tienen una apariencia visual que se aleja del patrón que aparece en la bobina sin defectos. Las características de la línea de producción enumeradas anteriormente, provocan la aparición de zonas en la imagen que inicialmente son consideradas como defecto y que realmente no lo son. Es por ello que se incluye este tipo de defecto en el proceso de clasificación. También se debe tener en cuenta que un mismo tipo de defecto puede presentar una apariencia visual muy diversa, tal y como se puede ver en algunos ejemplos en la **Figura 55**.

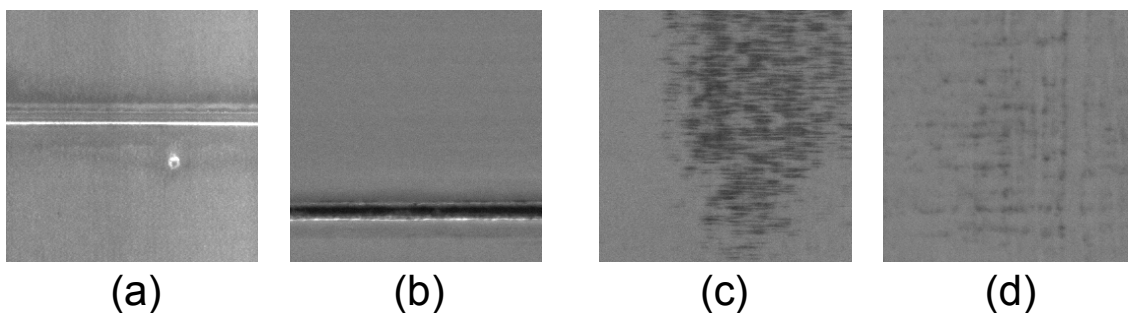


Figura 55. Dos soldaduras (a) y (b); y dos marcas de transporte (c) y (d). Puede verse la gran variabilidad en la apariencia visual de defectos del mismo tipo

Un *software* específicamente diseñado para este propósito ha sido desarrollado para entrenar el sistema. Permite el análisis de una gran cantidad de características de las imágenes y estudiar sus correlaciones, de tal manera que se pueden seleccionar las que mejor se adaptan al objetivo buscado.

Para poder entrenar el sistema, las muestras deben ser clasificadas manualmente. Para ello, el *software* detecta de forma automática los posibles defectos y lleva a cabo una preclasificación. Los resultados son mostrados a los responsables de calidad de la empresa para que acepten o modifiquen los resultados de la preclasificación. Estas imágenes, etiquetadas con el tipo de defecto asignado por los expertos en calidad, son las que se utilizan para el entrenamiento. El entrenamiento debe ser llevado a cabo con el suficiente cuidado de no provocar un sobreentrenamiento, es decir, la pérdida de generalidad de los algoritmos de clasificación debido a que se adaptan en exceso a las muestras utilizadas para entrenarlo. Por este motivo, solo la mitad de los defectos son utilizados para entrenar, mientras que la otra mitad se utiliza para verificar que el entrenamiento ha mejorado. Cuando el nuevo entrenamiento no mejora los resultados obtenidos por el entrenamiento anterior, se descarta y no se utiliza para la siguiente clasificación.

Además, el *software* desarrollado para hacer de interfaz entre el usuario y el sistema, permite al usuario analizar los resultados de la inspección de la bobina y visualizar las imágenes de la bobina en tiempo real y fuera de línea. Los resultados obtenidos una vez procesada la bobina también pueden ser analizados desde un lugar remoto.

La segmentación de las imágenes se ha ajustado de tal manera que se detectan una gran cantidad de candidatos a ser defectos, con el objetivo de perder la menor cantidad de defectos reales que sea posible. En realidad, aproximadamente el 75% de los posibles defectos detectados finalmente no son considerados como defectos reales. Se han inspeccionado 12 bobinas seleccionadas por el personal de calidad, con una longitud total de más de 20km. Varios tipos de acero con diferentes acabados superficiales han sido utilizados.

Además, se han incluido en la base de datos de defectos 20 soldaduras de otras bobinas. En total, las bobinas contenían 18.532 defectos: 20 soldaduras, 2.268 mohos blancos, 1.298 marcas de transporte, 6.690 calvas, 2.680 grumos de zinc y 5.576 impresiones. El sistema etiquetó 71.286 zonas de la imagen como posibles defectos de los cuales 52.754 no eran defectos. La mitad de estas muestras, elegidas al azar, se utilizan para realizar el entrenamiento mientras que la otra mitad son utilizadas para verificar sus resultados.

7. Resultados

La clasificación de los defectos analizados se realiza utilizando tres clasificadores distintos. Los resultados obtenidos por cada uno de los clasificadores se encuentran desglosados en las siguientes tablas. En cada una de estas tablas se encuentra el resultado obtenido por cada uno de los clasificadores desarrollados en esta tesis al clasificar los defectos que previamente fueron clasificados por los expertos. En estas tablas se muestran los resultados obtenidos para las muestras test, es decir, las utilizadas para verificar los resultados del entrenamiento.

A continuación se muestran los resultados obtenidos al aplicar los tres clasificadores propuestos en esta tesis. En las siguientes tres tablas se muestran las matrices de confusión para cada uno de los clasificadores tal y como se describe en la sección 4.2.2. En la **Tabla 9**, se muestran los resultados del clasificador que utiliza una red neuronal, en la **Tabla 10** los obtenidos al aplicar el clasificador de vecinos cercanos y en la **Tabla 11** se los correspondientes al clasificador Bayesiano. Por último, en la **Tabla 12**, se muestra el resultado de combinar los tres clasificadores anteriores, en los que se aprecia una considerable mejora en los resultados obtenidos.

Los resultados obtenidos con cada uno de los tres clasificadores propuestos en este proyecto en términos de sensibilidad y especificidad se encuentran en la **Tabla 13**, así como los obtenidos utilizando la combinación de los tres.

Tabla 9. Matriz de confusión del clasificador red neuronal (ANN)

		Clasificación del clasificador						
		Soldadura	Moho blanco	Marcas de transporte	Calvas	Grumos de cinc	Impresiones	Sin defecto
Clasificación del experto	Soldadura	10	0	0	0	0	0	0
	Moho blanco	12	849	85	6	12	0	170
	Marcas de transporte	1	54	485	5	7	0	97
	Calvas	0	15	16	2.844	25	17	428
	Grumos de cinc	15	31	24	20	1.164	2	84
	Impresiones	0	0	1	30	17	1.820	920
	Sin defecto	92	86	86	59	92	117	25.845

Tabla 10. Matriz de confusión del clasificador de vecinos cercanos (KNN)

		Clasificación del clasificador						
		Soldadura	Moho blanco	Marcas de transporte	Calvas	Grumos de cinc	Impresiones	Sin defecto
Clasificación del experto	Soldadura	10	0	0	0	0	0	0
	Moho blanco	22	883	104	9	58	0	58
	Marcas de transporte	2	54	496	4	8	1	84
	Calvas	0	15	26	2.852	33	40	379
	Grumos de cinc	11	31	34	20	1.180	3	61
	Impresiones	0	0	1	30	47	1.975	735
	Sin defecto	92	149	161	119	155	148	25.553

Tabla 11. Matriz de confusión del clasificador Bayesiano

		Clasificación del clasificador						
		Soldadura	Moho blanco	Marcas de transporte	Calvas	Grumos de cinc	Impresiones	Sin defecto
Clasificación del experto	Soldadura	10	0	0	0	0	0	0
	Moho blanco	15	857	73	11	20	0	158
	Marcas de transporte	2	13	414	9	13	2	196
	Calvas	0	9	17	3.082	27	18	192
	Grumos de cinc	7	18	27	7	1.193	2	86
	Impresiones	0	0	0	21	21	1.961	785
	Sin defecto	141	157	180	108	98	162	25.531

7. Resultados

Tabla 12. Matriz de confusión de la combinación de los tres clasificadores empleados

		Clasificación del clasificador						
		Soldadura	Moho blanco	Marcas de transporte	Calvas	Grumos de cinc	Impresiones	Sin defecto
Clasificación del experto	Soldadura	10	0	0	0	0	0	0
	Moho blanco	2	957	45	0	3	0	127
	Marcas de transporte	0	1	527	0	2	0	119
	Calvas	0	0	7	3.210	0	1	127
	Grumos de cinc	0	2	21	0	1.272	1	44
	Impresiones	0	0	0	2	0	2.124	662
	Sin defecto	1	0	6	1	0	1	26.368

Tabla 13. Resultados de sensibilidad (Sen.) y especificidad (Esp.) de las muestras utilizadas para la red neuronal (ANN), el método de los vecinos cercanos (KNN), el clasificador Bayesiano y la combinación de los tres, aplicado a cada uno de los defectos estudiados

	ANN		KNN		Bayesiano		Combinado	
	Sen.	Esp.	Sen.	Esp.	Sen.	Esp.	Sen.	Esp.
Soldadura	1,000	0,997	1,000	0,996	1,000	0,995	1,000	1,000
Moho blanco	0,749	0,995	0,779	0,993	0,756	0,994	0,844	1,000
Marcas de transporte	0,747	0,994	0,764	0,991	0,638	0,992	0,812	0,998
Calvas	0,850	0,996	0,853	0,994	0,921	0,995	0,960	1,000
Grumos de cinc	0,869	0,996	0,881	0,991	0,890	0,995	0,949	1,000
Impresiones	0,653	0,996	0,708	0,994	0,703	0,994	0,762	1,000
Sin defecto	0,980	0,817	0,969	0,858	0,968	0,847	1,000	0,884

En la **Tabla 13** se puede observar que la sensibilidad tiene un valor elevado en todos los casos. El valor de la especificidad es elevado para todos los grupos excepto para el grupo sin defecto. Este grupo es el que mayor cantidad de muestras tiene, lo que provoca que el número de verdaderos negativos para el resto de los grupos sea muy elevado. Sin embargo, el número de verdaderos negativos para el grupo sin defecto es relativamente bajo. Se puede comprobar que la sensibilidad y la especificidad consiguen mejores resultados para el clasificador combinado para todos los grupos de defectos y el grupo de no defectos.

Los defectos para los que se obtienen peores valores en la clasificación son las marcas de transporte y las impresiones. El grupo de sin defectos tiene una elevada sensibilidad pero tiene un elevado valor de falsos positivos, manteniéndose siempre por debajo del 20%. En este caso, la clasificación combinada alcanza prácticamente el 100% de especificidad con un valor de sensibilidad del 88%.

También se puede observar que el 100% de las soldaduras se clasifican correctamente, con un valor de especificidad elevado. La detección de las soldaduras es de vital importancia porque si una soldadura entra en una máquina de prensado puede ocasionar serios daños en la misma. Las impresiones se clasifican correctamente solo en el 76,2% de los casos debido a la dificultad de extraer características de un defecto tan pequeño. Como en la práctica este tipo de defecto aparece de forma cíclica en una bobina, su detección está prácticamente asegurada. Por ejemplo, si aparecen tres impresiones consecutivas, la probabilidad de detectar al menos una de ellas alcanza el 98,7%, y este valor se eleva al 99,9% en el caso de que aparezcan 5 de forma consecutiva.

La clasificación que utiliza la combinación de los otros tres métodos de clasificación mejora los resultados de cada uno de los métodos de forma individual. En concreto, el número de falsos positivos desciende drásticamente hasta ser casi nulo en todos los tipos de defectos excepto en las marcas de transporte. Esto seguramente es debido a que existe un gran parecido entre las marcas de transporte y el moho blanco. A pesar de ello, los resultados en este caso son aceptables. Además la mayoría de los posibles defectos que en realidad no son defectos son clasificados como sin defecto por el clasificador combinado.

El 87,42% de los defectos utilizados para verificar el comportamiento de los clasificadores han sido clasificados correctamente, lo que puede considerarse un buen resultado. El comportamiento del sistema también puede evaluarse en términos de exactitud, es decir, el número de imágenes correctamente clasificadas del total de imágenes. Una exactitud global del 96,70% ha sido obtenida para el clasificador combinado, lo que confirma la fiabilidad del sistema.

En términos generales, los índices de acierto obtenidos son mejores que los obtenidos en otros trabajos, aunque para comparar correctamente estos resultados se requeriría que hubieran sido aplicados al mismo tipo de material, tipo de defectos y bajo las mismas condiciones. En [Yaz08] se utiliza una red neuronal para clasificar 5 defectos obteniendo una exactitud global del sistema del 97,19%, similar a la que se obtiene en el presente trabajo, pero en la que se considera que a la fase de clasificación solo llegan defectos. En [Mas12] se clasifican 7 tipos de defectos con una red neuronal alcanzando un 93,03% de exactitud global. Al igual que en el caso anterior se considera que todas las imágenes utilizadas para realizar el test contienen alguno de los 7 defectos estudiados. En [Dup97] se alcanza una exactitud global del 79,3% en la clasificación de 8 tipos de defecto mediante el algoritmo de vecinos cercanos. En [Yun11] se clasifican 6 tipos de defectos utilizando tres algoritmos distintos: el algoritmo de vecinos cercanos, una red neuronal y una máquina de soporte vectorial, obteniendo una exactitud global del 80%, 82,86% y 88,57% respectivamente. Estos autores proponen la utilización del algoritmo AdaBoost para combinar varios clasificadores alcanzando el 94% de exactitud global.

Resulta complicado realizar una comparación objetiva con los resultados obtenidos en estos trabajos, principalmente porque la mayoría de los autores presuponen que en la fase de clasificación solo se encuentran imágenes de los defectos esperados. Sin embargo, en esta tesis se utiliza la información obtenida en una instalación real en la que gran cantidad de los posibles defectos que llegan a la fase de clasificación no son realmente defectos. De esta forma se consigue que la mayor parte de los defectos presentes sobre la bobina lleguen a la fase de clasificación. Aun así, los resultados están en línea con los obtenidos en estos trabajos.

Por otro lado, el sistema está funcionando en dos líneas de producción real, y los responsables de calidad de la empresa no han informado de que se haya producido una disminución en el rendimiento del sistema a lo largo del tiempo.

7.1.3 Conclusiones

Se ha abordado la aplicación de filtros de Gabor para la detección y clasificación de defectos sobre chapa laminada. Para ello se ha desarrollado e implantado un sistema de inspección visual automático en dos líneas de corte de chapa laminada, resolviendo los problemas encontrados en un entorno industrial real. Este trabajo se centra en la detección y clasificación de seis tipos de defectos: soldaduras, moho blanco, marcas de transporte, calvas, grumos de cinc e impresiones.

Las principales técnicas de procesamiento que han permitido alcanzar los objetivos planteados son la homogenización de las imágenes y la combinación de características extraídas de las imágenes filtradas por filtros de Gabor con otro tipo de características

más habituales. Los defectos que se pretende clasificar, presentan diferentes características de tamaño y forma para un mismo tipo de defecto, lo que dificulta la utilización de características habituales para la clasificación.

Se ha diseñado un banco de filtros de Gabor que alcance una frecuencia central óptima y permita ajustar el solapamiento. Las imágenes han sido filtradas por filtros de Gabor, y se ha calculado el valor medio del módulo de la imagen filtrada utilizando sólo los píxeles marcados como defectos en la segmentación. Estas características, junto a otras más habituales relacionadas con la forma y el brillo, han sido utilizadas como entrada de tres clasificadores. Se han combinado estos tres clasificadores obteniendo mejores resultados que utilizándolos de forma separada, tal y como puede verse analizando los valores de sensibilidad y especificidad calculados con las muestras test.

La solución global ha sido implementada y probada en una línea de producción industrial, utilizando una arquitectura *hardware/software* de tres niveles. Los resultados obtenidos muestran que el sistema se comporta correctamente para la detección y clasificación de los defectos objetivo en diferentes tipos de aceros.

La comparación con otros métodos similares resulta compleja debido a la gran diversidad de tipos de acero existentes en el mercado, a que las técnicas analizadas son aplicadas a distintos tipos de defectos, las condiciones de adquisición de las imágenes no suelen ser en instalaciones reales en funcionamiento y los resultados son medidos utilizando diversas métricas según el objetivo buscado. Además la mayoría de los trabajos presuponen que a la fase de clasificación solo llegan imágenes de los defectos esperados. En estas condiciones los trabajos analizados presentan una exactitud global que varía entre el 80% y el 97%, por lo que se puede concluir que los resultados obtenidos en la presente tesis son satisfactorios, ya que la exactitud global del clasificador combinado desarrollado es del 96,70%, clasificando correctamente el 87,42% de los defectos. Estos resultados han sido obtenidos con imágenes adquiridas en condiciones normales de funcionamiento. Las soldaduras, las calvas y los grumos de cinc se clasifican correctamente por encima del 95% de los casos. El grupo sin defectos, que corresponde a las zonas de la imagen que han sido etiquetadas en la fase de segmentación como posiblemente defectuosas pero que realmente no lo son, prácticamente es clasificado correctamente en todos los casos. Los defectos que suelen aparecer de forma cíclica, las marcas de transporte y las impresiones, aunque presentan a priori peores valores de detección, en la práctica los índices de detección son elevados, ya que, la probabilidad de detectar alguno de una secuencia relativamente corta, es prácticamente el 100%.

7.2 Detección y clasificación de grietas en la superficie de la carretera

Para la detección y clasificación de grietas en la superficie de la carretera se utiliza el sistema de adquisición de imágenes LRIS (*Laser Road Imaging System*) de INO, [INO15], el cual permite adquirir imágenes de la carretera con alta calidad.

Sin embargo, las condiciones variables de luz ambiental y las sombras que pueden aparecer, y sobre todo, la naturaleza del asfalto que está compuesto por áridos y betún, materiales que aparecen en las imágenes con un alto contraste entre ellos y cuya distribución se ve afectada por la degradación de la carretera, hacen que las habituales técnicas basadas en segmentación no funcionen satisfactoriamente y la detección eficiente de grietas sigue siendo un reto abierto.

El sistema de adquisición de imágenes utilizado permite obtener imágenes con una influencia mínima de las condiciones lumínicas ambientales. A pesar de ello, no puede despreciarse la influencia del ángulo de iluminación, por lo que se ha desarrollado un algoritmo de preprocesamiento de imágenes para mitigarlo.

En este trabajo se propone un nuevo método para la detección de grietas transversales y longitudinales utilizando filtros de Gabor. La utilización de filtros de Gabor para la detección de defectos ofrece un amplio rango de posibilidades que son analizadas en toda su amplitud. Para ello se evalúan diversos métodos para elegir la forma de utilizar las imágenes filtradas. También se utiliza el algoritmo AdaBoost para la combinación de clasificadores, que dan como resultado un nuevo clasificador con mejores resultados que un clasificador sencillo.

A continuación se describen las técnicas de procesamiento de imágenes utilizadas, los resultados obtenidos al aplicar estas técnicas y una comparativa con los resultados obtenidos en otros trabajos.

7.2.1 Técnicas de procesamiento de imagen utilizadas

El procesamiento de las imágenes obtenidas por el equipo LRIS es llevado a cabo mediante 4 pasos fundamentales. En primer lugar, se preprocesa la imagen para corregir las diferencias de luminosidad en las distintas zonas de la imagen. A continuación, se han definido una serie de bancos de filtros de Gabor utilizados para la extracción de características de las imágenes. Estas características son utilizadas para crear clasificadores binarios que permitan diferenciar las zonas de las imágenes que contienen grietas. Se buscan dos tipos de grietas, grietas longitudinales y grietas transversales, por lo que se han creado dos tipos de clasificadores, una para cada tipo de grieta. Finalmente, se aplica el algoritmo de entrenamiento AdaBoost, para la

combinación de clasificadores binarios, mejorando el comportamiento de un clasificador aislado.

7.2.1.1 Preprocesamiento

Al comienzo del proceso de adquisición de imágenes, la ganancia y el tiempo de exposición de cada cámara se ajusta individualmente al valor medio entre los dos valores extremos que puede adquirir la cámara, es decir, al valor 128 que es el valor medio entre 0 y 255. De esta manera es posible detectar defectos que supongan tanto un incremento como un decremento en los valores de luminosidad de las imágenes adquiridas. Sin embargo, el valor de luminosidad que se mide a lo largo de una línea no es constante, debido al hecho de que la luz y los ángulos de incidencia de la misma no son idénticos en todos los puntos de dicha línea. Además, la reflectividad de la carretera puede variar al desplazarse por ella y pueden aparecer sombras sobre la superficie, incluso provocadas por el propio vehículo. Para incrementar la homogeneidad de la imagen el nivel de gris de cada imagen es corregido para cada columna de la imagen siguiendo el siguiente proceso:

$$I'_{ij} = 128 \frac{I_{ij}}{A_i^k} \quad i = 1, 2, \dots, M \quad j = 1, 2, \dots, N$$

$$A_i^k = \alpha A_i^{k-1} + (1 - \alpha) a_i^k \quad i = 1, 2, \dots, M$$

Ec. 44

$$\text{donde } \alpha = \beta \frac{k-1}{k}$$

I_{ij} es el valor de luminosidad del pixel de la imagen original de la columna i y de la fila j . I'_{ij} es el valor de luminosidad corregido. M es el número de columnas, N es el número de filas y k es el número de imagen. Además, A_i^k es el valor ponderado de luminosidad de la columna i después de k imágenes y a_i^k es el valor medio de la luminosidad para la columna i de la imagen k . α es una variable de ponderación y β es una constante para controlar el peso de medidas cercanas en el tiempo que ha sido fijada en 0,75. Se puede comprobar que α se aproxima al valor de β cuando el valor de k tiende a infinito. Dos imágenes adquiridas por el LRIS, una sin defectos y otra con defectos, y sus correspondientes imágenes preprocesadas usando el algoritmo desarrollado pueden verse en la siguiente figura.

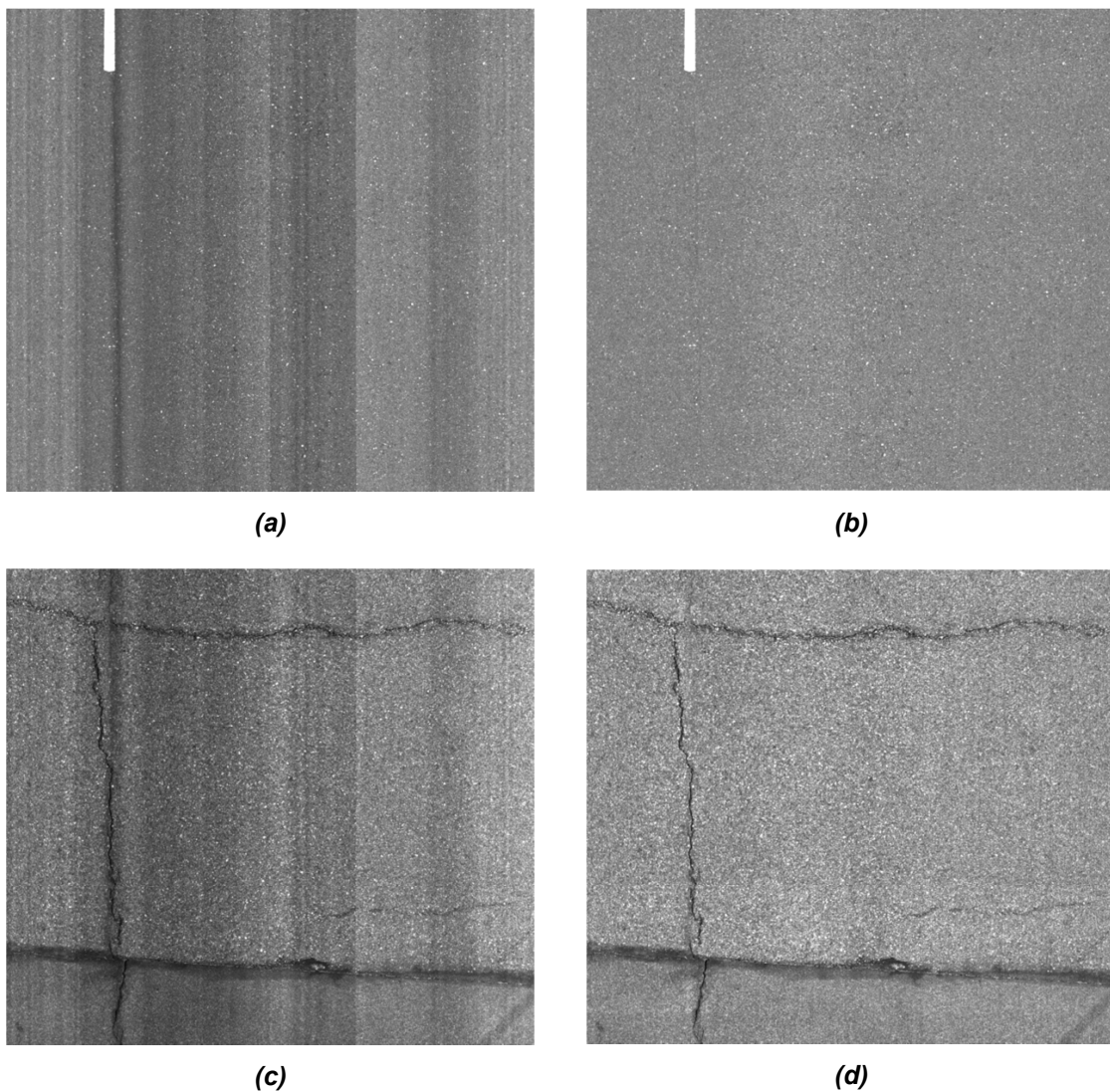


Figura 56. (a) imagen adquirida con el LRIS sin defectos. (b) Imagen preprocesada. (c) Imagen adquirida con el LRIS con presencia de grietas. (d) Imagen preprocesada

7.2.1.2 Extracción de características mediante bancos de filtros de Gabor

En la sección 5, se introdujeron los filtros de Gabor y se expuso cómo se diseñaba un banco de filtros de Gabor, los cuales permiten abarcar todo el espectro en frecuencia y orientación de una forma metódica. Se han aplicado dichas técnicas para diseñar una serie de bancos de filtros de Gabor que permitan la detección y clasificación de grietas longitudinales y transversales presentes en la superficie de la carretera. Las imágenes obtenidas por el equipo LRIS, una vez preprocesadas con el algoritmo presentado en la sección anterior, son filtradas mediante los filtros seleccionados de un conjunto de bancos de filtros de Gabor. De las imágenes filtradas se extraen características que son utilizadas en su clasificación.

Como ya se detalla en la sección 5, el diseño de un banco de filtros de Gabor consiste en la selección de un conjunto de valores para los parámetros del filtro F , θ , σ_x y σ_y . Sustituyendo este conjunto de parámetros en la Ec. 19 se obtienen las distintas ecuaciones del filtro de Gabor.

Siguiendo la metodología expuesta en la sección 5.3 para el diseño de bancos de filtros de Gabor, se ha fijado el número de orientaciones, n_θ , y el ancho de banda de la frecuencia, B_F . Asimismo se han fijado los solapamientos radiales, K_x , y circunferenciales, K_y . Se ha establecido un tamaño de filtro de 128 x 128 píxeles, lo que permite albergar los tamaños de grietas que se encuentran habitualmente en la carretera.

Se ha calculado el ancho de banda de la orientación, B_θ , estableciendo que los filtros están separados uniformemente en la orientación, es decir, dividiendo 180° entre el número de orientaciones, n_θ . Utilizando la Ec. 27, se ha calculado la frecuencia central más alta, F_M , de tal manera que la frecuencia más alta de la iso-curva formada por los puntos cuyo valor del módulo es la mitad del valor máximo alcanzado por el módulo (que para los filtros de Gabor es 1), alcance, pero no sobrepase, el valor de la frecuencia de Nyquist, 0,5.

Se ha fijado el número de frecuencias, n_F , de tal manera que la frecuencia más baja sea superior al inverso del tamaño del filtro, 128^{-1} , ya que las imágenes no pueden contener una frecuencia más baja que el inverso de su tamaño.

Con estos valores se pueden generar todos los filtros del banco de filtros de Gabor, tal y como se explicó en la sección 5.3.

Cuatro bancos de filtros se emplean para la detección de grietas transversales y longitudinales. Los parámetros utilizados para crear estos bancos de filtros de Gabor pueden verse en la **Tabla 14**.

Por tanto, el número total de filtros en los 4 bancos propuestos es 400. El banco 1 consta de 96 filtros (3 x 32), el banco 2 de 96 (6 x 16), el banco 3 de 80 (10 x 8) y el banco 4 de 128 (16 x 8).

Tabla 14. Parámetros de los bancos de filtros de Gabor

	Banco 1	Banco 2	Banco 3	Banco 4
Número de frecuencias (n_F)	3	6	10	16
Número de orientaciones (n_θ)	32	16	8	8
Ancho de banda de la frecuencia (B_F) (octavas)	2	1	0,5	0,25
Ancho de banda de la orientación (B_θ) (°)	5,625	11,25	22,5	22,5
Solapamiento radial (K_x)	1	0,5	0,125	2
Solapamiento circunferencial (K_y)	1	0,5	4	1

Las unidades del ancho de banda de la frecuencia están expresadas en octavas. La segunda frecuencia más alta es igual a la frecuencia más alta dividido entre 2 elevado al ancho de banda de la frecuencia, y así sucesivamente.

Las variables de solapamiento no tienen unidades. Las iso-curvas se solapan cuando las variables de solapamiento son menores de 1 y hay un espacio cuando es mayor de 1.

La representación de dos de los filtros seleccionados puede verse en la **Figura 57**. Puede observarse que existen importantes diferencias entre ambos. El objetivo de estas diferencias es cubrir con estos filtros el mayor rango posible de frecuencias y orientaciones. Debe tenerse en cuenta que el objetivo de los bancos de filtros es que algunos filtros tenga un alto valor de salida cuando se aplican a imágenes con defectos mientras que otros lo tengan cuando la imagen no contenga defectos.

Estos filtros se aplican a porciones de imágenes de 128 x 128 píxeles, por lo que es necesario dividir la imagen en porciones de ese tamaño. Para cada porción de la imagen se evalúa el módulo de la imagen filtrada, asignando como característica el mayor valor del módulo alcanzado dentro de la imagen filtrada.

Se dispone de una amplia base de datos con imágenes etiquetadas como grietas transversales, longitudinales y sin defecto. Las imágenes de las grietas transversales tienen un tamaño de 384 x 128, mientras que el tamaño de las imágenes que contienen grietas longitudinales es de 128 x 384. Se dispone de imágenes de ambos tamaños sin defecto. Todas estas muestras se pueden dividir en tres porciones de 128 x 128 píxeles para aplicarles los filtros de Gabor. El máximo valor del módulo en estas tres porciones es el valor asignado a cada imagen que se utiliza como característica en los clasificadores que se presentan la siguiente sección.

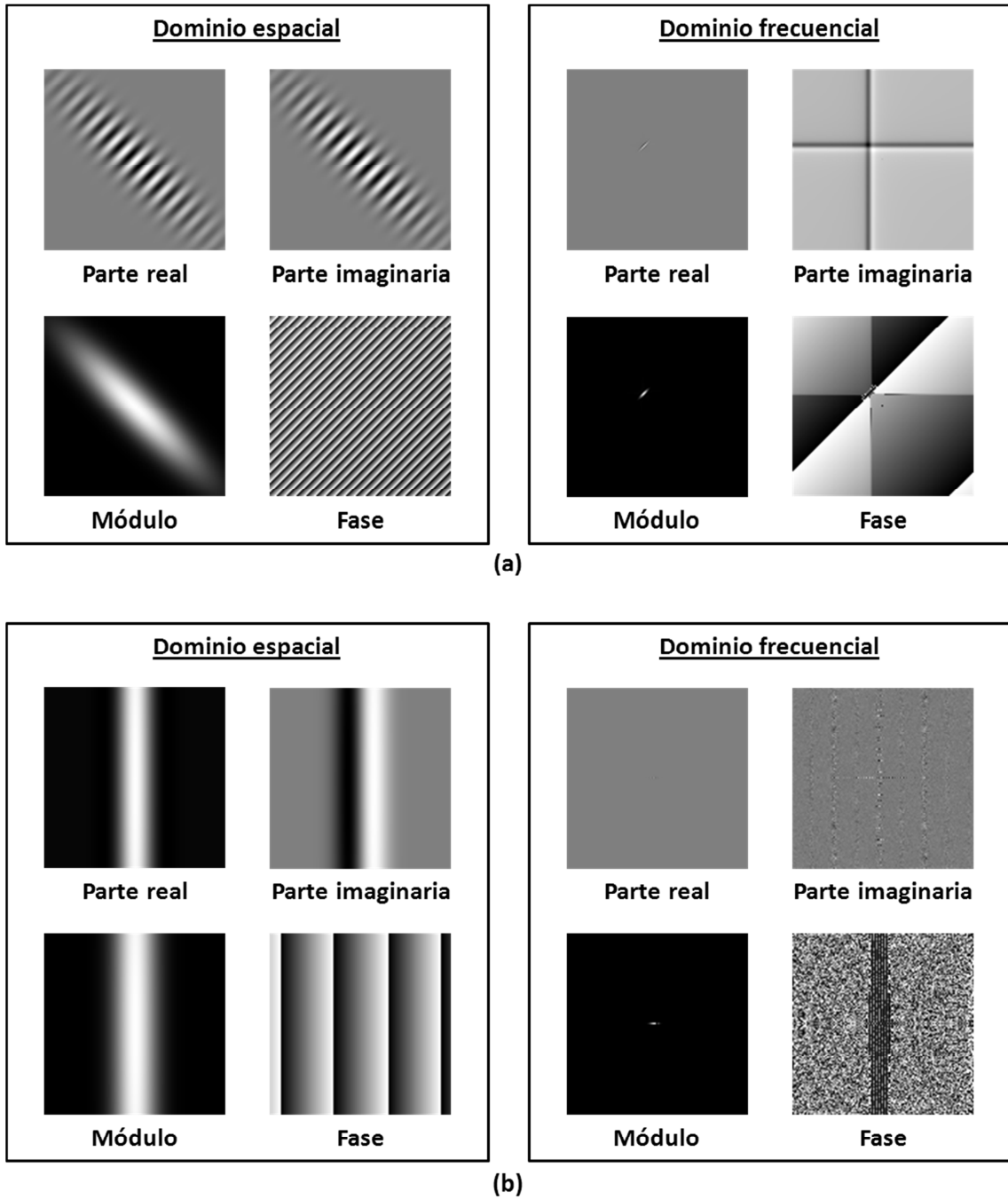


Figura 57. Dos filtros diferentes en el dominio espacial y frecuencial. El rango de valores de la parte real e imaginaria es de -1 a 1, el del módulo de 0 a 1 y la fase de -180° a 180° . (a) Filtro del banco 4. La frecuencia central es $0,010 \text{ píxeles}^{-1}$ y la orientación es 45° . (b) Filtro del banco 3. La frecuencia central es $0,013 \text{ píxeles}^{-1}$ y la orientación es 0° .

7.2.1.3 Clasificadores binarios

Para cada tipo de grietas, transversales y longitudinales, se emplean tres tipos de clasificadores binarios, diferenciados por el método de cálculo del umbral. Estos tres métodos se describen con más detalle en la sección 4.2.1.

- **Método 1:** El umbral es el valor medio del valor medio de las muestras positivas y del valor medio de las muestras negativas, como se puede ver en la Ec. 4.
- **Método 2:** El umbral se fija minimizando la distancia con el clasificador ideal.
- **Método 3:** El umbral se fija maximizando la suma de la sensibilidad y la especificidad, que equivale a maximizar la exactitud ponderada.

El valor utilizado para cada una de las muestras es el descrito en la sección anterior, es decir, para cada clasificador se utiliza la imagen filtrada de uno de los 400 filtros de Gabor descritos en el apartado anterior y se calcula el mayor valor del módulo de la imagen filtrada. El umbral para determinar si una muestra pertenece al grupo de grietas o al grupo de no defecto se determina empleando los tres métodos, de tal manera que se puedan comparar los resultados obtenidos.

7.2.1.4 Entrenamiento mediante el algoritmo AdaBoost

Se ha utilizado el algoritmo Adaboost para realizar el entrenamiento de un clasificador fuerte a partir de los clasificadores binarios, llamados también clasificadores débiles, tal y como se describe en la sección 4.4.

Para cada tipo de defectos y para cada método de clasificación se aplica este entrenamiento y se comparan los resultados obtenidos. En la siguiente sección se muestran los resultados obtenidos.

7.2.2 Resultados obtenidos

Del conjunto de imágenes adquiridas y preprocesadas con el equipo LRIS se han seleccionado un conjunto de imágenes que contienen defectos de grietas longitudinales y transversales, así como otras que estaban libres de defectos. Se pueden ver algunas muestras de grietas transversales en la **Figura 58** y de grietas longitudinales en la **Figura 59**. Las imágenes seleccionadas han sido utilizadas para diseñar un clasificador de grietas transversales y un clasificador de grietas longitudinales. En el primero de los casos las imágenes tienen un tamaño de 384 x 128 mientras que en el segundo son imágenes de tamaño 128 x 384, tanto para las imágenes que contienen defectos como para las que no los contienen.

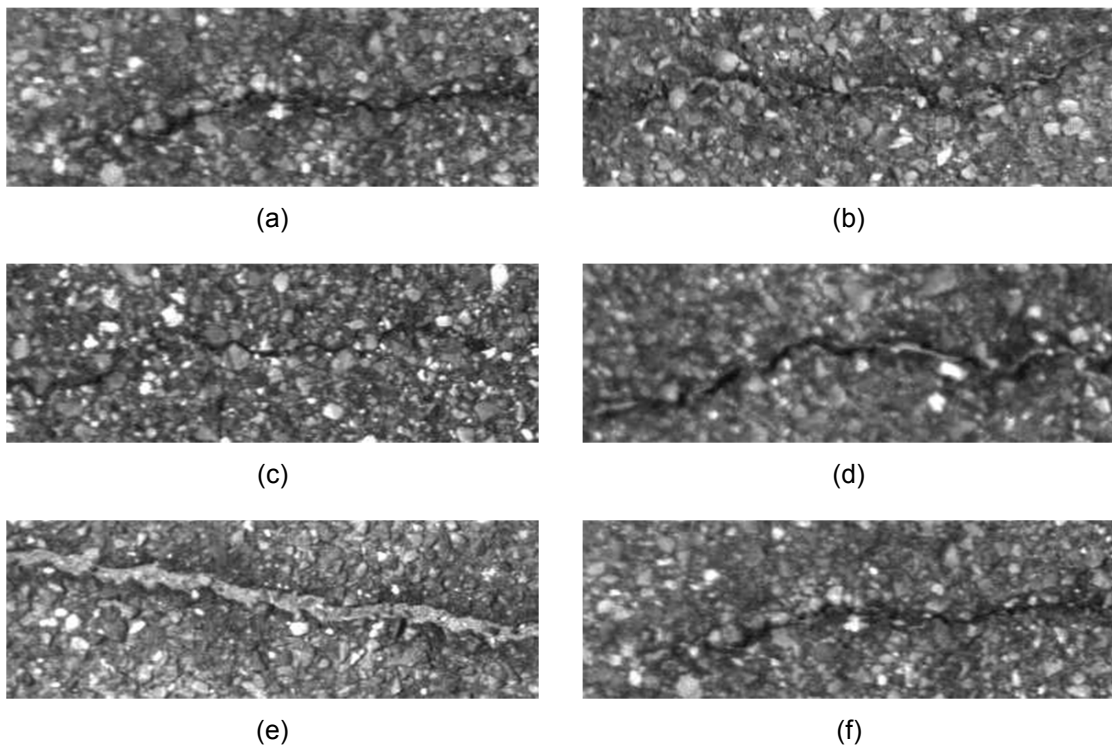


Figura 58. Imágenes de grietas transversales utilizadas en el entrenamiento

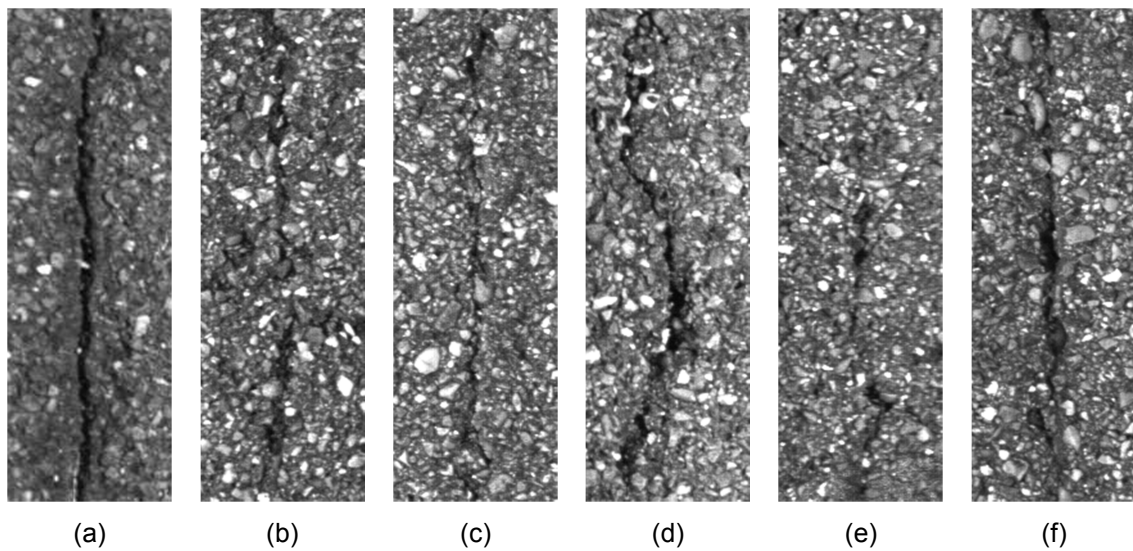


Figura 59. Imágenes de grietas longitudinales utilizadas en el entrenamiento

En las imágenes anteriores se puede ver una pequeña muestra de las imágenes utilizadas para entrenar al sistema. Las imágenes ponen de manifiesto la dificultad existente para la detección de estas grietas así como la gran variabilidad que existe en

la apariencia visual de grietas del mismo tipo, lo cual dificulta la automatización de cualquier procesamiento automático de imágenes.

El conjunto de imágenes utilizadas para realizar el clasificador de grietas longitudinales está compuesto por un total de 2.922 imágenes de las cuales 239 contienen una grieta longitudinal y 2.683 están libres de ellas. Para el clasificador de grietas transversales se dispone de un total de 6.777 imágenes, de las cuales 567 contienen grietas transversales y 6.210 están libres de defectos.

Tanto el conjunto de imágenes de grietas longitudinales como el conjunto de imágenes de grietas transversales, se han dividido en dos subgrupos, uno para entrenar el sistema y otro para testarlo. Las imágenes que pertenecen a cada subgrupo han sido seleccionadas al azar, asignando el 80% al primer grupo y el 20% al segundo. La cantidad de imágenes que contienen cada grupo se encuentra en la **Tabla 15**:

Tabla 15. Muestras utilizadas en el entrenamiento y el test para la detección de grietas transversales y longitudinales

		Longitudinal	Transversal
Entrenamiento	Positivos	191	454
	Negativos	2.146	4.968
	Total	2.337	5.422
Test	Positivos	48	113
	Negativos	537	1.242
	Total	585	1.355
Total	Positivos	239	567
	Negativos	2.683	6.210
	Total	2.922	6.777

Los filtros que mejor se han comportado para cada uno de los tres métodos estudiados para fijar el umbral se muestran a continuación en la **Tabla 16**.

Tabla 16. *Filtros individuales que han alcanzado los menores errores para los tres métodos utilizados para grietas transversales*

	Método 1	Método 2	Método 3
Banco	2	3	3
Frecuencia (píxeles ⁻¹)	0,009375	0,009314	0,009314
Orientación (°)	90	90	90
Dispersión en X (píxeles)	29,98	14,66	14,66
Dispersión en Y (píxeles)	101,47	404,61	404,61
Umbral	2,72	0,9457	0,9457
Signo	1	1	1
Error	0,1143	0,1045	0,1045

Signo indica la dirección de la inecuación: -1 indica que el defecto tiene una respuesta de mayor valor que la imagen sin defecto.

El error es el error ponderado, que equivale a $1 - \text{media}$ entre la sensibilidad y la especificidad.

Tabla 17. *Filtros individuales que han alcanzado los menores errores para los tres métodos utilizados para grietas longitudinales*

	Método 1	Método 2	Método 3
Banco	3	3	3
Frecuencia (píxeles ⁻¹)	0,013171	0,013171	0,013171
Orientación (°)	0	0	0
Dispersión en X (píxeles)	10,37	10,37	10,37
Dispersión en Y (píxeles)	286,10	286,10	286,10
Umbral	1,5970	1,4384	1,4384
Signo	-1	-1	-1
Error	0,1387	0,1230	0,1230

Signo indica la dirección de la inecuación: -1 indica que el defecto tiene una respuesta de mayor valor que la imagen sin defecto.

El error es el error ponderado, que equivale a $1 - \text{media}$ entre la sensibilidad y la especificidad.

7. Resultados

Se puede ver que los clasificadores individuales alcanzan errores que oscilan entre 0,1045 y el 0,1387, lo cual implica unos buenos resultados a priori. Se puede observar en las dos tablas anteriores que 5 de los 6 filtros seleccionados pertenecen al banco de filtros número 3, mientras que el otro (el seleccionado para grietas transversales por el método 1) pertenece al banco de filtros número 2. La orientación seleccionada en el caso de la grietas transversales es de 90° mientras que en el caso de grietas longitudinales es de 0°, tal y como cabía esperar debido a la orientación media de este tipo de defectos. Los resultados obtenidos con los métodos 2 y 3 son exactamente iguales para el mejor clasificador, mientras que el método 1 presenta resultados ligeramente peores ya que su error es algo mayor.

Los resultados anteriormente mostrados han sido obtenidos con las muestras utilizadas para realizar el entrenamiento. Se van a comparar los resultados de comportamiento de estos filtros para las muestras utilizadas en el entrenamiento con las muestras que se han reservado para la realización del test. Para comparar los resultados se han calculado para cada filtro seleccionado la sensibilidad, especificidad, exactitud y la precisión, así como la exactitud ponderada, que fueron definidas en la sección 4.2. En las siguientes tablas se muestran los resultados comparativos para las muestras utilizadas en el entrenamiento y en el test, cuando se utiliza un solo clasificador.

Tabla 18. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas transversales por el método 1

	Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
Entrenamiento	0,8238	0,9477	0,9373	0,9833	0,8857
Test	0,8407	0,9372	0,9292	0,9848	0,8890

Tabla 19. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas transversales por el método 2 y 3

	Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
Entrenamiento	0,8877	0,9034	0,9021	0,9888	0,8955
Test	0,9204	0,8977	0,8996	0,9920	0,9091

Tabla 20. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas longitudinales por el método 1

	Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
Entrenamiento	0,7906	0,9320	0,9204	0,9804	0,8613
Test	0,7292	0,9385	0,9214	0,9749	0,8339

Tabla 21. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del filtro seleccionado para detectar grietas longitudinales por el método 2 y 3

	Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
Entrenamiento	0,8639	0,8900	0,8879	0,9866	0,8770
Test	0,8958	0,8734	0,8752	0,9895	0,8846

Se puede concluir que los resultados obtenidos con las muestras de entrenamiento y las muestras para el test son muy similares. En algunos de los casos se obtienen incluso mejores resultados en las muestras de los test que las propias muestras utilizadas para el entrenamiento aunque estas pequeñas variaciones son propias de la evaluación de muestras seleccionadas de forma aleatoria.

Al comprobar que las muestras utilizadas en el test proporcionan resultados similares se demuestra que el sistema de entrenamiento es válido y que el clasificador generado no está sobreentrenado (un clasificador sobreentrenado se ajusta extremadamente bien a las muestras usadas para entrenar a costa de perder generalidad y por tanto no sería válido para futuras muestras).

También se puede comprobar que tanto el método 2 como el 3 ofrecen mejores índices de exactitud ponderada que el método 1, tanto para grietas transversales como para grietas longitudinales. Se puede observar como el método 1 obtiene mejores valores de especificidad pero mucho peores de sensibilidad, es decir, clasifica mejor las muestras que no son defecto pero peor las que sí lo son. Como el grupo de muestras utilizadas contienen una menor cantidad de imágenes con grietas que de imágenes sin defecto, los valores de exactitud son más altos para el método 1 que para los métodos 2 y 3. Esto es debido, como ya se explicó anteriormente, a que la medida de la exactitud tiene en cuenta por igual todas las muestras, mientras que la exactitud

7. Resultados

ponderada es la media aritmética del acierto en las muestras positivas (sensibilidad) y del acierto en las muestras negativas (especificidad).

Aunque los resultados de los filtros individuales muestran unos buenos resultados, se ha intentado mejorar aún más estos resultados. Para ello se ha utilizado el método de entrenamiento de Adaboost, que busca la forma de mejorar los resultados obtenidos por un clasificador único, combinando diversos clasificadores de la forma más apropiada.

En las siguientes figuras se representa la evolución del error ponderado al ir aumentando el número de clasificadores utilizados para cada uno de los métodos de cálculo de umbral que se han propuesto y para cada tipo de grieta.

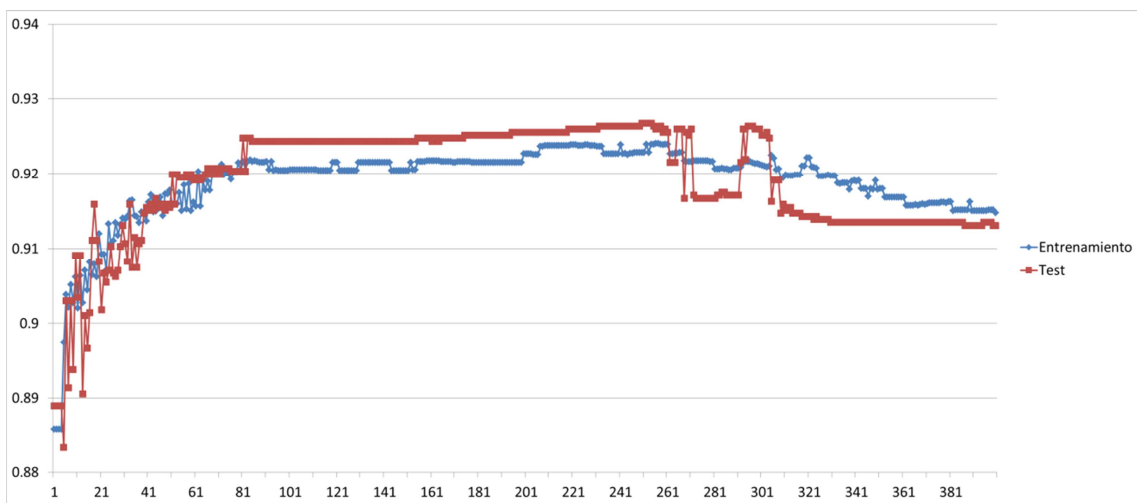


Figura 60. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 1 para grietas transversales

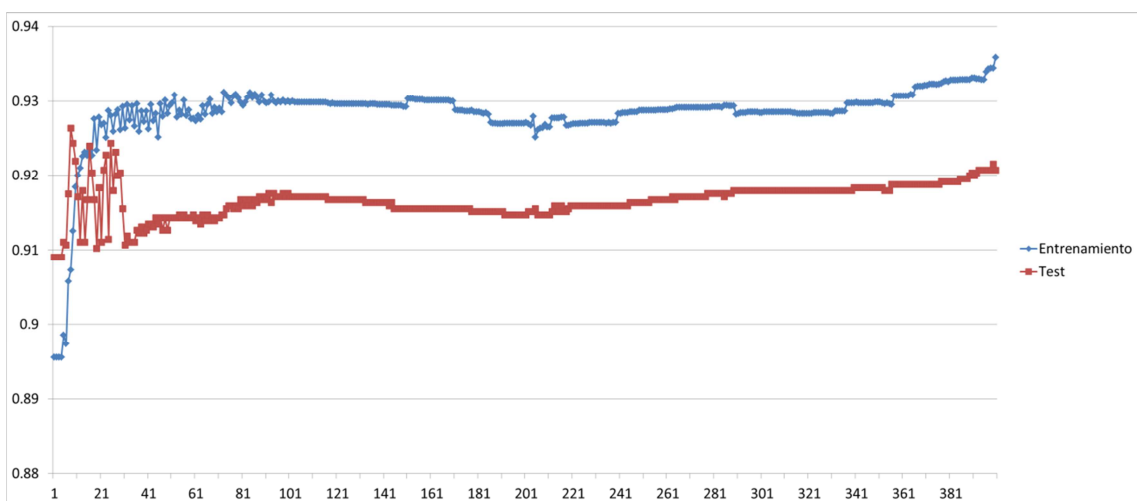


Figura 61. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 2 para grietas transversales

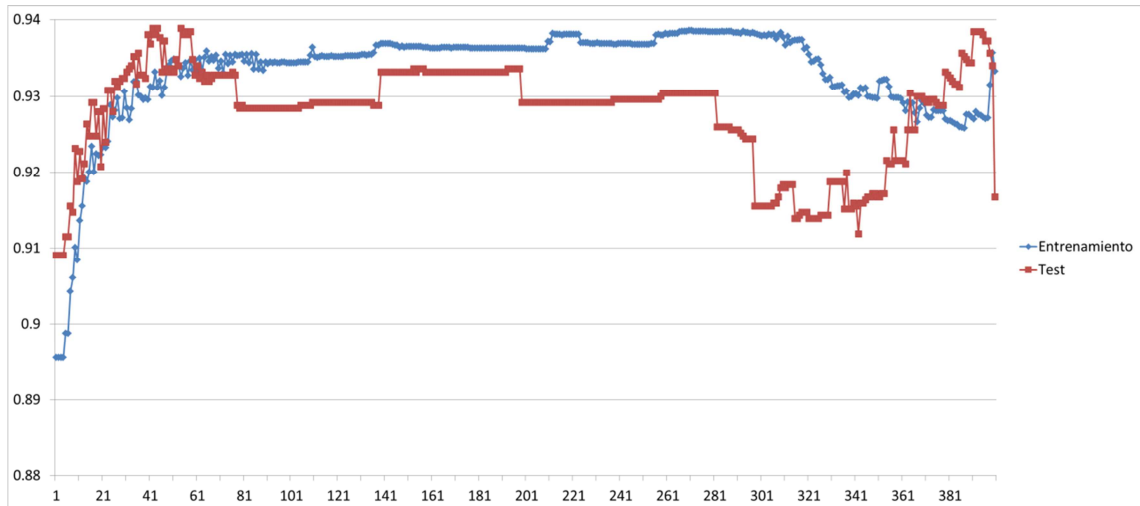


Figura 62. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 3 para grietas transversales

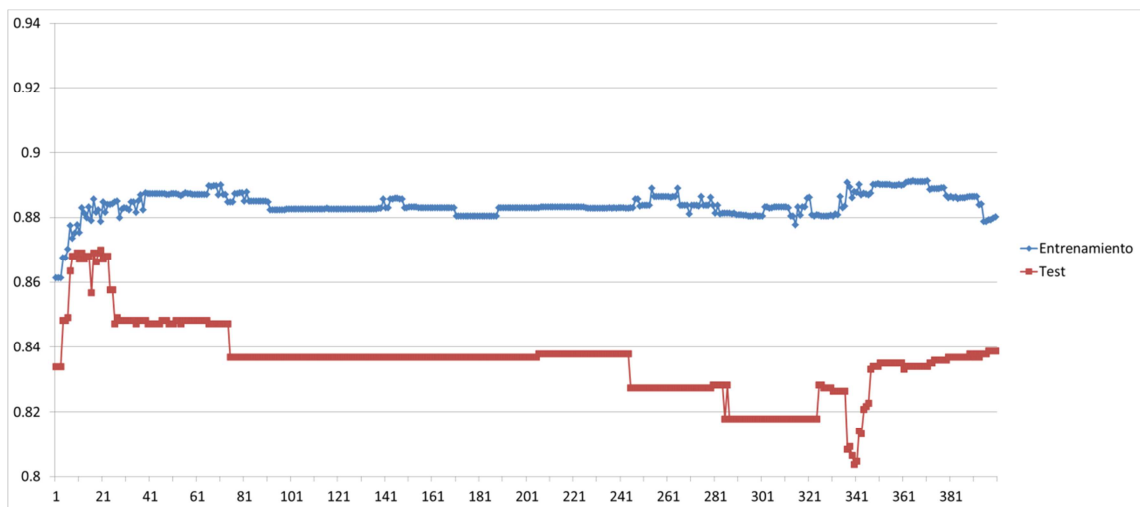


Figura 63. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 1 para grietas longitudinales

7. Resultados

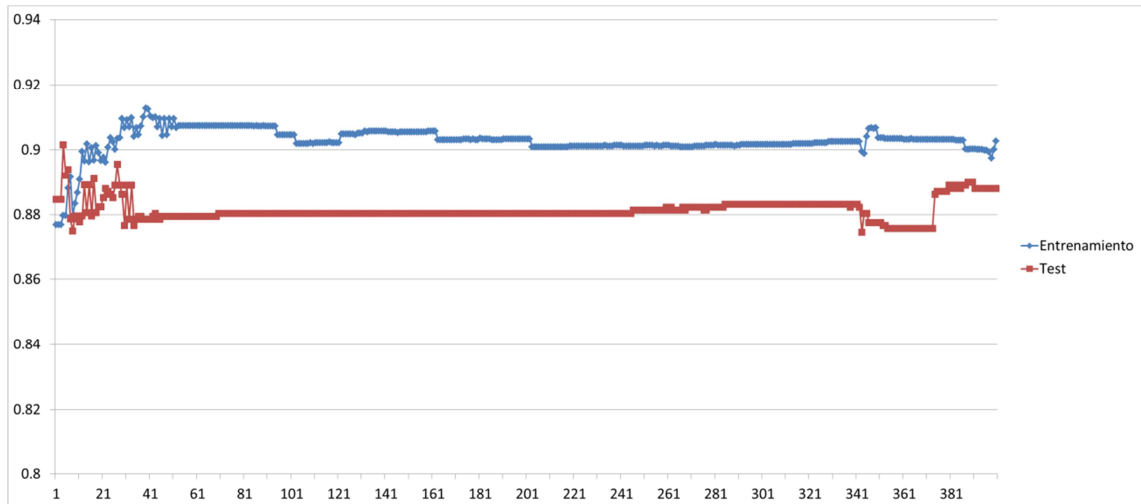


Figura 64. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 2 para grietas longitudinales

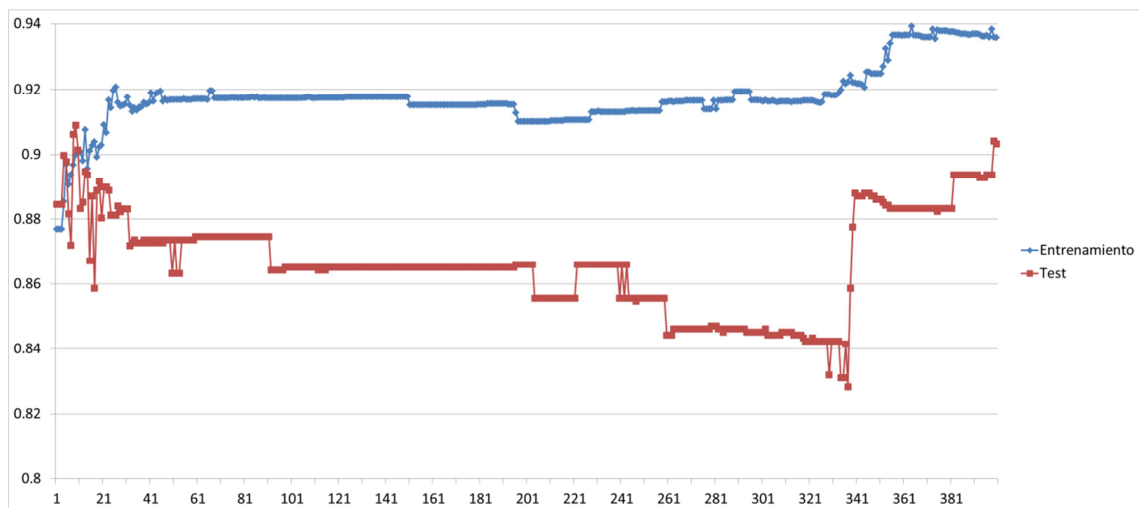


Figura 65. Evolución de la exactitud ponderada al aumentar el número de clasificadores generado por el método 3 para grietas longitudinales

Las gráficas anteriores ofrecen información del comportamiento del clasificador fuerte resultante de la combinación de varios clasificadores mediante el algoritmo de AdaBoost. Se ha podido extraer información de cuántos clasificadores se deben utilizar, de cuál es el índice de acierto que se puede alcanzar y si este índice de acierto es similar para las muestras de entrenamiento y para las muestras test. A continuación se van a analizar cada una de estas gráficas, así como los resultados y conclusiones extraídas.

1. Grietas transversales:

- a. **Método 1:** En la **Figura 60** se puede ver que utilizando un único clasificador se obtiene un índice de acierto inicial aproximado de 0,885, tanto para las muestras de entrenamiento como para las muestras test. Según se va incrementando el número de clasificadores se incrementa este índice de acierto de forma similar para los dos grupos de muestras hasta llegar a estabilizarse aproximadamente al llegar a los 80 clasificadores, donde se alcanzan valores por encima de 0,92, siendo ligeramente superior para las muestras test, lo cual muestra claramente la no existencia de sobreentrenamiento.
- b. **Método 2:** Se puede observar en la **Figura 61** que existe una discrepancia en los valores de inicio entre las muestras de entrenamiento y las muestras test, en el que las muestras test presentan un comportamiento mejor que las de entrenamiento, siendo el primero de los casos de 0,895 y 0,909 en el caso de las muestras test. Las muestras de entrenamiento presenta una evolución positiva, muy similar al caso anterior hasta estabilizarse en torno al 0,93 al llegar a utilizar unos 25 clasificadores. Sin embargo, el comportamiento de las muestras test difiere del de las muestras de entrenamiento. Se puede ver que se incrementa con los primeros clasificadores, alcanzando rápidamente un máximo cuando se utilizan 8 clasificadores, disminuyendo a continuación y estabilizándose para una mayor cantidad de clasificadores. El comportamiento para este entorno de número de clasificadores es similar al de las muestras de entrenamiento, siendo significativamente peor a partir de los 30 clasificadores, lo cual muestra que el clasificador fuerte, a partir de este número de filtros, sufre el problema de sobreentrenamiento y no sería apropiada su utilización.
- c. **Método 3:** El punto de partida es el mismo que el método 2, anteriormente explicado, como se puede comprobar en la **Figura 62**. Sin embargo, en este caso el comportamiento de ambos grupos de muestras es similar hasta llegar aproximadamente a los 50 clasificadores, mejorando su comportamiento al incrementarse el número de clasificadores. La convergencia en los valores de las dos muestras es muy rápida, manteniéndose el mejor comportamiento de las muestras test respecto a las usadas para el entrenamiento hasta los 50 clasificadores. El grupo de muestras utilizado para el entrenamiento alcanza un valor máximo de exactitud ponderada de 0,935, mientras que las muestras test rozan los 0,94, siendo los mejores valores de los tres métodos estudiados.

2. Grietas longitudinales

- a. **Método 1:** En este caso las muestras de entrenamiento presentan mejores resultados de inicio que las muestras test, siendo los índices de

exactitud ponderada de 0,861 y 0,834 respectivamente, tal y como aparecen en la **Figura 63**. Ambos grupos presentan al principio una mejora al incrementar el número de clasificadores utilizado, convergiendo rápidamente sus valores. Con 8 clasificadores se obtienen unos índices de exactitud ponderada de 0,877 y 0,868. A partir de este punto, al incrementar el número de clasificadores no mejora el comportamiento para las muestras test, aunque sí que lo hace para las muestras de entrenamiento, lo que pone de manifiesto que a partir de este número de clasificadores se está produciendo el efecto del sobreentrenamiento, y no se deberían usar más filtros.

- b. Método 2:** En este método se parte de unos valores iniciales de 0,877 y 0,885 para los grupos de entrenamiento y test. Como se puede observar en la **Figura 64** es superior para el grupo de muestras test. El grupo de entrenamiento mejora su comportamiento hasta los 40 clasificadores, alcanzado en ese punto una exactitud ponderada de 0,912. Sin embargo, el grupo de muestras test alcanza su valor máximo al cabo de tan solo 4 clasificadores siendo el valor de exactitud ponderada de 0,901, disminuyendo a partir de ese valor y no mejorando posteriormente.
- c. Método 3:** Como se muestra en la **Figura 65**, el punto de partida para este método es el mismo que en el método 2. El grupo de muestras de entrenamiento alcanza un máximo para 25 clasificadores, con un índice de exactitud ponderada de 0,92, mientras que las muestras test alcanzan su máximo para 9 clasificadores, siendo su índice de exactitud ponderada próximo a 0,91.

Se puede concluir que el método 2 es el que muestra una peor evolución al aumentar el número de clasificadores. Los métodos 1 y 3 presentan una mejor evolución, y como los resultados para un solo clasificador son mejores para el método 3, el clasificador fuerte final obtenido también presenta mejores resultados.

En general, se puede observar que, para todos los métodos, a partir de cierto número de clasificadores el clasificador fuerte no mejora significativamente para las muestras de entrenamiento, incluso llegando a empeorar para las muestras test. Esto es debido a dos motivos, en primer lugar, a que algunos de los clasificadores ofrecen una información redundante, y en segundo lugar, a que al aumentar el número de clasificadores el clasificador fuerte se "amolda" excesivamente a los datos usados para entrenar, de tal manera que el clasificador pierde generalidad.

El número de clasificadores recomendado para alguno de los métodos puede llegar a superar los 50. Esto supone un incremento significativo del tiempo necesario para el procesamiento de las imágenes. El procesamiento de las imágenes se lleva a cabo fuera de línea por lo que el tiempo de procesamiento no se ha considerado una

variable crítica, pero podría tenerse en cuenta a la hora de obtener los resultados en la mayor brevedad de tiempo posible.

También se puede observar que el clasificador fuerte para grietas longitudinales no ha funcionado tan bien como el clasificador fuerte para grietas transversales, especialmente para la base de datos usada para el test. Esto puede ser debido al buen comportamiento del mejor clasificador débil y porque el número de muestras no es lo suficientemente alto.

Por tanto, se ha considerado que tanto para las grietas transversales como para las longitudinales el método 3 es el que presenta los mejores resultados y es el que se está utilizando para analizar las imágenes. En el caso de las grietas transversales el clasificador fuerte consta de 56 clasificadores mientras que en el caso de las grietas longitudinales consta de 9 clasificadores. Los datos de comportamiento de estos filtros se pueden ver en las dos siguientes tablas.

Tabla 22. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del clasificador fuerte seleccionado para detectar grietas transversales

	Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
Entrenamiento	0,9207	0,9481	0,9458	0,9924	0,9344
Test	0,9292	0,9477	0,9461	0,9932	0,9384

Tabla 23. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada de las muestras utilizadas en el entrenamiento y en el test del clasificador fuerte seleccionado para detectar grietas longitudinales

	Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
Entrenamiento	0,8900	0,9096	0,9080	0,9894	0,8998
Test	0,9167	0,9013	0,9026	0,9918	0,9090

Comparando la **Tabla 19** con la **Tabla 22** y la **Tabla 21** con la **Tabla 23**, es decir, comparando la utilización de un solo clasificador frente a la utilización del clasificador

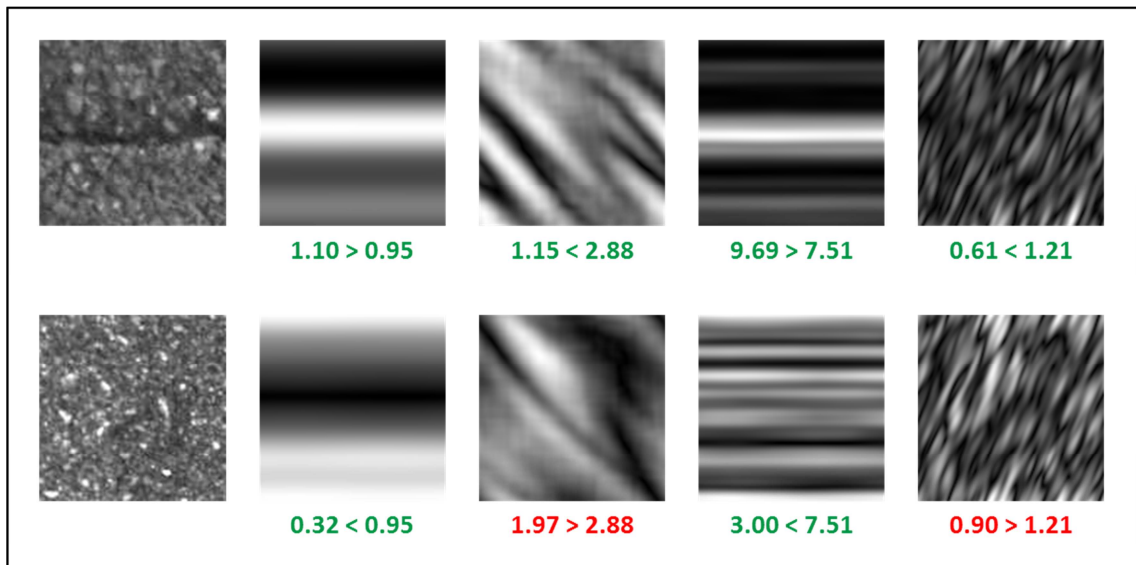
fuerte, se puede comprobar que se han incrementado significativamente el rendimiento de los clasificadores, tanto para las muestras de entrenamiento como para el grupo de muestras utilizado para el test.

En la **Figura 66** se muestra una grieta transversal y una grieta longitudinal, así como dos imágenes que no tienen defectos. La grieta transversal y una de las imágenes sin defecto han sido filtradas por los cuatro primeros filtros de Gabor seleccionados por el clasificador fuerte de grietas transversales, mientras que la grieta longitudinal y la otra imagen sin defectos han sido filtradas por los cuatro primeros filtros de Gabor seleccionados por el clasificador fuerte de grietas longitudinales. Debajo de cada imagen se encuentra una inequación con el máximo valor del módulo de la imagen filtrada y el umbral determinado por el método 3, tal y como se explica en la sección 4.2.1. Las inequaciones que se cumplen están marcadas en verde, mientras que las que no se cumplen se marcan en rojo. Se puede observar que no se cumplen todas las condiciones, pero para que un clasificador fuerte clasifique una muestra correctamente no es necesario que cumpla todas las condiciones, tal y como se explica en la sección 4.4.

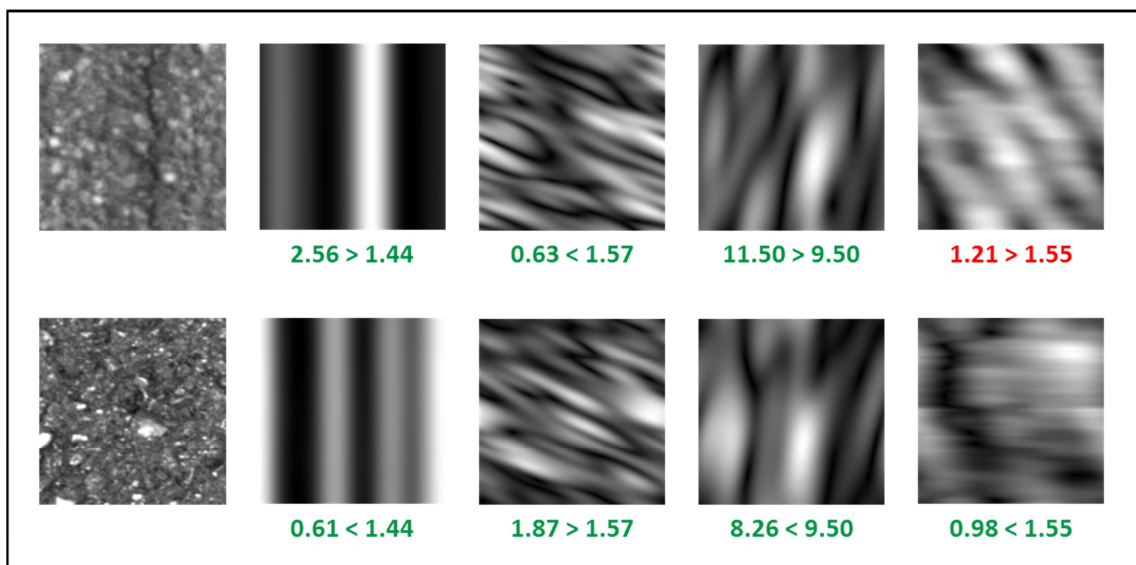
7.2.3 Conclusiones

Se ha abordado la detección y clasificación de grietas en carreteras analizando las imágenes adquiridas por el sistema de adquisición de imágenes LRIS. Para ello se ha adaptado el algoritmo AdaBoost para combinar el análisis de las imágenes filtradas por un extenso conjunto de filtros de Gabor. El método desarrollado ha sido aplicado y verificado con un gran número de imágenes de grietas longitudinales y transversales obtenidos en carreteras en condiciones de iluminación no controlada. Los resultados han sido satisfactorios, alcanzando los distintos métodos desarrollados una exactitud ponderada en torno al 0,9.

La comparación de resultados con otros trabajos es siempre una tarea difícil debido a diferentes motivos: Los objetivos no son exactamente los mismos, la forma de medir el comportamiento de los clasificadores no es exactamente el mismo, las bases de datos no suelen ser públicas y la implementación de los algoritmos suele ser compleja.



(a)



(b)

Figura 66. De izquierda a derecha, imagen preprocesada y el valor del módulo de la imagen filtrada por los cuatro mejores filtros de Gabor seleccionados por cada clasificador fuerte. Debajo de cada imagen, la inecuación que debe cumplir cada clasificador; las condiciones que son satisfechas están escritas en verde, mientras que las que no están en rojo. (a) Grietas transversales. (b) Grietas longitudinales

7. Resultados

Como puede verse en la **Tabla 24**, dos trabajos recientes han sido tenidos en cuenta para comparar sus resultados con los obtenidos por los métodos descritos en esta tesis. El método de Oliveria y Correia [Oli09a] y el método de Cord y Chambon [Cor12], que serán llamados a partir de ahora método de Oliveira y método de Cord respectivamente.

Tabla 24. Comparativa de distintos clasificadores

	Sensibilidad	Especificidad	Exactitud ponderada
1. Método de Oliveira aplicado a BD1 de Oliveira	0,9480	No disponible	No disponible
2. Método de Oliveira aplicado a BD2 de Oliveira	0,9560	No disponible	No disponible
3. Método de Cord aplicado a la BD de Cord	0,9520	0,9230	0,9375
4. Clasificador débil con el método 1 para grietas transversales	0,8407	0,9372	0,8910
5. Clasificador débil con los métodos 2 y 3 para grietas transversales	0,9204	0,8977	0,9091
6. Clasificador débil con el método 1 para grietas longitudinales	0,7292	0,9385	0,8339
7. Clasificador débil con los métodos 2 y 3 para grietas longitudinales	0,8958	0,8734	0,8846
8. Clasificador fuerte con el método 3 para grietas transversales	0,9292	0,9477	0,9384
9. Clasificador fuerte con el método 3 para grietas longitudinales	0,9167	0,9013	0,9090

El método y las bases de datos BD1 y BD2 de Oliveira se encuentran en [Oli09].

El método y la base de datos de Cord se encuentran en [Cor12]

Los clasificadores propuestos en esta tesis (del 4 al 9) se aplican a bases de datos propias.

El método de Oliveira se basa en la medida de la entropía y es aplicado a sus propias bases de datos, llamadas por los autores BD1 y BD2. BD1 es una base de datos usada para verificar su método, compuesta por 20 imágenes de 4.096 x 2.048 píxeles obtenidas por un sistema LRIS. BD2 es otra base de datos que también utilizan para verificar su método que está compuesta por 56 imágenes de 2.048 x 1.536 píxeles obtenidas con una cámara digital con su eje óptico perpendicular a la superficie de la carretera. En ambos casos, el ancho total de un carril de la carretera, de

aproximadamente 4 metros, es adquirido, por lo que la resolución de las imágenes es de 1 x 1 mm y 2 x 2 mm por píxel respectivamente.

El método de Cord aplica el algoritmo de AdaBoost a descriptores de textura. Este método minimiza el error ponderado (equivalente a maximizar la exactitud ponderada) aplicándolo a su propia base de datos, la cual está compuesta de 6.875 sub-imágenes de 128 x 128 píxeles, que contienen cinco tipos diferentes de defectos. Estas sub-imágenes han sido extraídas de imágenes de 1.920 x 480 píxeles de un carril completo de la carretera de aproximadamente 4 metros de ancho, por lo que su resolución es de 2 x 2 mm por píxel.

Los métodos propuestos en esta tesis han sido verificados con bases de datos de grietas transversales y grietas longitudinales que contienen imágenes adquiridas con los equipos descritos en la sección 6.3. La base de datos usada para verificar los clasificadores de grietas transversales está formada por 585 imágenes de 384 x 128 píxeles, mientras que la base de datos usada para verificar los clasificadores de grietas longitudinales está formada por 1.355 imágenes de 128 x 384 píxeles. La resolución de estas imágenes es de 1 x 1 mm por píxel.

El método de Oliveira divide las imágenes en sub-imágenes de 75 x 75 píxeles, mientras que el método de Cord lo hace en 128 x 128 píxeles. Ambos métodos clasifican las sub-imágenes en imágenes con defecto e imágenes sin defecto. Sin embargo, el método propuesto en esta tesis determina si una imagen contiene un defecto concreto o no lo contiene.

Se ha seleccionado la sensibilidad, la especificidad y la exactitud ponderada para comparar su comportamiento. No hay un método convencional para medir el comportamiento de un clasificador porque, dependiendo del objetivo buscado, una u otra medida es más útil. Además, se utilizan diferentes nombres para llamar a la misma medida y los valores complementarios de éstas suelen utilizarse frecuentemente. Se ha extraído la información disponible de los artículos de referencia para comparar los resultados. El método de Oliveira no proporciona la especificidad ni la exactitud ponderada. En su lugar, proporcionan la precisión y el valor-F (*F-measure*).

Puede observarse en la **Tabla 24** que nuestros métodos de clasificación alcanzan buenos resultados tanto para los clasificadores débiles como para los clasificadores fuertes. Además, los clasificadores fuertes mejoran los resultados de los clasificadores débiles, tal y como era de esperar.

El clasificador fuerte obtenido al aplicar el método de Cord alcanza una exactitud ponderada de 0,9375. Este método utiliza una estrategia de clasificación similar a nuestros clasificadores fuertes, pero el método que se propone en esta tesis necesita

más clasificadores débiles y los filtros utilizados son más complejos, por lo que necesita más tiempo de computación. Utilizando solo un clasificador débil, se obtiene una exactitud ponderada de 0,9091 y 0,8846 para grietas transversales y longitudinales respectivamente, cercano al valor de 0,9375 obtenido con el método de Cord, pero teniendo que procesar solo un filtro de Gabor, con lo que el tiempo de procesamiento es mucho menor que el utilizado en calcular todas las características propuestas en el trabajo de Cord. Más aún, el clasificador fuerte propuesto en este trabajo alcanza una exactitud ponderada de 0,9384 para grietas transversales y 0,9090 para grietas longitudinales. El primer valor es ligeramente superior al alcanzado por el método de Cord, mientras que el segundo es un poco menor. Se debe tener en cuenta que la base de datos utilizada por Cord contiene grietas de cocodrilo (17,5%), sin reparar (38,3%) y reparadas (44,2%), mientras que la base de datos utilizada en esta tesis solo contiene grietas sin reparar, que son sin duda las más difíciles de detectar. De todas formas, el comportamiento de nuestro método es similar al método de Cord pero utilizando solo un tipo de filtros y siendo capaz de detectar grietas de la mitad de abertura, ya que la resolución de nuestras imágenes es el doble. Además, el método que se ha propuesto en esta tesis, no solo detecta la presencia de grietas, si no que las clasifica en dos grupos en un solo paso, grietas transversales y grietas longitudinales, lo cual supone una mejora para un sistema de gestión de carreteras.

7.3 Detección y clasificación de grietas en la superficie de túneles

Se ha desarrollado un prototipo para la inspección de grietas en la superficie de túneles. El prototipo desarrollado tiene por objetivo determinar la viabilidad de la detección de grietas en la superficie de túneles recubiertos de hormigón. El sistema propuesto podría aplicarse a cualquier tipología de túnel de hormigón (ferroviario, carretero, hidráulico...), habiéndose evaluado el prototipo capturando imágenes en dos tipos de túneles: Uno ferroviario y otro carretero. Estas imágenes han sido utilizadas para el desarrollo de diversas técnicas de procesado de imágenes. En esta sección se van a mostrar los resultados obtenidos al aplicar a estas imágenes algunas de las técnicas descritas a lo largo de este proyecto.

Al igual que en el procesamiento de imágenes de chapa laminada o carreteras, las imágenes son inicialmente preprocesadas con el fin de homogeneizar la distribución de la iluminación. Las imágenes preprocesadas han sido utilizadas para aplicar los algoritmos descritos a continuación.

Para entrenar y comprobar el funcionamiento de los algoritmos propuestos, se han segmentado manualmente un conjunto de imágenes que contienen grietas, clasificando los píxeles de las imágenes como píxeles con grietas o píxeles sin defectos.

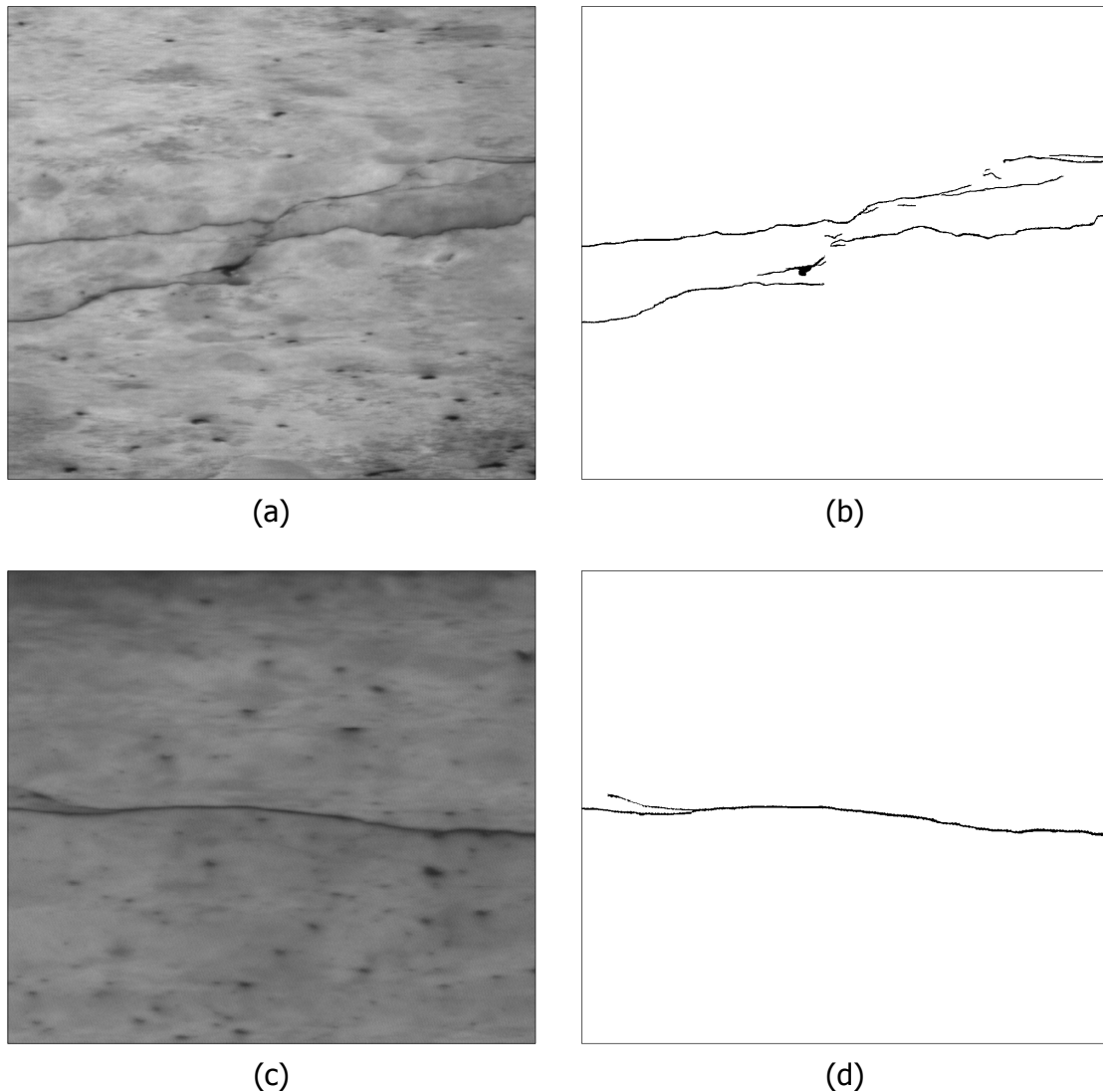


Figura 67. (a) y (c) son imágenes preprocesadas de grietas en túneles. (b) y (d) son la segmentación manual de estas grietas

Se han seleccionado aleatoriamente un total de 100.000 píxeles, 20.000 de los cuales son píxeles con grietas. Estos píxeles han sido utilizados para ajustar los parámetros de los algoritmos propuestos y verificar el funcionamiento del sistema.

En esta sección se muestran los resultados obtenidos al aplicar distintas técnicas para la segmentación de grietas al mismo grupo de imágenes, de tal forma que se puedan comparar directamente los resultados obtenidos.

7.3.1 Filtro de Gabor invariante a la rotación

En primer lugar, se ha utilizado un filtro de Gabor invariante a la rotación ajustando los parámetros del filtro utilizando un algoritmo genético denominado Evolución Diferencial, tal y como se definió en la sección 5.5.

Las grietas que pueden aparecer sobre la superficie de un túnel de hormigón no tienen ninguna dirección predominante, como ocurría en carreteras, en las que se dividían en grietas longitudinales y transversales. Un filtro de Gabor queda definido por 4 parámetros, la frecuencia F , la orientación θ y las desviaciones estándar de la función Gaussiana en la dirección de la onda y perpendicular a ella, σ_x y σ_y . La modificación propuesta en este trabajo consiste en combinar distintas orientaciones de un mismo filtro de Gabor, de tal forma que a cada pixel de la imagen filtrada se le asigna el máximo valor del módulo del filtro de Gabor entre las distintas orientaciones. Para abordar la detección de grietas en túneles, el número de direcciones del filtro de Gabor ha sido 16 y el algoritmo genético denominado Evolución Diferencial ha sido implementado utilizando las siguientes condiciones:

- La generación inicial, en lugar de generarse al azar, se realiza definiendo un banco de filtros, del tal forma que los filtros del banco abarquen todo el espacio frecuencial. Los parámetros del banco de filtros se encuentran en la **Tabla 25**

Tabla 25. Parámetros del banco de filtros de Gabor invariante a la rotación de la generación inicial utilizada por el algoritmo genético de Evolución Diferencial

Parámetro	Unidades	Valor
Tamaño (MxN)	Píxeles x Píxeles	128 x 128
Número de Frecuencias (n_F)	-	12
Número de Orientaciones (n_θ)	-	16
Ancho de Banda en Frecuencia (B_F)	Octavas	0,5
Ancho de Banda en Frecuencia (B_θ)	Grados	11,25
Constante de solapamiento en x (K_x)	-	1
Constante de solapamiento en y (K_y)	-	1
Frecuencia central máxima (F_M)	Píxeles ⁻¹	0,427

El número de frecuencias, n_F , se calcula de tal forma que la frecuencia más baja del banco de filtros no sea inferior al inverso del tamaño del filtro, tal y como se propuso en la sección 5.3

La frecuencia central máxima, F_M , se calcula de tal manera que la isocurva de la mitad del valor de pico (valor máximo) del módulo de la transformada de Fourier, para la frecuencia máxima, tenga 0,5 como mayor valor de la frecuencia, que es la frecuencia de Nyquist, tal y como se propuso en la sección 5.3

- En cada generación hay 12 individuos, $N_p = 12$. Este valor se ha hecho coincidir con el número de frecuencias del banco de filtros descrito en el apartado anterior
- La probabilidad de mutación se ha establecido en el 50%, $CR = 0,5$.
- El vector de escala de la mutación, F , no es un valor fijo, sino que es función de la generación actual y del número de generaciones sin mejora, de acuerdo a la siguiente ecuación:

$$F = 1 - \frac{\text{n}^\circ \text{ de generación}}{\text{n}^\circ \text{ máximo de generaciones}} + \frac{\text{n}^\circ \text{ de generaciones sin mejora}}{\text{n}^\circ \text{ máximo de generaciones sin mejora}} \quad \text{Ec. 45}$$

De esta manera, el factor de escala disminuye al aumentar el número de generación actual y aumenta al transcurrir generaciones sin mejora. El valor de F siempre es mayor de 0 y menor 2.

- La selección del valor del umbral del clasificador se realiza minimizando el error ponderado del clasificador. Para ello, se ordena la respuesta de todas las muestras de entrenamiento y se calcula el valor del error ponderado para cada valor intermedio entre dos muestras, seleccionando finalmente el umbral con menor error ponderado.
- El criterio de parada se ha establecido para detener la búsqueda cuando se alcanza el número máximo de generaciones, que se ha sido fijado en 500, o el número máximo de generaciones sin mejora, que es de 50.

Al aplicar el algoritmo descrito a la base de datos establecida, el algoritmo alcanza los criterios de parada en la generación 191, habiendo obtenido el mejor resultado 50 generaciones antes, tras 67 horas de procesamiento. El error porcentual alcanzado es del 4,7325%, definido en términos de error ponderado, es decir, el valor complementario de la exactitud ponderada definida en Ec. 3 en la sección 4.2.1.

En la **Tabla 26** se muestra la matriz de confusión con los resultados obtenidos tras el entrenamiento, mientras que en la **Tabla 27** se muestran los valores de sensibilidad, especificidad, exactitud, precisión y exactitud ponderada. Estos estadísticos que fueron definidos en la sección 4.2.1.

7. Resultados

Tabla 26. Matriz de confusión obtenida en la detección de grietas en túneles mediante la utilización de filtros de Gabor invariantes a la rotación

		Clasificación del experto	
		Positivos	Negativos
Clasificación del clasificador	Positivos	TP = 19.050	FP = 3.772
	Negativos	FN = 950	TN = 76.228

Tabla 27. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante un filtro de Gabor invariante a la rotación

Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
0,9525	0,9529	0,9373	0,9876	0,9527

En la **Figura 68** la evolución del error ponderado al aumentar el número de generaciones cuando se aplica el algoritmo descrito. Se puede ver en la gráfica que a partir de la generación 19 el resultado se estabiliza, por lo que se podría haber obtenido un error ponderado del 4,80% en menos de 10 horas.

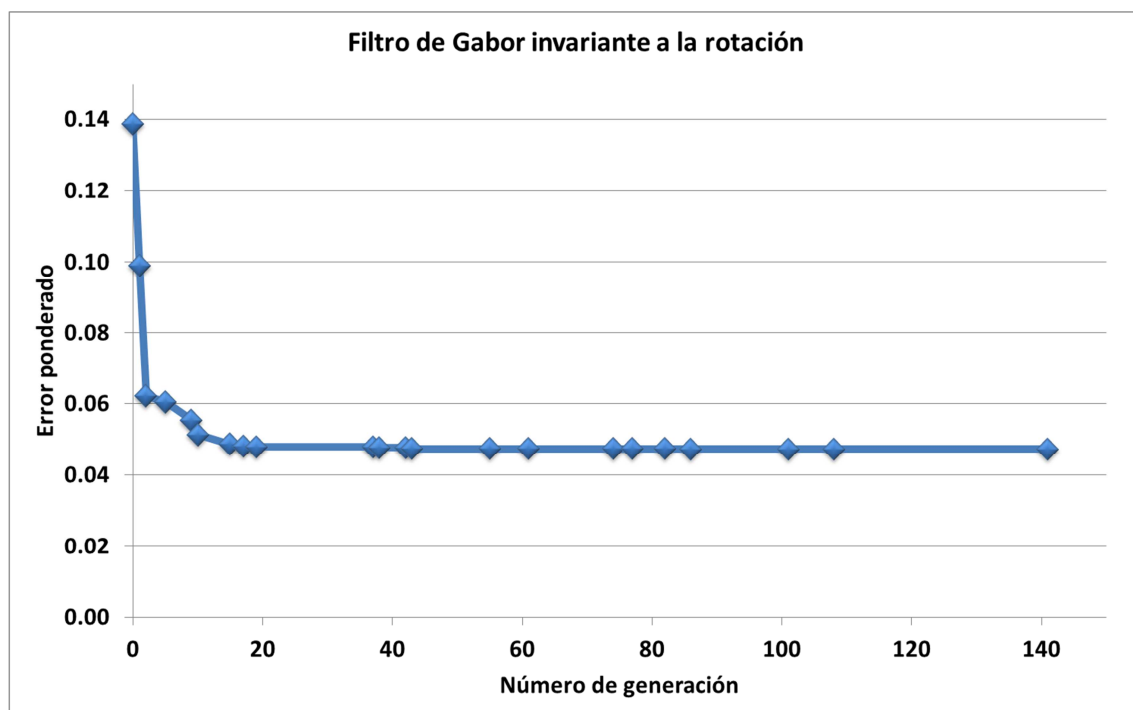


Figura 68. Evolución del error ponderado frente al número de generaciones al aplicar el algoritmo genético Evolución Diferencial para establecer los parámetros del filtro de Gabor invariante a la rotación para la detección de grietas en túneles

Los parámetros del filtro de Gabor obtenidos tras aplicar el algoritmo de Evolución Diferencial descrito en las generaciones 19 y 141 se encuentran en la **Tabla 28**.

Tabla 28. Parámetros de los filtros de Gabor en las generaciones 19 y 141

Generación	Frecuencia (píxeles ⁻¹)	Dispersión en X (píxeles)	Dispersión en Y (píxeles)	Umbral
19	0,1712	2,0566	17,7258	3,0623
141	0,1599	2,0813	17,1607	3,7301

Utilizando cualquiera de los dos resultados del entrenamiento anterior, el tiempo necesario para procesar una imagen de 1Mbyte para aplicar este algoritmo se sitúa en 65,6seg, es decir, que la velocidad de procesamiento es de 15,6kbytes/seg. Algunos resultados del procesamiento utilizando esta técnica se muestra en la **Figura 69**.

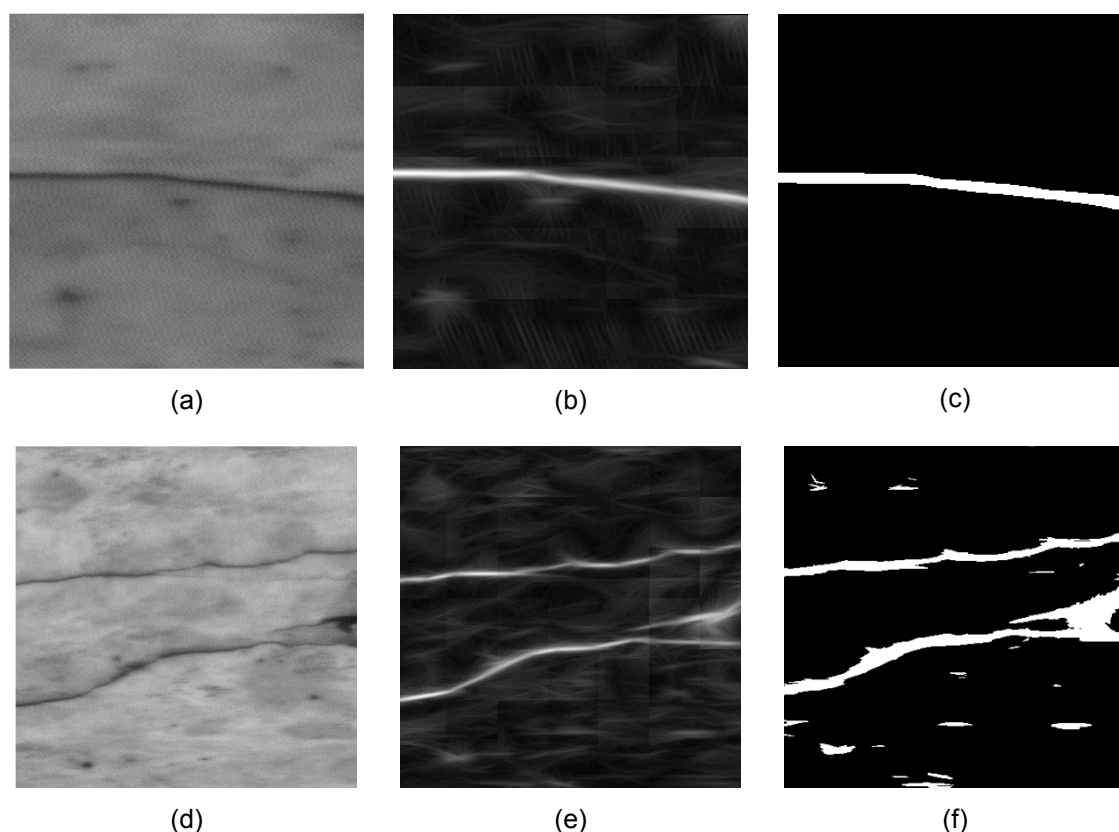


Figura 69. Ejemplo de aplicación del filtro de Gabor invariante a la rotación. (a) y (d) son las imágenes originales normalizadas; (d) y (e) son las imágenes filtradas; (c) y (f) las imágenes segmentadas

En la **Figura 69** se muestran dos ejemplos donde se aplica la técnica descrita con los parámetros obtenidos en la generación número 141. En esta figura se muestra la imagen original normalizada, la imagen filtrada y la imagen segmentada. Los valores de la imagen filtrada de la fila superior oscilan entre 0 y 9,91, pero para visualizarlos se han ajustado al intervalo 0-255. En la fila inferior oscilan entre 0 y 15,87 y han sido ajustados al mismo intervalo.

7.3.2 Textura de anisotropía condicional

Para la detección de grietas en túneles también se ha utilizado la textura de anisotropía condicional (*Conditional Texture Anisotropy*, CTA). Este método de procesamiento de imágenes es válido para detectar defectos que tienen una dirección predominante pero no conocida como ocurre con las grietas. En un primer lugar se ha probado el algoritmo desarrollado por Nguyen et al., [Ngu09], en el que el vector de características está formado por la media y la desviación estándar de los niveles de gris de la imagen en una dirección. El algoritmo se describe con detalle en la sección 5.6.

El número de direcciones utilizado para aplicar este algoritmo es de 8 ($k=8$) y la recta donde se evalúan las características seleccionadas tiene una longitud de 49 píxeles ($l=49$). Por tanto, este algoritmo analiza en cada píxel de la imagen el valor de la media y la desviación estándar de los píxeles pertenecientes a las 8 líneas rectas de 49 píxeles que pasan por dicho píxel en 8 direcciones. La dirección de cada recta forma un ángulo distinto y están separados de forma homogénea, de tal forma que cada recta tiene una inclinación de 22,5° más que la anterior.

El algoritmo descrito presenta un error porcentual ponderado del 6,90% alcanzado el resultado en menos de 1 hora de entrenamiento. Una vez realizado el entrenamiento, el tiempo necesario para procesar una imagen de 1Mbyte con este algoritmo es de 810seg, es decir, se procesan 1,2kbytes/seg. La matriz de confusión resultado de aplicar el algoritmo en la base de datos disponible para la detección de grietas en túneles se muestran a continuación en la **Tabla 29**.

Tabla 29. Matriz de confusión obtenida en la detección de grietas en túneles mediante el algoritmo CTA descrito por Nguyen et al.

		Clasificación del experto	
		Positivos	Negativos
Clasificación del clasificador	Positivos	TP = 18.508	FP = 5.068
	Negativos	FN = 1.492	TN = 74.932

En la **Tabla 30** se muestran los valores de sensibilidad, especificidad, exactitud, precisión y exactitud ponderada, estadísticos que fueron definidos en la sección 4.2.1.

Tabla 30. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante el algoritmo CTA de Nguyen et al.

Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
0,9254	0,9367	0,9344	0,9805	0,9310

El umbral del valor CTA, que está definido en la Ec. 34 de la sección 5.6, se calcula de tal forma que se minimice el error ponderado, o lo que es lo mismo que se maximice la exactitud ponderada. El valor del umbral obtenido al aplicar esta técnica es 0,1048. En la **Figura 70** se muestran dos ejemplos donde se aplica esta técnica.

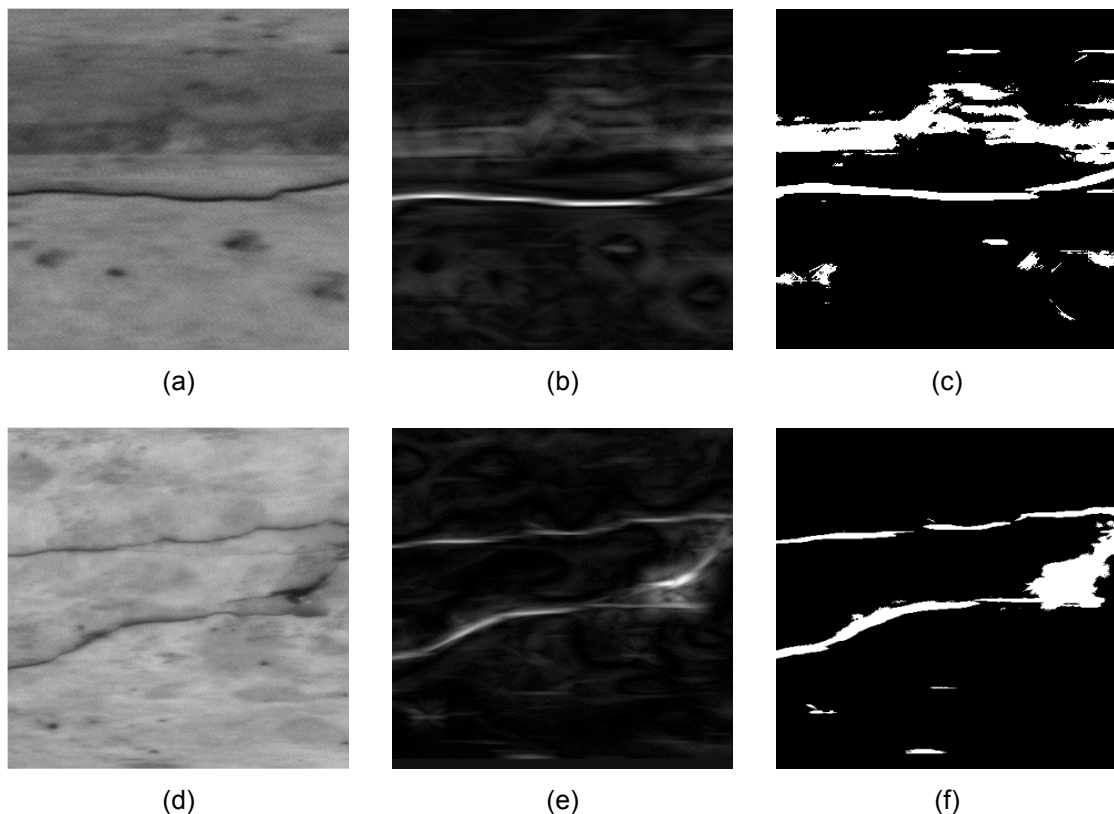


Figura 70. Ejemplo de aplicación del método original de textura de anisotropía condicional. (a) y (d) son las imágenes originales normalizadas; (b) y (e) son las imágenes filtradas; (c) y (f) las imágenes segmentadas

En **Figura 70** se muestra la imagen original normalizada, la imagen filtrada y la imagen segmentada. Los valores de la imagen filtrada de la fila superior oscilan entre 0 y 0,59, pero para visualizarlos se han ajustado al intervalo 0-255. En la fila inferior oscilan entre 0 y 0,64 y han sido ajustados al mismo intervalo.

El cálculo del valor de la desviación estándar del nivel de gris en cada línea es un proceso muy lento. Aunque el entrenamiento se realiza a una velocidad razonable, ese tiempo es similar al que sería necesario para aplicar el filtro. En la sección 5.6 se propuso que el vector de características solamente contuviera una variable, la media. Además también se propuso que la media de la imagen se calculará en un entorno circular de cada punto con un diámetro de 255 píxeles.

De esta manera se consigue reducir el tiempo de entrenamiento a menos de 10 minutos y obteniendo un error porcentual ponderado del 7,69%, muy próximo al error del método original. La velocidad de procesamiento de la imagen se reduce de forma similar, requiriendo 135seg para procesar una imagen de 1Mbyte, es decir, 7,56kbytes/seg. La matriz de confusión con los resultados obtenidos se muestra en la **Tabla 31**.

Tabla 31. Matriz de confusión obtenida en la detección de grietas en túneles mediante el algoritmo CTA con un único valor del vector de características y utilizando la media de la imagen en un entorno circular

		Clasificación del experto	
		Positivos	Negativos
Clasificación del clasificador	Positivos	TP = 18.346	FP = 5.694
	Negativos	FN = 1.654	TN = 74.306

En la **Tabla 32** se muestran los valores de sensibilidad, especificidad, exactitud, precisión y exactitud ponderada, estadísticos que fueron definidos en la sección 4.2.1.

Tabla 32. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante el algoritmo CTA con un único valor del vector de características y utilizando la media de la imagen en un entorno circular

Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
0,9173	0,9288	0,9265	0,9782	0,9231

El valor del umbral para el valor CTA que permite distinguir las grietas del resto de la imagen es 0,1435. Para su cálculo se emplea el mismo método que para el algoritmo original.

En la **Figura 71** se muestran dos ejemplos donde se aplica la modificación descrita. En esta figura se muestra la imagen original normalizada, la imagen filtrada y la imagen segmentada. Los valores de la imagen filtrada de la fila superior oscilan entre 0 y 0,66, pero para visualizarlos se han ajustado al intervalo 0-255. En la fila inferior oscilan entre 0 y 0,57 y han sido ajustados al mismo intervalo.

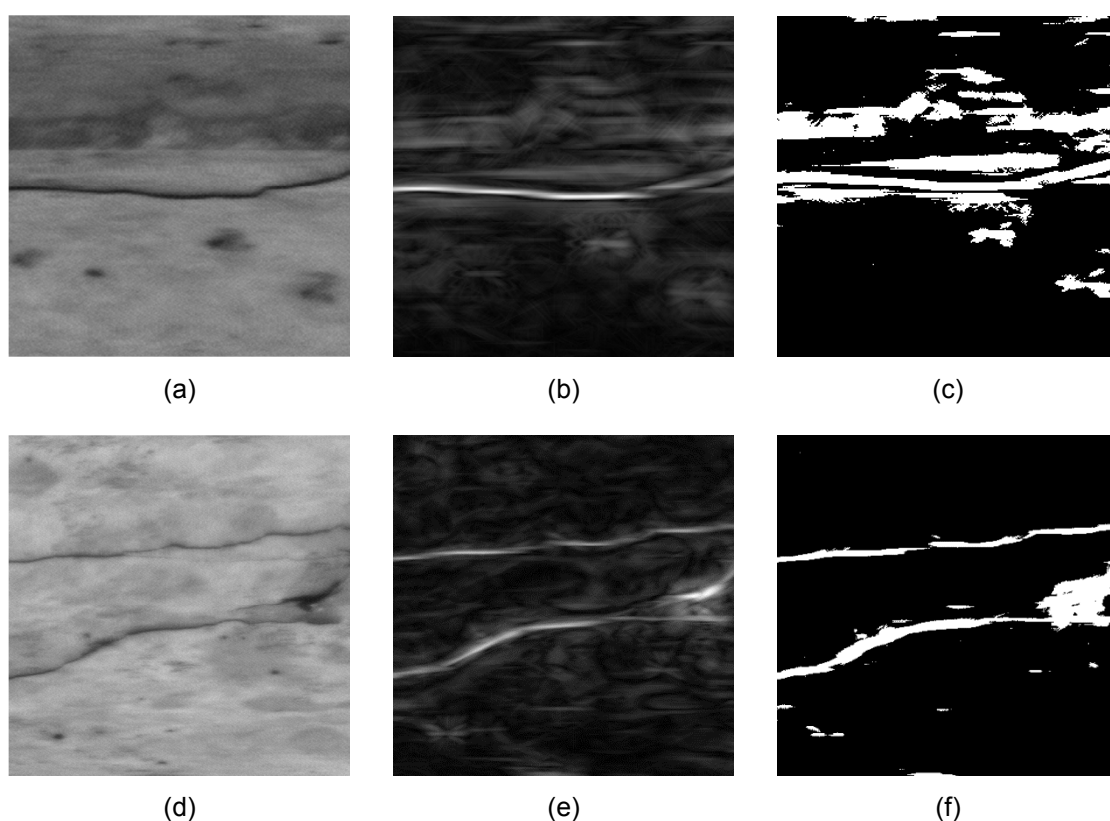


Figura 71. Ejemplo de aplicación de variante del método de textura de anisotropía condicional en el que solo se utiliza la media. (a) y (d) son las imágenes originales normalizadas; (d) y (e) es la imagen filtrada; (c) y (f) las imágenes segmentadas

Como se puede comprobar, el resultado obtenido con el nuevo método propuesto es muy similar al resultado obtenido utilizando el algoritmo original de Nguyen et al., pero con un tiempo de procesamiento mucho menor. Por tanto, en este caso el valor de la desviación estándar no aporta ninguna ventaja a la detección.

El cálculo de la media en un entorno circular es la operación que mayor tiempo consume en el proceso de cálculo en el anterior método propuesto. Por ello, se ha buscado la forma de eliminar esa operación manteniendo los resultados. En la sección

5.6 se analizó la ecuación que describía la probabilidad de que una zona de la imagen no contuviera defectos, llegando a la conclusión de que no era la más apropiada y se propuso la utilización de una nueva función de probabilidad que se ajustaba a la forma de una sigmoide, Ec. 37.

Aplicando esta nueva modificación del algoritmo CTA, se consigue reducir el error porcentual ponderado al 5,40% y el tiempo de entrenamiento se reduce a 9 segundos. El tiempo de procesamiento una vez entrenado el sistema también se reduce significativamente requiriendo 1,16seg para procesar una imagen de 1Mbyte, es decir, la velocidad de procesamiento se incrementa hasta los 882,3Kbytes/seg.

Tabla 33. Matriz de confusión obtenida en la detección de grietas en túneles mediante el algoritmo CTA cuya función de probabilidad se ajusta a una sigmoide

		Clasificación del experto	
		Positivos	Negativos
Clasificación del clasificador	Positivos	TP = 18.832	FP = 3.969
	Negativos	FN = 1.168	TN = 76.031

En la **Tabla 34** se muestran los valores de sensibilidad, especificidad, exactitud, precisión y exactitud ponderada, estadísticos que fueron definidos en la sección 4.2.1.

Tabla 34. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante el algoritmo CTA cuya función de probabilidad se ajusta a una sigmoide

Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
0,9416	0,9504	0,9486	0,9849	0,9460

En la **Figura 72** se muestran dos ejemplos donde se aplica la variante al método de textura de anisotropía condicional descrita, en el que se ha introducido una nueva forma de calcular la probabilidad. En esta figura se muestra la imagen original normalizada, la imagen filtrada y la imagen segmentada. Los valores de la imagen filtrada de la fila superior oscilan entre 0 y 0,59, pero para visualizarlos se han ajustado al intervalo 0-255. En la fila inferior oscilan entre 0 y 0,65 y han sido ajustados al mismo intervalo.

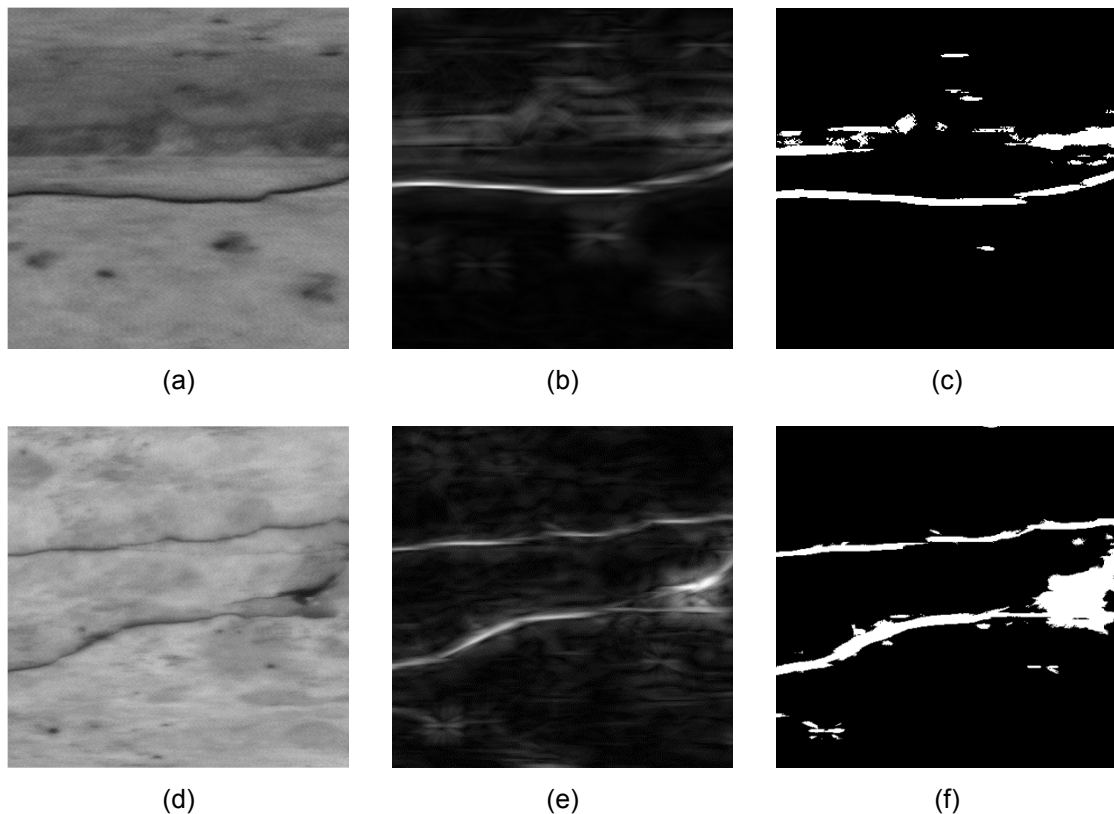


Figura 72. Ejemplo de aplicación de variante del método de textura de anisotropía condicional en el que se utiliza la función sigmoide para calcular la probabilidad. (a) y (d) son las imágenes originales normalizadas; (b) y (e) es la imagen filtrada; (c) y (f) las imágenes segmentadas

Este método consigue resultados significativamente superiores a los anteriores. Si se comparan con los obtenidos con el filtro de Gabor invariante a la rotación se puede ver que son similares, ya que se obtuvo un error porcentual ponderado del 4,73% al aplicar el filtro de Gabor invariante a la rotación y el 5,40% con la modificación del algoritmo CTA propuesta en este trabajo. En la siguiente sección se propone combinar estos dos algoritmos para intentar mejorar aún más los resultados.

7.3.3 Textura de anisotropía condicional de la imagen filtrada por un filtro de Gabor invariante a la rotación

Tanto el filtro de Gabor invariante a la rotación como la modificación al algoritmo CTA propuesto alcanzan buenos resultados de forma independiente. Se ha propuesto también aplicar la modificación del algoritmo CTA a la imagen filtrada por un filtro de Gabor invariante a la rotación. Para ello, es necesario modificar la función de probabilidad del algoritmo CTA que mide la probabilidad de que una zona de la imagen no contenga una grieta, ya que una imagen filtrada por el filtro de Gabor invariante a

7. Resultados

la rotación presenta valores altos en las zonas con defectos, justo al revés que la imagen sin filtrar.

Para su aplicación, es necesario ajustar de nuevo los parámetros del filtro de Gabor para lo que se utiliza el algoritmo genético de Evolución Diferencial. Se utiliza exactamente la misma configuración que la descrita en la sección 7.3.1. El algoritmo se ha detenido en la generación número 100, habiendo alcanzado su mejor valor en la generación 50, tras 31 horas de procesamiento. El error porcentual alcanzado es del 4,33%, cuatro décimas menos que al aplicar el filtro de Gabor invariante a la rotación. El tiempo necesario para procesar 1Mbyte es 68,8seg, aproximadamente la suma del tiempo empleado por los dos algoritmos que se han combinado, y por tanto su velocidad de procesamiento es de 14,88Kbytes/seg.

Los resultados detallados se muestran a continuación. En la **Tabla 35** se muestra la matriz de confusión.

Tabla 35. Matriz de confusión obtenida en la detección de grietas en túneles mediante la aplicación de una modificación del algoritmo CTA a imágenes filtradas por un filtro de Gabor invariante a la rotación

		Clasificación del experto	
		Positivos	Negativos
Clasificación del clasificador	Positivos	TP = 18.920	FP = 2.603
	Negativos	FN = 1.080	TN = 77.397

En la **Tabla 36** se muestran los valores de sensibilidad, especificidad, exactitud, precisión y exactitud ponderada, estadísticos que fueron definidos en la sección 4.2.1.

Tabla 36. Sensibilidad, especificidad, exactitud, precisión y exactitud ponderada en la detección de grietas en túneles mediante la aplicación de una modificación del algoritmo CTA a imágenes filtradas por un filtro de Gabor invariante a la rotación

Sensibilidad	Especificidad	Exactitud	Precisión	Exactitud ponderada
0,9460	0,9675	0,9632	0,9862	0,9567

A continuación, en la **Figura 73** se muestra la evolución del error ponderado al aumentar el número de generaciones cuando se aplica el algoritmo descrito. Se puede ver en la gráfica que a partir de la generación 27 el resultado se estabiliza, alcanzando

un error porcentual ponderado del 4,38% frente al 4,33% alcanzado al cumplirse el criterio de parada establecido.

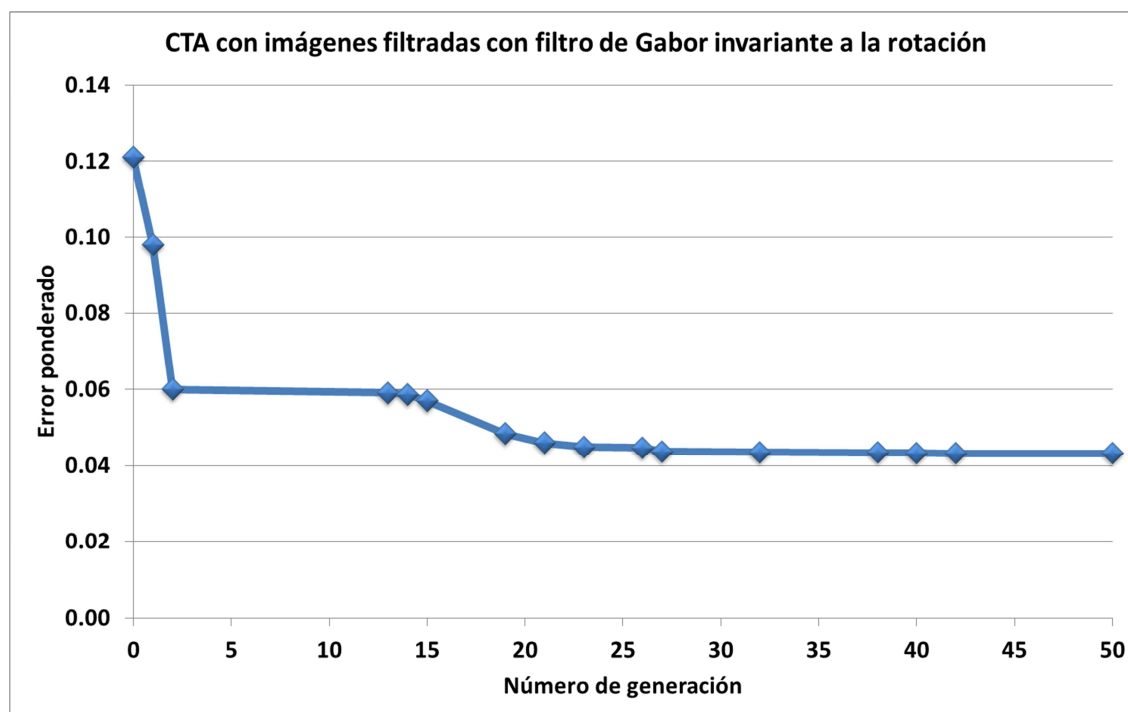


Figura 73. Evolución del error ponderado frente al número de generaciones al aplicar el algoritmo genético Evolución Diferencial para establecer los parámetros del filtro de Gabor invariante a la rotación utilizado para filtrar las imágenes de entrada al algoritmo CTA para la detección de grietas en túneles

Los parámetros del filtro de Gabor obtenidos tras aplicar el algoritmo de Evolución Diferencial descrito en las generaciones 27 y 50 son los siguientes:

Tabla 37. Parámetros de los filtros de Gabor en las generaciones 27 y 50

Generación	Frecuencia (píxeles ⁻¹)	Dispersión en X (píxeles)	Dispersión en Y (píxeles)	Umbral
27	0,1605	2,0500	10,6012	0,1369
50	0,1717	2,0813	7,2576	0,1181

En la **Figura 74** se muestran dos ejemplos donde se aplica la técnica descrita con los parámetros obtenidos en la generación número 50. En esta figura se muestra la imagen original normalizada, la imagen filtrada y la imagen segmentada. Los valores

de la imagen filtrada pueden oscilar entre 0 y 1. Para representarlos en una imagen de 8 bits se multiplica su valor por 255.

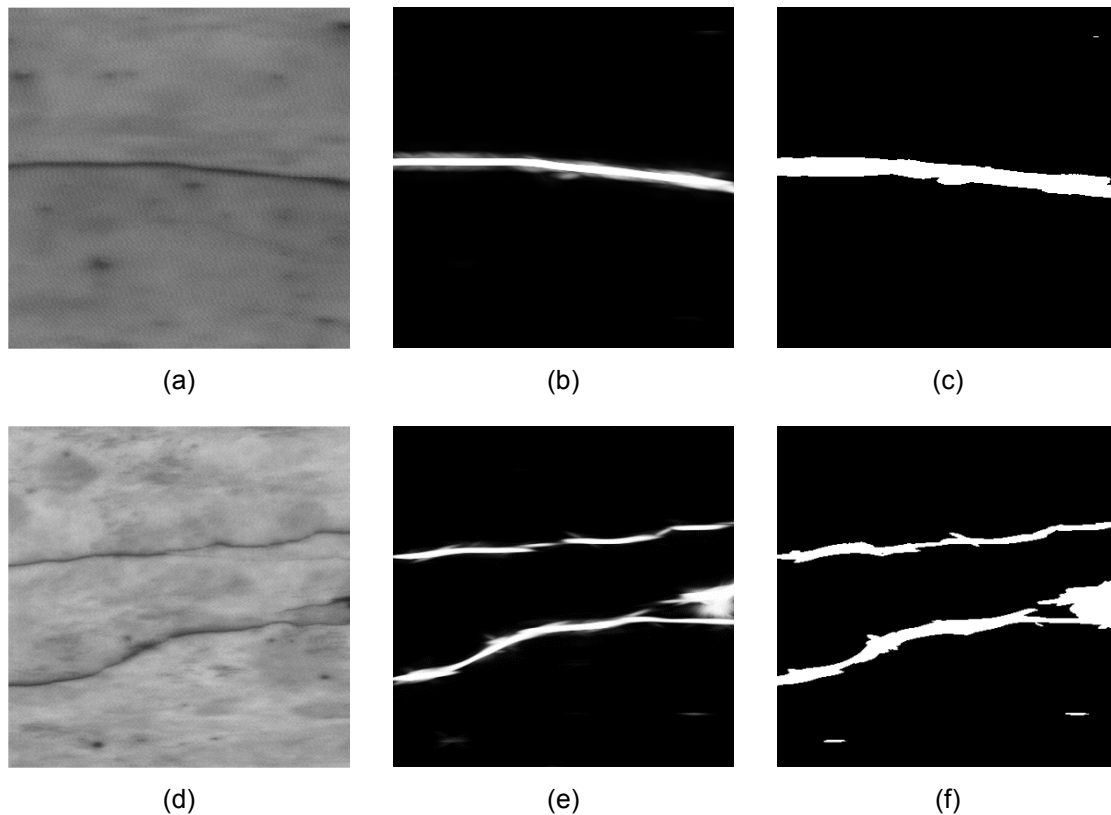


Figura 74. Ejemplo de aplicación de la textura de anisotropía condicional de la imagen filtrada de un filtro de Gabor invariante a la rotación. (a) y (d) son las imágenes originales normalizadas; (d) y (e) es la imagen filtrada; (c) y (f) las imágenes segmentadas

7.3.4 Conclusiones

Se ha abordado la detección de grietas en la superficie de túneles de hormigón mediante el análisis de imágenes adquiridas con un prototipo adaptado para la inspección de túneles ferroviarios y carreteros.

Para el análisis de las imágenes se han comprobado varios métodos de análisis de imágenes. En un primer lugar se ha propuesto una nueva estrategia para utilizar los filtros de Gabor en la detección de defectos que no tienen una dirección conocida y que se ha denominado filtro de Gabor invariante a la rotación. Este filtro alcanza una exactitud ponderada de 0,9527.

Se han comparado los resultados con los que se obtienen de aplicar un método de procesamiento para la detección de grietas que utiliza la textura de anisotropía condicional. Este método fue propuesto por Nguyen et al., [Ngu09], para la detección

de grietas en carreteras. Su aplicación a las mismas imágenes anteriormente descritas alcanza una exactitud ponderada de 0,9310.

A pesar de ser un buen resultado, no alcanza el obtenido por el filtro de Gabor invariante a la rotación propuesto en esta tesis. Tras analizar con detalle el funcionamiento de este método, se han propuesto varias modificaciones para mejorar los resultados y disminuir los tiempos de procesamiento.

En la primera de estas modificaciones se excluye la desviación estándar del vector de características propuesto por Nguyen et al., y se propone substituir el valor de la media en toda la imagen por el valor de la media en un entorno circular de cada pixel. El tiempo de procesamiento se reduce a la sexta parte, con una exactitud ponderada de 0,9231, ligeramente inferior al método original.

La segunda modificación consiste en modificar la forma de calcular la probabilidad de que la superficie no contenga defecto, uno de los elementos clave en la aplicación del método, eliminando la necesidad de calcular la media en un entorno circular de cada pixel. La disminución del tiempo de procesamiento es muy elevada, analizando las imágenes 67 veces más rápido. Los resultados también mejoran sensiblemente, alcanzado una exactitud ponderada de 0,9460, muy próxima a la obtenida por el filtro de Gabor invariante a la rotación.

Finalmente se han combinado estas dos técnicas. Esta combinación consiste en utilizar las imágenes filtradas por el filtro de Gabor invariante a la rotación a la modificación del método de textura de anisotropía condicional propuesta en esta tesis. El tiempo para el entrenamiento de esta técnica es muy elevado, obteniendo una exactitud ponderada de 0,9567, algo superior a la obtenida al utilizar únicamente el filtro de Gabor circular.

Comparando los resultados con otros trabajos, se puede comprobar que los resultados obtenidos son satisfactorios. Fujita y Hamamoto, en [Fuj11] obtenían una sensibilidad de 0,80 y una especificidad de 0,992, es decir, una exactitud ponderada de 0,8960, por debajo de todos los métodos analizados. Zhang et al., en [Zha14] alcanzaban una exactitud aproximada de 0,9 en los diversos métodos que analizaban, mientras que los métodos que se han analizado en esta sección varían entre el 0,9265 del método de la textura de anisotropía condicional propuesto por Nguyen et al. y el 0,9632 obtenido al aplicar el método de anisotropía condicional a la imagen filtrado por el filtro de Gabor circular, tal y como se ha propuesto en esta tesis. Por último, el método propuesto por Zhao et al. en [Zha15] obtiene una sensibilidad que varía entre el 0,8 y el 0,9, dependiendo de la forma de evaluar su comportamiento en las proximidades de las grietas. En los métodos analizados en esta sección la sensibilidad

se encuentra entre el 0,8948 y el 0,9525, por tanto, superior a la alcanzada por Zhao et al.

8 CONCLUSIONES

La presente tesis propone una metodología para la **inspección de superficies continuas** que comprende las etapas de adquisición de imágenes, su procesamiento y la utilización de la información obtenida. Las técnicas descritas en esta tesis pueden ser utilizadas en distintos ámbitos de inspección, aunque lógicamente cada aplicación requerirá una adaptación específica al tipo de superficie inspeccionada y a los defectos que sobre ella pueden aparecer. La implantación de tres sistemas de inspección, donde se emplean las metodologías propuestas, permite la evaluación de las técnicas de inspección desarrolladas en diferentes ámbitos de la vida real.

La adquisición de imágenes de superficies continuas se realiza preferiblemente utilizando **cámaras lineales**. Estas cámaras permiten obtener imágenes de gran calidad y resolución a alta velocidad y simplifican el sistema de iluminación. Para ello es necesario sincronizar la captura de las imágenes mediante un codificador incremental e implementar sistemas eficaces para el almacenamiento de la ingente cantidad de información que generan.

El procesamiento de las imágenes se realiza mediante un novedoso enfoque en el que las imágenes son analizadas mediante la **caracterización de anomalías en el dominio espacio-frecuencial**. Las posibilidades que ofrece el **filtro de Gabor** para el análisis de imágenes son explotadas para la detección y clasificación de defectos, que mediante la combinación, adaptación y desarrollo de otras técnicas de procesamiento de imágenes han permitido obtener unos resultados de detección y clasificación de defectos satisfactorios.

Los resultados del procesamiento son mostrados en un **entorno de visualización** que permite el acceso sencillo a la información, la presentación de información complementaria sincronizada cuando es necesario y la generación de informes a medida para el cliente.

Estos tres pilares básicos constituyen el tronco común sobre el cual se ha abordado la resolución de tres problemas reales de inspección superficial, adaptándose a las peculiaridades de cada aplicación.

El primer problema abordado ha sido la **inspección superficial de chapa de acero laminado**. Se ha desarrollado un sistema que realiza la inspección por ambas caras de la superficie utilizando 5 cámaras por cada cara y 13 equipos informáticos que permiten capturar, almacenar y procesar las imágenes ofreciendo los resultados de inspección en tiempo real. El sistema de inspección sigue una arquitectura jerárquica

que permite coordinar todo el proceso. El emplazamiento del sistema, el enmascaramiento de los defectos provocado por el aceite utilizado para lubricar la chapa, las vibraciones, la presencia de suciedad, la alta cadencia de la línea y la gran variabilidad de la apariencia visual de los defectos son los principales retos que se abordan en esta aplicación. La clasificación de defectos suele realizarse utilizando características de forma y nivel de gris de las agrupaciones de píxeles extraídas en la etapa de segmentación. En esta tesis se combinan tres clasificadores cuyas entradas también incluyen características extraídas de imágenes filtradas por filtros de Gabor. El método propuesto para extraer estas características consiste en calcular el valor medio del módulo de la imagen filtrada en un entorno de la zona del defecto cuyo nivel de gris se aleje más del valor esperado de la zona sin defectos, pero utilizando sólo los píxeles marcados como defectos en la segmentación. De esta forma se consigue extraer información más precisa del defecto y se reduce el tiempo de procesamiento. El sistema clasifica 6 tipos de defectos (soldaduras, moho blanco, marcas de transporte, calvas, grumos de cinc e impresiones) identificando también las agrupaciones de píxeles extraídas en la etapa de segmentación que realmente no corresponden a defectos. La exactitud global del clasificador combinado es del 96,70%, clasificando correctamente el 87,42% de los defectos, en línea con los obtenidos por otros trabajos que alcanzan una exactitud global entre el 80% y el 97%, aunque en estos trabajos las verificaciones se realizan utilizando grupos de imágenes que solo contienen los defectos esperados.

El segundo problema abordado ha sido la **inspección superficial de carreteras**. Con este fin se ha instalado en un vehículo el sistema de adquisición de imágenes LRIS, cuya información es enriquecida con dos perfilómetros láser, una cámara de video convencional de alta resolución para visualizar el entorno y un GPS para georreferenciar toda la información. Todos los dispositivos adquieren la información de forma sincronizada utilizando un codificador incremental que, junto a la electrónica de control apropiada, permite adquirir la información con distintas frecuencias de muestreo. La naturaleza del asfalto que está compuesto por áridos y betún, materiales que aparecen en las imágenes con un alto contraste entre ellos y cuya distribución se ve afectada por la degradación de la carretera, provoca que las habituales técnicas basadas en segmentación no funcionen satisfactoriamente. El procesamiento de las imágenes se realiza analizando las imágenes filtradas por un conjunto de bancos de filtros de Gabor que se combinan utilizando el algoritmo AdaBoost. Una base de datos compuesta por imágenes de grietas longitudinales y transversales así como zonas sin ningún deterioro se utiliza para entrenar y verificar el funcionamiento del algoritmo propuesto. Dos clasificadores distintos son utilizados para detectar grietas longitudinales y transversales, realizando la detección y clasificación en un solo paso. El clasificador fuerte propuesto en esta tesis para detectar grietas transversales alcanza una exactitud ponderada de 0,9384 mientras que el encargado de realizar la

clasificación de las grietas longitudinales obtiene un 0,9090. Estos datos están en línea o incluso mejoran los obtenidos en otros trabajos analizados en esta tesis, más aún si se observa que en otros trabajos se incluyen en la detección otros tipos de deterioros como grietas selladas cuya detección es significativamente más sencilla.

Por último, también se ha implementado un **prototipo para la inspección superficial del revestimiento de túneles**. Las grietas son los defectos objeto de estudio en esta aplicación, por ser uno de los defectos más frecuente y difíciles de detectar. En esta tesis se analizan las distintas posibilidades para la inspección completa de la superficie de túneles con distintas tipologías y con objetivos de detección variables. Se describen soluciones factibles para la detección de grietas de hasta 0,4mm de abertura, pudiéndose inspeccionar grietas de hasta 0,1mm en el caso de túneles con diámetros más pequeños, aunque para ello la resolución de la imagen sería de 0,05mm, lo que supone una ingente cantidad de información que manejar. El prototipo desarrollado permite la adquisición de una pequeña porción del diámetro del túnel, aunque la solución propuesta es escalable. Las imágenes adquiridas de dos túneles reales se utilizan para verificar el funcionamiento de un conjunto de algoritmos propuestos para la detección de grietas y fisuras. Por un lado, se propone la utilización de lo que se denomina en esta tesis filtro de Gabor invariante a la rotación. Consiste en la combinación de filtros de Gabor de distintas orientaciones con el objetivo de realzar en la imagen la presencia de grietas y fisuras, independientemente de la dirección que éstas presenten. Obtiene una exactitud ponderada del 0,9527. Por otro lado, se utiliza la textura de anisotropía condicional con el mismo objetivo. Se proponen algunas modificaciones al método original de Nguyen et al. para mejorar los resultados y aumentar la velocidad de procesamiento. Utilizando una nueva función de probabilidad y analizando únicamente la media de la imagen, se consigue reducir la velocidad de procesamiento 90 veces incrementando la exactitud ponderada desde el 0,9310 que se obtiene aplicando el método original, al 0,9460 obtenida con la modificación propuesta. Finalmente se han combinado los dos métodos propuestos anteriormente aplicando el algoritmo de textura de anisotropía condicional propuesto en esta tesis a las imágenes filtradas por el filtro de Gabor invariante a la rotación. El resultado obtenido alcanza el 0,9567, mejorando también los resultados individuales. Los resultados obtenidos en otros trabajos en los que el objetivo es la detección de grietas en superficies de hormigón obtienen una exactitud ponderada entorno al 0,9, inferior al alcanzado en esta tesis.

Por tanto, los resultados obtenidos en la presente tesis son satisfactorios aunque son susceptibles de ser ampliados y mejorados, lo que deja la puerta abierta a **líneas futuras de investigación**. En este trabajo, la inspección de las superficies se basa en el análisis de las imágenes, si bien es cierto que otras tecnologías podrían complementar y mejorar los resultados obtenidos así como ampliar el rango de

defectos que se pudieran detectar. Algunos de los defectos estudiados también muestran anomalías geométricas, por lo que la incorporación de **información tridimensional** mejoraría sin duda los resultados. Algunos equipos comerciales para la inspección de grietas y fisuras en carreteras y en túneles mediante la extracción de información tridimensional están comenzando a aparecer en el mercado, aunque no siempre alcanzan la resolución y precisión requerida, y su precio sigue siendo elevado. El procesamiento automático de la información tridimensional será sin duda un reto que se abordará en los próximos años. Aunque no es objetivo de esta tesis, ya hemos dado los primeros pasos para abordar este problema. Por un lado, se ha desarrollado una herramienta informática para el mapeado la textura de una superficie captura con una cámara digital sobre el modelo tridimensional generado con un escáner láser, [Zal11]. El resultado ofrecido abre la puerta al análisis combinado de imágenes sobre superficies tridimensionales. Por otro lado, en el ámbito de la inspección de carreteras, se ha propuesto un método que combina información visual con información geométrica para la detección de grietas, [Med14].

Las **imágenes termográficas** también podrían aportar información relevante a la inspección de ciertas superficies, especialmente en la detección de defectos en los que aparece humedad. Las mejoras en los últimos años en esta tecnología permiten disponer de equipos a un coste razonable, aunque la velocidad de adquisición a alta resolución sigue siendo muy inferior a la que se alcanza en el espectro visible.

Aunque esta tesis se centra en la detección y clasificación de los defectos más frecuentes en las superficie estudiadas, sobre éstas pueden aparecer otros tipos de defectos, tal y como se describió en el capítulo 2. Algunos de estos defectos se manifiestan de forma evidente en las imágenes aunque hay otros que resulta difícil de apreciar. Disponer de una base de datos con defectos que aparecen con una frecuencia baja es siempre una tarea complicada pero necesaria para desarrollar y verificar las técnicas de procesamiento más apropiadas para cada caso.

Por último, también se ha detectado la necesidad de caracterizar la **intensidad de los defectos** detectados. Aunque algunas variables de estos defectos relacionados con su tamaño ya son medidas de forma objetiva, en muchas ocasiones la intensidad del defecto es una medida subjetiva que solo un operador humano es capaz de asignar. Realizar esta caracterización de forma automática es un reto para el que todavía no se ha alcanzado una solución.

BIBLIOGRAFÍA

- [Aba12] Abadi, M., Koudeir, M., y Marchand, S. (2012), Gabor filter-based texture features to archaeological ceramic materials characterization. *Proceedings of 5th International Conference on Image and Signal Processing, ICISP 2012*, Agadir, Morocco, June 28-30, 2012, 7340:333-342, doi: 10.1007/978-3-642-31254-0_38
- [Abd03] Abdel-Qader, I., Abudayyeh, O. y Kelly, M., (2003), Analysis of Edge-Detection Techniques for Crack Identification in Bridges, *Journal of Computing in Civil Engineer*, 17(4):255–263, doi: 10.1061/(ASCE)0887-3801(2003)17:4(255)
- [Ahm04] Ahmadian, A., Mostafa, A., Abolhassani, M.D. y Alan, N.R., (2004), A method for texture classification of ultrasonic liver images based on Gabor Wavelet, *IEEE 7th International Conference on Signal Processing, ICPS04*, Troia, Turkey, aug 31 – sep 19, 2004, 2:971-974, ISBN: 0-7803-8406-7, doi: 10.1109/ICOSP.2004.1441482
- [Ben75] Bentley, J., (1975), Multidimensional binary search trees used for associative searching, *Communications of the ACM*, 18(9):509–517, doi: 10.1145/361002.361007
- [Ber96] Berchtold S., Keim D.A. y Kriegel H.P., (1996), The X-tree: an index structure for high-dimensional data. *Proceedings of 22nd International Conference on Very Large Data Bases, VLDB96*, 28–39, ISBN:1-55860-382-384
- [Bia12] Bianchini, A., (2012), Fuzzy representation of pavement condition for efficient pavement management, *Computer-Aided Civil and Infrastructure Engineering*, 27(8):608–19, doi:10.1111/j.1467–8667.2012.00758.x

- [Bia06] Bianconi, F. y Fernández, A. (2006), Granite texture classification with Gabor filters, *Proceedings of the 18th International Congress on Graphical Engineering*, Sitges, Spain, May 31 – June 2, 2006
- [Bla08] Blackledge, J. y Dubovitskiy, D., (2008), A Surface Inspection Machine Vision System that Includes Fractal Texture Analysis, *International Society for Advanced Science and Technology, Journal of Intelligent Systems*, 3(2):76-89, ISSN 1797-2329
- [Bro10] Brodersen, K.H., Ong, C.S., Stephan, K.E. y Buhmann, J.M., (2010), The balanced accuracy and its posterior distribution, *Proceedings of the 20th International Conference on Pattern Recognition (ICPR2010)*, Istanbul, Turkey, 23-26 Ago, 2010, 3121–3124, doi: 10.1109/ICPR.2010.764
- [Bul13] Bulnes, F.G., Usamentiaga, R., Garcia, D.F., Molleda, J., Rendueles, J.L., (2013), Periodic defects in steel strips: Detection through a vision-based technique, *IEEE Industry Applications Magazine*, 19(2):39-46 DOI: 10.1109/MIAS.2012.2215638
- [Ces87] Cestnik, G., Kononenko, I. y Bratko, I., (1987), Assistant-86: A Knowledge-Elicitation Tool for Sophisticated Users, *Progress in Machine Learning*, 31-45
- [Cha11] Chambon, S. y Moliard, J.M. (2011), Automatic road pavement assessment with image processing: review and comparison, *International Journal of Geophysics*, Volume 2011, Article ID 989354, DOI: 10.1155/2011/989354.
- [Cho14] Choi, D.C., Jeon, Y.J., Lee, S.J., Yun, J.P., y Kim, S.W., (2014), Algorithm for detecting seam cracks in steel plates using a Gabor filter combination method, *Applied Optics*, 53(22):4865-4872, doi: 10.1364/AO.53.004865
- [Chr08] Christmas, I., (2008), Steel's Climate Change Commitment, *World Steel Association*. <http://www.worldsteel.org/?action=storypages&id=294>. Consultada en Julio de 2010

- [Cla05] Clausi, D.A. y Deng, H., (2005), Design-Based texture feature fusion using gabor filters and co-occurrence probabilities, *IEEE Transactions on Image Processing*, 14(7): 925-936, doi: 10.1109/TIP.2005.849319
- [Con07] Cong, J.H., Yan, Y.H, Zhang, H.A y Li, J., (2007), Real-Time Surface Defects Inspection of Steel Strip Based on Difference Image, *International Symposium on Photoelectronic Detection and Imaging, Related Technologies and Applications. Proceedings of SPIE*, 6625:1-9, doi: 10.1117/12.790865
- [Cor10] Cord, A., Bach, F. y Jeulin, D., (2010), Texture classification by statistical learning from morphological image processing: application to metallic surfaces, *Journal of Microscopy*, 239(2):159–166, doi: 10.1111/j.1365-2818.2010.03365.x
- [Cor12] Cord, A. y Chambon, S., (2012), Automatic road defect detection by textural pattern recognition based on AdaBoost, *Computer-Aided Civil and Infrastructure Engineering*, 27:244–259, doi: 10.1111/j.1467–8667.2011.00736.x.
- [CSI15] CSIRO, (2015), *Automated Detection of Road Cracks*, <http://www.csiro.au/solutions/RoadCrack>, Consultado en Marzo 2015
- [Dib15] Dibit, (2015), Tunnel Scanning, <http://www.dibit.at/en/services/tunnel/tunnel-scanning/>, Consultado en Agosto de 2015
- [Don08] Dong, L., Yu, G., Ogunbona, P. y Li, W. (2008), "An efficient iterative algorithm for image thresholding", *Pattern Recognition Letter*, 29(9):1311-1316, doi: 10.1016/j.patrec.2008.02.001
- [Dud73] Duda, R.O. y Hart, P.E., (1973), *Pattern Classification and Scene Analysis*. John Wiley and Sons

- [Dup97] Dupont, F., Odet, C. y Carton, M., (1997), Optimization of the recognition of defects in flat steel products with the cost matrices theory, *NDT&E International*, 30(1):3–10, doi: 10.1016/S0963-8695(96)00045-X
- [Eur12] European Commission, (2012), EU Transport in Figures – Statistical PocketBook 2012, *Publications Office of the European Union*, Luxembourg, ISBN 978–92–79–21694–7, doi: 10.2832/52252
- [Faw06] Fawcett, T., (2006), An Introduction to ROC Analysis, *Pattern Recognition Letters*, 27(8):861–874, doi:10.1016/j.patrec.2005.10.010
- [Fre95] Freund, Y. y Schapire, R.E., (1995), A decision-theoretic generalization of on-line learning and an application to boosting, *Proceedings of the Second European Conference on Computational Learning Theory (EuroCOLT '95)*, Barcelona, España, Mar 1995, 23–37, doi: 10.1007/3-540-59119-2_166
- [Fug15] Fugro Roadware, (2015), *Automatic Road Analyzer (ARAN)*, http://www.roadware.com/products/survey_equipment/, consultado en Marzo de 2015
- [Fuj11] Fujita, Y. y Hamamoto, Y., (2011), A robust automatic crack detection method from noisy concrete surfaces, *Machine Vision and Applications*, 22:245–254, doi: 10.1007/s00138-009-0244-5
- [Fuk90] Fukuhara, T., Terada, K., Nagao, M., Kasahara, A., y Ichihashi, S., (1990), Automatic Pavement-Distress-Survey System, *Journal of Transportation Engineering*, 116(3):280–286, doi: 10.1061/(ASCE)0733-947X(1990)116:3(280)
- [Gab46] Gabor, D., (1946), Theory of communication, *Journal of Institute for Electrical Engineering – Part III: Radio and Communication Engineering*, 93(26):429-457, doi: 10.1049/ji-3-2.1946.0074

- [Gam91] Gammerman, A. y Thatcher, A.R., (1991), Bayesian diagnostic probabilities without assuming independence of symptoms, *Methods of Information in Medicine*, 30(1):15-22
- [Gar15] García, C., Vara, E., Medina, E., Llamas, J., Gómez-García-Bermejo, J. y Zalama, E.,(2015), Fusión de Información Visual y Dimensional para la Detección de Grietas y Fisuras en la Superficie de Carretera, *X Jornada Nacional de ASEFMA*, Zaragoza, España, Jul. 2015, Comunicación 32
- [Gay06a] Gayubo, F., González, L.M., Olmedo, D., Medina, R., Gómez-García-Bermejo, J. y Perán, J.R., (2006), Detección Visual en Línea de Roturas y Poros en Piezas de Chapa Embutida, *XXVII Jornadas de Automática*, Almería, España, Sep. 2006
- [Gay06b] Gayubo, F., González, L.M., Olmedo, D., Medina, R., Gómez-García-Bermejo, J. y González, J.L., (2006), Inspección visual en línea de defectos en piezas de chapa embutida, *XVI Congreso de Máquinas-Herramienta y Tecnologías de Fabricación*, San Sebastián, España, Oct. 2006
- [Gra93] Gramling, W.L., y Hunt, J. E., (1993), *Photographic Pavement Distress Record Collection and Transverse Profile Analysis*, Report SHRP-P-660, Strategic Highway Research Program, Transportation Research Board, National Research Council, Washington, D.C.
- [Guo14] Guo, J.M., Prasetyo, H y Wong, K, (2014), Vehicle Verification Using Gabor Filter Magnitude with Gamma Distribution Modeling, *IEEE Signal Processing Letters*, 21(5):600-604, doi: 10.1109/LSP.2014.2311132
- [Gut84] Guttman, A., (1984), R-trees: a dynamic index structure for spatial searching, *In Proceedings ACM SIGMOD International Conference on Management of Data, SIGMOD84*, 47–57, doi: 10.1145/602259.602266.
- [Han09] Hand, D.J., (2009), Measuring classifier performance: A coherent alternative to the area under the ROC curve, *Matching Learning*, 77(1):103-123, doi: 10.1007/s10994-009-5119-5

- [Hor89] Hornik K., Stinchcombe M. y White H., (1989), Multilayer feedforward networks are universal approximators, *Neural Networks*, 2(5):359-366, doi: 10.1016/0893-6080(89)90020-8
- [Hua06] Huang, Y., y Xu, B., (2006), Automatic inspection of pavement cracking distress, *Journal of Electronic Imaging*, 15(1):13-17, doi: 10.1117/1.2177650
- [Hub65] Hubel, D.H. y Wiesel, T.N., (1965), Receptive fields and functional architecture in two nonstriate visual areas visual areas 18 and 19 of the cat, *Journal of Neurophysiology*, 28:229-289
- [INO15] INO, (2015), Laser Road Imaging System (LRIS), [http://www.ino.ca/en/examples/laser-road-imaging-system\(Iris\)](http://www.ino.ca/en/examples/laser-road-imaging-system(Iris)), consultado en Marzo de 2015
- [Int15] International Cybernetics Corporation, (2015), *Pavement Imaging*, http://www.intlcybernetics.com/pavement_imaging.html, consultado en Marzo de 2015
- [Jah11] Jahanshahi, M., Masri, S. y Sukhatme, G., (2011), Multi-image stitching and scene reconstruction for evaluating defect evolution in structures, *Structural Health Monitoring*, 10(6):643-657, doi: 10.1177/1475921710395809
- [Jai90] Jain, A.K. y Farrokhnia, F., (1990), Unsupervised texture segmentation using Gabor filters, *IEEE International Conference on Systems, Man and Cybernetics*, Los Angeles, EE.UU, Nov 4-7,1990, 14-19, ISBN 0-87942-597-0, doi: 10.1109/ICSMC.1990.142050
- [Jai00] Jain, A.K., Duin, R.P.W., and Mao, J., (2000), Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Recognition and Machine Intelligence*, 22(1):4-37, 10.1109/34.824819

- [Kan05] Kang, G.W. y Liu, H.B., (2005), Surface defects inspection of cold rolled strips based on neural network, *Proceedings of 2005 International Conference on Machine Learning and Cybernetics*, Guangzhou, China, Aug 2005, 8:5034–5037, doi: 10.1109/ICMLC.2005.1527830
- [Kon90] I Kononenko I., (1990), Comparison of Inductive and I Bayesian Learning Approaches to Automatic Knowledge Acquisition, *Current Trends in Knowledge Acquisition*. IOS Press
- [Kon07] Kong, Z. y Cai, Z., (2007), Advances of research in fuzzy integral for classifiers' fusion, *Proceedings of the 8th ACIS international conference on software engineering, artificial intelligence, networking and parallel/distributed computing, SNPD 2007*, 2:809–814, ISBN: 978-0-7695-2909-7, doi: 10.1109/SNPD.2007.422
- [Kun04] Kuncheva, L.I., (2004), *Combining Pattern Classifiers: Methods and Algorithms*, John Wiley and Sons, ISBN:0471210781
- [Lau14] Laurent, J., Fox-Ivey, R., Sanchez Dominguez, F. y Ramos García, J.A., (2014), Use of 3D Scanning Technology for Automated Inspection of Tunnels, *World Tunnel Congress 2014 (WTC2014)*, Foz do Iguacu, Brasil, May 2014
- [Lax01] Lax, R., (2001), Keeping Europe Moving, European Union Road Federation. Available at: <http://www.publicservice.co.uk/featurestory.asp?id=1586>, consultado en julio de 2013
- [Lee04] Lee, B.J., y Lee, H.D, (2004), Position-Invariant Neural Network for Digital Pavement Crack Analysis, *Computer-Aided Civil and Infrastructure Engineering*, 19(2):115-118, doi: 10.1111/j.1467-8667.2004.00341.x
- [Lep99] Lepert, P., Hildebrand, G. y Christensen, J., (1999), Development of models for economic evaluation of pavement maintenance: the PAV-ECO project: providing an efficient and socially acceptable road transport network, in *Roskilde (ed.)*, VI rapport, 96, ISBN: 87-90145-59-3

- [Liu02] Liu, Z., Suandi, S.A., Ohashi, T. y Ejima, T., (2002), Tunnel crack detection and classification system based on image processing, Proceedings of the SPIE 4664, Machine Vision Applications in Industrial Inspection X, 145, San Jose, California, USA, Mar 2002; doi: 10.1117/12.460191
- [Liu10] Liu, Y., Kong, I., Wang, X. and Jiang F., Research on Image Acquisition of Automatic Surface Vision Inspection Systems for Steel Sheet, *3^d International Conference on IEEE Advanced Computer Theory and Engineering (ICACTE2010)*, Chengdu, China, Aug 2010, 6:189–192, doi: 10.1109/ICACTE.2010.5579393
- [Mah91] Mahler, D.S., Kharoufa, Z.B., Wong, E.K. y Shaw, L.G., (1991), Pavement distress analysis using image processing techniques, *Computer-Aided Civil and Infrastructure Engineering*, 6(1):1–14, DOI: 10.1111/j.1467–8667.1991.tb00393.x.
- [Mal03] Malamas, E.L., Petrakis, E.G.M., Zervakis, M., Petit, L. y Legat, J.D., (2003). A Survey on Industrial Vision Systems, Applications and Tools, *Image and Vision Computing*, 21(2):171-188, doi: 10.1016/S0262-8856(02)00152-X
- [Mar10] Martins, L.A.O, Padua, F.L.C, y Almeida, P.E.M., (2010), Automatic detection of surface defects on rolled steel using Computer Vision and Artificial Neural Networks, *IECON 2010 – 36th Annual Conference on IEEE Industrial Electronics Society*, Glendale, Arizona, EE.UU, Nov 2010, 1081–1086, doi: 10.1109/IECON.2010.5675519
- [Mas12] Masci, J., Meier, U., Ciresan, D., Schmidhuber, J. y Fricout, G., (2012), Steel Defect Classification with Max-Pooling Convolutional Neural Networks, *The 2012 International Joint Conference on Neural Networks (IJCNN)*, Brisbane, Australia, Jun 2012, doi: 10.1109/IJCNN.2012.6252468
- [McC43] McCulloch, W.S., y Pitts, W., (1943), A logical calculus of the ideas imminent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4):115-133, ISSN: 0007-4985, doi: 10.1007/BF02478259

- [Med08] Medina, R., Gayubo, F., González-Rodrigo, L.M., Olmedo, D., Gómez-García-Bermejo, J., Zalama, E. y Perán, J.R. (2008), Surface Defects Detection on Rolled Steel Strips by Gabor Filters, *VISAPP 2008: Proceedings of the Third International Conference on Computer Vision Theory and Applications*, Funchal, Madeira, Portugal, Jan 22-25, 2008, 1:479-485, ISBN:978-989-8111-21-0
- [Med10] Medina, R., Gómez-García-Bermejo, J. y Zalama, E. (2010), Automated visual inspection of road surface cracks, in *Proceedings of the 27th International Symposium on Automation and Robotics in Construction, ISARC'10*, Bratislava, Slovakia, Jun 25-27, 2010, 155-164, ISBN:978-80-7399-974-2
- [Med11] Medina, R., Gayubo, F., González-Rodrigo, L.M., Olmedo, D., Gómez-García-Bermejo, J., Zalama, E. y Perán, J.R. (2011), Automated visual classification of frequent defects in flat steel coils, *The International Journal of Advanced Manufacturing Technology*, 57:1087-1097, doi: 10.1007/s00170-011-3352-0
- [Med14] Medina, R., Llamas, J., Zalama, E. y Gómez-García-Bermejo, J., (2014), Enhanced Automatic Detection Of Road Surface Cracks By Combining 2D/3D Image Processing Techniques, *IEEE International Conference on Image Processing, ICIP2014*, CNIT La Défense, Paris, France, October 27-30, 2014, 778-782, doi: 10.1109/ICIP.2014.7025156
- [MER15] MERMEC Group, (2015), T-Sight 5000, <http://www.mermecgroup.com/inspection-technology/tunnel-e-clearance/82/1/t-sight-5000.php>, Consultado en Agosto de 2015
- [Mih11] Mihajlovic, D. y Jokanovic, I., (2011), Road Data Collection and Database Establishment in Developing Countries—Three Case Studies, *Transportation and Development Institute Congress 2011*, Chicago, Illinois, EE.UU, Mar. 2011, 1-10, doi: 10.1061/41167(398)1

- [Mio13] Mioulet, L., Breckon, T.P., Mouton, A., Liang, H y Morie, T., (2013), Gabor features for real-time road environment classification, *IEEE International Conference on Industrial Technology (ICIT2013)*, Cape Town, South Africa, Feb 25-28, 2013, 1117-1121, doi: 10.1109/ICIT.2013.6505829
- [Miy07] Miyamoto, A., Konno, M. y Bruhwiler, E., (2007), Automatic crack recognition system for concrete structures using image processing approach, *Asian Journal of Information Technology*, 6(5):553–561, doi: ajit.2007.553.561
- [MOP89] MOPU, (1989), Catálogo de deterioros en firmes, *Dirección General de Carreteras del Ministerio de Obras Públicas y Urbanismo*, ISBN: 978-84-7433-600-9
- [Mue00] Muehlemann, M., (2000). Standardizing defect detection for the surface inspection of large web steel. *Illumination Technology white paper*.
- [Mül12] Müller, U., Kuhn, P. y Kadner, G., (2012), Crack detection: Potential savings in motion, *World Tunnelling*, Marzo 2012, 25-27
- [Ngu09] Nguyen, T.S., Avila, M. y Begot, S., (2009), Automatic detection and classification of defect on road pavement using anisotropy measure, *17th European Signal Processing Conference (EUSIPCO 2009)*, Glasgow, Scotland, Aug 24-28, 2009, 617–621
- [Nis12] Nishikawa, T., Yoshida, J., Sugiyama, T. y Fujino, Y., (2012), Concrete crack detection by multiple sequential image filtering, *Computer-Aided Civil and Infrastructure Engineering*, 27(1):29–47, doi: 10.1111/j.1467–8667.2011.00716.x.
- [Nyq28] Nyquist, H., (1928), Certain topics in telegraph transmission theory, *Transactions of the American Institute of Electrical Engineers*, 47(2):617-644, doi: 10.1109/T-AIEE.1928.5055024

- [OEC11] OECD, (2011), Trends in the Transport Sector 2011, *OECD Publishing*, Paris, France, ISBN: 978-92-821-0359-3, doi: 10.1787/trend transp-2011-en
- [Ohm88] Ohmann, C., Yang, Q., Kunneke, M., Stolzing, H., Thon, K. y Lorenz, W., (1988), Bayes theorem and conditional dependence of symptoms: different models applied to data of upper gastrointestinal bleeding, *Methods of Information in Medicine*, 27(2):73-83
- [Oli08] Oliveira, H. y Correia, P. L. (2008), Identifying and retrieving distress images from road pavement surveys, *15th IEEE International Conference on Image Processing, ICIP 2008*, San Diego, USA, Oct 2008, 57–60, DOI: 10.1109/ICIP.2008.4711690
- [Oli09a] Oliveira, H. y Correia, P. L. (2009a), Automatic road crack segmentation using entropy and image dynamic thresholding, *Proceedings of the 17th European Signal Processing Conference (EUSIPCO'09)*, Glasgow, United Kingdom, Aug 24-28, 2009, 622–626
- [Oli09b] Oliveira, H. y Correia, P.L., (2009), *Supervised Crack Detection and Classification in Images of Road Pavement Flexible Surfaces, Recent Advances in Signal Processing*, Ashraf A Zaher (Ed.), ISBN: 978-953-307-002-5, InTech, DOI: 10.5772/7448.
- [Olm08] Olmedo, D., Gayubo, F., González, J.L., Medina, R., González, L.M., Gómez-García-Bermejo, J., Zalama, E. y Perán, J.R., (2008), Detección automática en línea de defectos superficiales en chapa laminada lubricada mediante técnicas de visión artificial, *XVII Congreso de Máquinas-Herramienta y Tecnologías de Fabricación*, San Sebastián, España, Nov. 2008
- [Olm10] Olmedo, D., Gayubo, F., González, J.L., Medina, R., González L.M., Gómez-García-Bermejo J., Zalama, E. y Perán, J.R., (2010), Detección Rápida de Regiones Uniformes en Chapa Laminada para Control de Calidad, *XVIII Congreso de Máquinas-Herramienta y Tecnologías de Fabricación*, San Sebastian, España, Nov. 2010

- [Ots79] Otsu, N., (1979), A Threshold Selection Method from Gray-Level Histogram, *IEEE Transaction on Systems, Man, and Cybernetics*, 9(1): 62-66, doi: 10.1109/TSMC.1979.4310076
- [Paa06] Paar, G., Caballo-Perucha, M.P., Kontrus, H. y Sidla, O., (2006), Optical crack following on tunnel surfaces, *Proceedings of the SPIE 6382, Two- and Three-Dimensional Methods for Inspection and Metrology IV, 638207*, Boston, Massachusetts, USA, Oct 2006, doi: 10.1117/12.685987
- [Par82] Parker, D.B., (1982), *Learning logic*, Invention Report S81-64, File 1, Office of Technology Licensing, Stanford University, Stanford, CA.
- [PAV15] Pavemetrics, (2015), Laser Road Imaging System (LRIS), <http://www.pavemetrics.com/en/lris.html>, consultado en Marzo de 2015
- [PAV15b] Pavemetrics, (2015), Laser Tunnel Scanning System (LTSS), <http://www.pavemetrics.com/en/ltss.html>, consultado en Agosto de 2015
- [Pen09] Peng, K. y Zhang, X., (2009), Classification Technology for Automatic Surface Defects Detection of Steel Strip Based on Improved BP Algorithm, *IEEE Fifth International Conference on Natural Computation, (ICNC2009)*, Tianjin, China, Aug 2009, 110–114, doi: 10.1109/ICNC.2009.487
- [Per03] Pernkopf, F. y O’Leary, P., (2003). Image Acquisition Techniques for Automatic Visual Inspection of Metallic Surfaces, *NDT&E International*, 36(8):609–617, doi: 10.1016/S0963-8695(03)00081-1
- [Per04] Pernkopf, F., (2004), Detection of Surface Defects on Raw Steel Blocks Using Bayesian Network Classifiers, *Pattern Analysis and Applications*, 7(3):333–342, doi: 10.1007/s10044-004-0232-3
- [Pin03] Pinto, J., (2003), *Caracterização Superficial de Pavimentos Rodoviários*, Tesis de Doctorado, Facultad de Ingeniería, Universidad de Oporto, Portugal

- [Piu05] Piuri, V., Scotti, F. y Roveri, M., (2005), Computational intelligence in industrial quality control, *IEEE International Workshop on Intelligent Signal Processing*, Algarve, Portugal, Sep. 2005, 4–9, doi: 10.1109/WISP.2005.1531623
- [Pit47] Pitts, W. y McCulloch, W.S., How we know universals the perception of auditory and visual forms, *The Bulletin of Mathematical Biophysics*, 9(3):127-147, ISSN: 0007-4985, doi: 10.1007/BF02478291
- [Pyn99] Pynn, J., Wright, A. y Lodge, R., (1999), Automatic identification of cracks in road surfaces, *Seventh International Conference on Image Processing and Its Applications*, Manchester, Reino Unido, Jul. 1999, 2:671-675, doi: 10.1049/cp:19990408
- [Qin08] Qingquan, L. y Xianglong, L., (2008), Novel Approach to Pavement Image Segmentation Based on Neighboring Difference Histogram Method, *IEEE International Congress on Image and Signal Processing (CISP2008)*, Sanya, China, May 2008, 2:792-796, doi: 10.1109/CISP.2008.13
- [Ria13] Riaz, F., Hassan, A., Rehman, S. y Qamar, U., (2013), Texture Classification Using Rotation- and Scale-Invariant Gabor Texture Features, *IEEE Signal Processing Letters*, 20(6):607-610, doi: 10.1109/LSP.2013.2259622
- [Riv15] Rivera, J.P., Josipovic, G., Lejeune, E., Luna, B.N. y Whittaker, A.S., (2015), Automated Detection and Measurement of Cracks in Reinforced Concrete Components, *ACI structural Journal*, 112(3):397-406, doi: 10.14359/51687424
- [Rol96] Roli, F., (1996), Measure of texture anisotropy for crack detection on textured surfaces, *Electronics Letters*, 32(14):1274-1275, doi: 10.1049/el:19960872
- [Ros62] Rosenblatt, F., (1962), *Principles of neurodynamics: perceptrons and the theory of brain mechanisms*, Washington, Spartan Books

- [Rum86] Rumelhart, D.E., Hinton, G.E., y Williams, R.J., (1986), Learning internal representations by error propagation. *In Parallel distributed processing: explorations in the microstructure of cognition*, 1:318-362, ISBN: 0-262-68053-X
- [Sas07] Sasaki, T., Takada, H. y Tomura, Y., (2007), Automatic Surface Inspection System for Tin Mill Black Plate(TMBP), *JFE technical report*, 9:60-63
- [Seu01] Seulin, R., Merienne, F. y Gorria, P., (2001), Dynamic lighting system for specular surface inspection, *Proceedings of SPIE, Machine Vision Applications in Industrial Inspection IX*, San Jose, California, EE.UU, Ene. 2001, 4301:199-206; doi: 10.1117/12.420913
- [Sha49] Shannon, C.E., (1949), Communication in the presence of Noise, *Proceedings of the Investigative Reporters and Editors*, 37(1):10-21, doi: 10.1109/JRPROC.1949.232969
- [Sha08] Sharifzadeh, M., Alirezaee, S., Amirfattahi, R. y Sadri, S., (2008) *Detection of steel defect using the image processing algorithms. IEEE International Multitopic Conference, INMIC2008*, Karachi, Pakistan, dec 23-24, 2008, 125–127, doi: 10.1109/INMIC.2008.4777721
- [She04] Shen, L. y Bai, L., (2004), Adaboost Gabor feature selection for classification, *Proceedings of Image and Vision Computing*, New Zealand, 77–83
- [Sjö95] Sjögren, L., (1995), The Laser RST PAVUE system, a pavement image acquisition and analysis system, *In Proc., 20th World Road Congress*, World Road Association, Montreal, Canada, Sep. 1995
- [Sta94] Stauffer, D., (1994), Introduction to Percolation Theory. 2nd edition, Taylor & Francis, New York, ISBN: 978-0748402533

- [Sto97] Storn, R. y Price, K., (1997), Differential Evolution – A Simple and Efficient Heuristic for global Optimization over Continuous Spaces, *Journal of Global Optimization*, 11(4):341-359, doi: 10.1023/A:1008202821328
- [Suv12] Suvdaa, B. Ahn, J. y Ko, J., (2012), Surface defects detection and classification using SIFT and voting strategy, *International Journal of Software Engineering and Its Applications*, 6(2):168–170, ISSN: 1738-9984
- [Ter15] Terra Vermessungen, (2015) tCrack: Crack detection, <http://www.terra.ch/en/crack-detection.html>, consultado en agosto de 2015
- [Tod94] Todd, B.S. y Stamper, R., (1994), The relative accuracy of a variety of medical diagnostic programs, *Methods of Information in Medicine*, 33(4):402-16
- [Tur86] Turner, M.R., (1986), Texture discrimination by Gabor functions, *Biological Cybernetics*, 55(2-3):71-82, ISSN: 0340-1200, doi: 10.1007/BF00341922
- [Val82] De Valois, R.L., Albrecht, D.G., y Thorell, L.G., (1982), Spatial-frequency selectivity of cells in macaque visual cortex, *Vision Research*, 22(5):545–559, doi: 10.1016/0042-6989(82)90113-4
- [Wan00] Wang, K.C.P., (2000), Design and implementation of automated systems for pavement surface distress survey, *ASCE Journal of Infrastructure Systems*, 6(1):24–32, doi: 10.1061/(ASCE)1076-0342(2000)6:1(24)
- [Wan05] Wang, K.C.P., y Gong, W., (2005), Real-Time Automated Survey System of Pavement Cracking in Parallel Environment, *ASCE Journal of Infrastructure Systems*, 11(3):154-164, doi: 10.1061/(ASCE)1076-0342(2005)11:3(154)

- [Wei08] Weiwei, L., Yunhui, Y., Jun, L., Yao, Z. y Hongwei, S., (2008), Automated On-Line Fast Detection for Surface Defect of Steel Strip Based on Multivariate Discriminant Function, *IEEE Second International Symposium on Intelligent Information Technology Application, IITA2008*, Shanghai, China, Dic 2008, 2:493–497, doi: 10.1109/IITA.2008.67
- [Wer74] Werbos, P.J., (1974). *Beyond regression: New tools for prediction and analysis in the behavioral sciences*. PhD thesis, Harvard University.
- [Wor14] *World Steel Association*, www.worldsteel.org, consultada en diciembre de 2014.
- [Wu08] Wu, G., Kwak, H., Jang, S., Xu, K. y Xu, J., (2008), Design of Online Surface Inspection System of Hot Rolled strips, *Proceedings of the IEEE International Conference on Automation and Logistics (ICAL2008)*, Qingdao, China, Sep 2008, 2291–2295, doi: 10.1109/ICAL.2008.4636548
- [Xie08] Xie, X., Dai, Q., Lam, K.M. Zhao, H., (2008), Efficient rotation- and scale-invariant texture classification method based on Gabor wavelets, *Journal of Electronic Imaging*, 17(4), 043026, doi: 10.1117/1.3050071
- [Yam10] Yamaguchi, T. y Hashimoto, S., (2010), Fast crack detection method for large-size concrete surface images using percolation-based image processing, *Machine Vision and Applications*, 21:797–809, doi: 10.1007/s00138-009-0189-8
- [Yan10] Yang, Y., Li, Q. Chen, P. y Ang, X.Z., (2010), Strip Surface Defect Detection Algorithm Based on Background Difference, *Second Pacific-Asia Conference on IEEE Circuits, Communications and System (PACCS2010)*, Beijing, China, Aug 2010, 23–26, doi: 10.1109/PACCS.2010.5626901

- [Yaz08] Yazdchi, M.R., Mahyari, A.G. y Nazeri, A., (2008), Detection and Classification of Surface Defects of Cold Rolling Mill Steel Using Morphology and Neural Network, *2008 International Conference on Computational Intelligence for Modelling Control & Automation*, Viena, Austria, Dic 2008, 1071–1076, doi: 10.1109/CIMCA.2008.130
- [Yun11] Yunhui, Y., Kechen, S., Zhitao, X. y Xuehui, F., (2011), The Strip Steel Surface Defects Classification Method Based on Weak Classifier Adaptive Enhancement, *IEEE Third International Conference on Measuring Technology and Mechatronics Automation, (ICMTMA2011)*, Shanghai, China, Jan 2011, 958–961, doi: 10.1109/ICMTMA.2011.811
- [Zal11] Zalama, E., Gómez-García-Bermejo, J., Llamas, J. y Medina, R., (2011), An Effective Texture Mapping Approach for 3D Models Obtained from Laser Scanner Data to Building Documentation, *Computer-Aided Civil and Infrastructure Engineering*, 26(5):381–392, doi: 10.1111/j.1467-8667.2010.00699.x
- [Zal14] Zalama, E., Gómez-García-Bermejo, J., Medina, R. y Llamas, J. (2014), Road Crack Detection Using Visual Features Extracted by Gabor Filters, *Computer-Aided Civil and Infrastructure Engineering*, 29(5):342–358, doi: 10.1111/mice.12042
- [Zha02] Zhang, J., Tan, T., Ma, L., (2002), Rotation invariant texture segmentation via circular Gabor filters, *16th International Conference on Pattern Recognition*, Quebec, Canada, Aug. 2002, 2:901-904, doi: 10.1109/ICPR.2002.1048450
- [Zha06] Zhang, X., Krewet, C., Kuhlentötter, B., (2006), Automatic classification of defects on the product surface in grinding and polishing, *International Journal of Machine Tools and Manufacture*, 46(1):59-69, doi: 10.1016/j.ijmachtools.2005.03.013
- [Zha12] Zhang, X. y Gao, H., (2012), Determining an optimal maintenance period for infrastructure systems, *Computer-Aided Civil and Infrastructure Engineering*, 27(7):543–54, doi:10.1111/j.1467-8667.2011.00739.x

- [Zha14] Zhang, W., Zhang, Z., Qi, D. y Liu, Y., (2014), Automatic Crack Detection and Classification Method for Subway Tunnel Safety Monitoring, *Sensors (Basel, Switzerland)*, 14(10):19307-19328, doi: 10.3390/s141019307.
- [Zha10] Zhao, J., Yang, Y. y Li, G., (2010), The Cold Rolling Strip Surface Defect On-Line Inspection System Based on Machine Vision, *Second Pacific-Asia Conference on IEEE Circuits, Communications and System (PACCS2010)*, Beijing, China, Aug 2010, 402–405, doi: 10.1109/PACCS.2010.5626924
- [Zha15] Zhao, G., Wang, T. y Ye, J., (2015), Anisotropic clustering on surfaces for crack extraction, *Machine Vision and Applications*, 26(5):675-688, doi: 10.1007/s00138-015-0682-1
- [Zha09] Zhaoyun, S., Chaofan, W. y Aimin, S. (2009), Study of image-based pavement cracking measurement techniques, *Proceedings of the 9th International Conference on Measurement & Instruments (ICEMI09)*, Beijing, China, aug 2009, 2:140–43, doi: 10.1109/ICEMI.2009.5274633
- [Zou12] Zou, Q., Cao, Y., Li, Q., Mao, Q. y Wang, S., (2012), CrackTree: automatic crack detection from pavement images, *Pattern Recognition Letters*, 33(3):227–38, doi: 10.1016/j.patrec.2011.11.004