

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



PROYECTO FIN DE CARRERA

**"Detección de falsos positivos en sistemas de
videovigilancia debidos a cambios bruscos en la
iluminación de la escena"**

Sara Rodríguez González

Diciembre 2010

**"Detección de falsos positivos en sistemas de
videovigilancia debidos a cambios bruscos en la
iluminación de la escena"**

**AUTOR: Sara Rodríguez González
TUTOR: Manuel A. Sánchez-Montañés Isla**

Grupo de Neurocomputación Biológica (GNB)

**Dpto. de Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid**

Noviembre 2010

Palabras Clave

Videovigilancia, Segmentación, Detección de Movimiento, Sustracción de fondo, Modelo de fondo, *Foreground*, *Background*, Falsos Positivos, Cambios de Iluminación, Detección de intrusos.

Resumen

Este trabajo ha sido realizado en el contexto de INTEGRA, proyecto financiado por el CDTI dentro del programa CENIT. El Programa CENIT, cuyas siglas corresponden a "Consortios Estratégicos Nacionales en Investigación Técnica", contempla la financiación de grandes proyectos integrados de investigación industrial de carácter estratégico, gran dimensión y largo alcance científico-técnico orientados a una investigación planificada en áreas tecnológicas de futuro y con potencial proyección internacional.

El objetivo principal será reducir las falsas alarmas que ocasionan las variaciones de iluminación en sistemas de videovigilancia basados en detección de movimiento utilizados para la detección de intrusos, teniendo en cuenta que la solución que se proponga no puede afectar a dicha detección de intrusos u otras situaciones consideradas como alarmas verdaderas. Además se debe tener en cuenta la necesidad de encontrar soluciones de bajo coste computacional para poder aplicarlas en tiempo real, ya que el fin último será integrarlas en un sistema con aplicaciones comerciales en el campo de la seguridad.

Con este fin, se han estudiado, implementado y evaluado diferentes algoritmos aplicados a secuencias de imágenes provenientes de escenarios tanto interiores (*indoor*) como exteriores (*outdoor*), y en diferentes condiciones de iluminación. Para ello se ha hecho uso de herramientas para el procesamiento y segmentación de video basadas en técnicas de sustracción de fondo (*Background Subtraction*).

Los experimentos llevados a cabo, así como las pruebas de validación de resultados han arrojado cifras muy favorables, habiéndose reducido la tasa de falsos positivos en un 100% para secuencias capturadas en condiciones de iluminación alta o media y en un 70% para secuencias capturadas en condiciones de baja iluminación o en condiciones nocturnas, manteniendo la tasa de detección de verdaderos positivos en un 100%.

Algunos de los resultados obtenidos en este proyecto serán convenientemente integrados en un sistema comercial. Con el fin de proteger la investigación realizada se ha solicitado una patente con el título "Método y dispositivo de detección de cambio de iluminación para sistemas de visión" y número de solicitud *EP 10382281.3*.

Como resultado este proyecto ha contribuido a la mejora de los sistemas actuales de videovigilancia basados en detección de movimiento.

Key Words

Video Surveillance, Video Segmentation, Motion Detection System, Background Subtraction, Foreground, Background, False Positive, Lighting changes, Human Intrusion Detection.

Abstract

This work was carried out in the context of INTEGRA (Research in technologies for migration management), a project financed by the CDTI, which is part of the CENIT program. The CENIT (National Strategic Consortiums for Technological Research) program seeks to finance large-scale, broad-based strategic industrial research projects in technological areas of the future with potential international prominence.

The main objective of this Master Thesis (PFC) is to reduce false alarms caused by lighting changes in motion detection based surveillance systems focused on intrusion detection, considering that the proposed solution can not affect to the above intrusion detection process. It is also necessary to find out low computational cost methods for real-time applications, in order to be able to integrate them into a commercial system for security applications.

For this purpose, we have studied, implemented and evaluated different algorithms applied to images from both indoor and outdoor scenarios in different lighting conditions. Background Subtraction techniques have been used for image processing and video segmentation.

The experiments executed as well as the cross validation tests of the results reported favorable results, having reduced the rate of false positives to 100% for sequences captured in high or medium light conditions and to 70% for those captured in low light/night conditions, maintaining a 100% of true positive rate detection.

Some of achieved results in this project will be conveniently integrated in a commercial system. In order to protect the executed research a patent, with application number EP 10382281.3. and intitled: Method and device to detect lighting changes for vision system, has been applied for.

As a result, this project has contributed to the improvement of existing video surveillance systems based on motion detection.

Agradecimientos

Manuel Sánchez-Montañés: Gracias por darme la oportunidad de realizar este proyecto y haber confiado en mí, por la ayuda y por el apoyo prestado durante estos meses. **Luis Fernando Lago:** Gracias por toda tu colaboración, tus propuestas y tu ayuda para que este proyecto haya llegado a buen puerto. **EPS:** Gracias a todos los profesores, porque cada uno me enseñó algo y de algunos aprendí mucho (incluso de mí misma).

Eduardo Cermeño: Gracias por darme la oportunidad de colaborar con vosotros de forma tan cercana. Gracias por los consejos, las “lecciones magistrales”, por tratarme tan bien y por ofrecerme tu amistad. También por las croquetas...**Ana, Silvana, Alba y Lucía:** Gracias por esas tardes geniales de “transferencia de conocimientos varios”, incluido conocimientos de MATLAB ;) **Rafi, Jorge, Rafa, Carlos y José:** Gracias por hacerme reír tanto, por hacer más agradables las tardes y por recordarme “mis frases de coronación”. Mil gracias a todos!!!

Ana, Yera, Rocío y María: Gracias por caminar a mi lado todos estos años y por quererme “defectuosa”, me he sentido muy afortunada a vuestro lado.

Almu: Gracias por tu generosidad, por tu compañía, por tu amistad y por tirar de mí en los momentos malos. Ha sido genial encontrarte en el camino de la vida! **Borja:** Gracias por ser un gran compañero de clase, de prácticas, de viajes, de salidas nocturnas, de erasmus, de trabajo y sobre todo por ser un gran amigo. **Ali :** Gracias por las miles de horas que hemos pasado juntas, por estar ahí y por seguir ahí. **Alberto:** Gracias por alegrar mis mañanas café en mano y por preocuparte por mí (aunque te metas con Rafa...). **Bader:** MUCHAS gracias por MUCHAS cosas, sobre todo por ser así. **Isa:** Gracias por esos momentos “Ya no hay tortuga, ni caracola” que recuerdan lo genial que puede ser la vida. **Julita:** Gracias por escucharme SIEMPRE (incluyendo monólogos y tonterías). **Álvaro:** Gracias por todo lo que nos hemos reído, lo que hemos “discutido”, lo que hemos hablado y lo que no. **Sara:** Gracias por ver la película de mi vida...y varias veces. **David:** Gracias por todos los exámenes que hemos pasado juntos, porque algunos provocaban la risa y otros el llanto, pero los recuerdo todos. **Edu:** Gracias por todo lo que me has hecho reír (a pesar de nuestros comienzos...). **Clau:** Gracias por las meriendas-cenas de biblioteca donde arreglábamos el mundo. **Alex:** Gracias por...la tortilla de patata y la hamburguesa de “La Manga”... **Silvia:** Gracias por contagiarme la risa hasta el dolor de tripa. **Ramiro, Alfonso, Laura y Luisito:** Gracias por esas reuniones que a veces se convierten en el “oktoberfest de mayo”. **Luis:** Gracias por que un día decidiste estudiar teleco, llegar con tu violín y tu moto, y así pasar por mi vida aunque fuera poco tiempo. Muchas gracias a todos por acompañarme hasta aquí con la mejor de las sonrisas.

Laura, Tania y Alex: Gracias por estar ahí desde siempre y por conseguir que parezca que nos vemos todos los días.

Paola, Guada, Bea y Sara: Gracias por convertirnos en mi familia y por todo lo que vivimos durante ese gran año. LUNGARNO GIRLS!!

Vito y Loly: Gracias por todo lo que me habéis dado, lo que me habéis enseñado y lo que hacéis por mí cada día.

Familia: Gracias a mis abuelos por ser ABUELOS. Gracias a mis tíos que me han dado siempre buenos consejos y de los que he aprendido mucho. Gracias por vuestra generosidad. Gracias a mis primos y mis PRIMAS, porque yo nunca he sido hija única. **Antonio:** Gracias por haber sido ese hermano mayor que toda familia quisiera tener. **Antonio:** Gracias Abuelito por creer en mí más que nadie. **Manuel:** Gracias por todas las conversaciones, por preocuparte, por “los helados”, por cuidarme toda la vida...De todas las personas aquí citadas sé que serías el más contento al ver este proyecto. Gracias padrino.

Índice de contenido

Palabras Clave.....	I
Resumen.....	I
Key Words.....	II
Abstract.....	II
Agradecimientos.....	IV
1 Introducción.....	1
1.1 Motivación.....	1
1.2 Objetivos.....	2
1.3 Organización de la memoria.....	3
2 Estado del arte.....	4
2.1 Análisis de los sistemas de videovigilancia.....	4
2.1.1 Introducción.....	4
2.1.2 Evolución tecnológica de los sistemas de videovigilancia.....	5
2.1.3 Clasificación funcional de los sistemas de videovigilancia.....	7
2.2 Técnicas de análisis y segmentación de videos.....	9
2.2.1 Segmentación mediante técnicas de Background Subtraction.....	9
2.2.2 Modelado de fondo.....	10
2.2.2.2 Modelos que estiman una distribución de probabilidad.....	13
2.2.3 Detección de frente.....	15
2.3 Limitaciones de los sistemas de videovigilancia automáticos basados en detección de movimiento.....	16
2.3.1 Tipología de los problemas derivados de la etapa de segmentación.....	16
2.3.2 Aproximaciones existentes al problema de la iluminación.....	18
2.3.3 Propuestas actuales para la reducción de los problemas debidos a cambios de iluminación.....	19
3 Detección de intrusos y problemas debidos a los cambios de iluminación en sistemas de videovigilancia.....	21
3.1 Aplicaciones en las que se enfocará este proyecto.....	21
3.2 Sistemas de videovigilancia basados en detección de movimiento.....	22
3.3 Generación de alarmas en sistemas basados en detección de movimiento.....	23
3.4 Efectos de los cambios de iluminación y falsas alarmas debidas a estos efectos negativos.....	24
3.4.1 Efectos en escenarios Indoor.....	25
3.4.2 Efectos en escenarios Outdoor.....	27
3.4.2.1 Grabaciones diurnas.....	28
3.4.2.2 Grabaciones nocturnas.....	30
3.5 Conclusiones.....	35
4 Diseño.....	36
4.1 Etapa preliminar.....	36
4.1.1 Segmentación de vídeos.....	36
4.1.2 Elección del modelo de color.....	37
4.1.3 Elección de herramientas de desarrollo.....	38
4.2 Estrategias de diseño.....	39
4.2.1 Medida de Dependencias: Correlación.....	39
4.2.2 Arquitectura del sistema.....	40
4.3 Diseño de algoritmos base.....	42

4.3.1	Básico.....	42
4.3.2	En cuadrantes.....	42
4.3.3	En bounding boxes completos.....	43
4.3.4	En labeling píxeles dentro del bounding box.....	45
4.3.5	Adición de un tiempo de guarda antes de la generación de alarma.....	46
4.4	Diseño de algoritmos mejorados para condiciones de baja iluminación o nocturnidad.....	47
4.4.1	Estudio de otras regiones de cálculo.....	47
4.4.2	Eliminar gradiente de iluminación (versión1).....	47
4.4.3	Eliminar gradiente de iluminación (versión2).....	48
4.4.4	Eliminar gradiente de iluminación en el bounding box (versión1).....	48
4.4.5	Eliminar gradiente de iluminación en el bounding box (versión2).....	49
4.4.6	Eliminación de píxeles con valores saturados a blancos dentro del bounding box.....	49
4.4.7	Cambios de escala.....	50
4.4.8	Cálculo de intensidad media de los píxeles vecinos.....	51
4.4.9	Estudio de otros factores.....	51
5	Pruebas y resultados experimentales.....	53
5.1	Curvas ROC.....	53
5.1.1	Introducción.....	53
5.1.2	Clasificadores binarios.....	53
5.1.3	El espacio ROC.....	55
5.2	Diseño de pruebas experimentales.....	56
5.2.1	Procedimiento 1, clasificación del evento en cada frame de la secuencia (P1).....	56
5.2.2	Procedimiento 2, clasificación del evento en cada frame de la secuencia, analizando W frames sucesivos (P2).....	57
5.2.3	Procedimiento 3, clasificación del evento dada una secuencia completa (P3).....	58
5.2.4	Procedimiento 4, clasificación del evento dada una secuencia completa, analizando W frames sucesivos (P4).....	59
5.3	Resultados caso Indoor.....	60
5.3.1	Secuencias utilizadas para las pruebas.....	60
5.3.2	Curvas ROC.....	60
5.3.3	Curvas ROC comparativas.....	65
5.3.4	Tabla de clasificación.....	68
5.4	Resultados caso Outdoor diurno.....	69
5.4.1	Secuencias utilizadas para las pruebas.....	69
5.4.2	Curvas ROC.....	69
5.4.3	Curvas ROC comparativas.....	73
5.4.4	Tabla de clasificación.....	76
5.5	Resultados caso Outdoor nocturno.....	77
5.5.1	Secuencias utilizadas para las pruebas.....	77
5.5.2	Curvas ROC de diferentes métodos probados para condiciones nocturnas.....	77
5.5.3	Curvas ROC del método con resultados más favorables.....	84
5.5.4	Tabla de clasificación.....	87
5.6	Validación de Resultados.....	87
5.6.1	Métodos de validación cruzada.....	87
5.6.2	Diseño de pruebas para la validación de resultados.....	88
5.6.3	Validación de resultados del caso Indoor.....	90
5.6.4	Validación de resultados del caso Outdoor diurno.....	91
5.6.5	Validación de resultados del caso Outdoor nocturno.....	92

6 Conclusiones y trabajo futuro.....	93
6.1 Conclusiones.....	93
6.2 Trabajo futuro.....	95
Referencias.....	I
ANEXO A: Presupuesto	V
ANEXO B: Pliego de Condiciones	VI
ANEXO C: Acreditación de méritos.....	XI

Índice de ilustraciones

Figura 1-1: Ejemplo de detección de intruso en un área videovigilada [1].....	1
Figura 2-1: Sistema CCTV con VCR.....	5
Figura 2-2: Sistema CCTV con DVR.....	5
Figura 2-3: Sistema CCTV con DVR de red.....	6
Figura 2-4: Sistemas de video IP con cámaras analógicas.....	6
Figura 2-5: Sistemas de video IP con cámaras IP.....	6
Figure 2-6: Videowall controlado por operador.....	7
Figure 2-7: Videowall de sistema con detección de movimiento.....	8
Figura 2-8: Diagrama de bloques de la técnica de sustracción de fondo. <i>t</i> : Tiempo. <i>N</i> : Número de imágenes. <i>B</i> : Imagen de fondo. <i>I</i> : Imagen actual [15].....	10
Figura 2-9: Modelado de fondo mediante MoG.....	14
Figura 2-10: Detector de frente basado en diferencia.....	16
Figura 2-11: Técnicas de post-procesado para la detección de sombras.....	17
Figura 2-12: Ejemplo de cambio de iluminación debido al paso de nubes.....	18
Figura 2-13: Detección de cambios de iluminación utilizando dos vistas.....	20
Figura 3-1: Arquitectura básica de un sistema de videovigilancia basado en detección de movimiento.....	22
Figura 3-2: Falsa alarma provocada por el movimiento de una telaraña situada delante del objetivo de la cámara.....	23
Figura 3-3: Muestra de las falsas alarmas provocadas por insectos.....	25
Figura 3-4: Respuesta del sistema ante cambio iluminación leve.....	26
Figura 3-5: Respuesta del sistema ante cambio iluminación brusco.....	27
Figura 3-6: Respuesta del sistema ante cambio iluminación muy brusco.....	28
Figura 3-7: Respuesta del sistema ante un caso de alarma real.....	28
Figura 3-8: Respuesta del sistema ante cambio iluminación natural.....	29
Figura 3-9: Respuesta del sistema ante la desaparición de una fuente luminosa.....	29
Figura 3-10: Respuesta del sistema ante un caso de alarma real.....	30
Figura 3-11: Escena grabada con una cámara sin iluminación infrarroja y con iluminación infrarroja [6].....	30
Figura 3-12: Ejemplo de cámara IR utilizada en escenarios exteriores.....	31
Figura 3-13: Efectos sobre la escena de la iluminación en movimiento.....	31
Figura 3-14: Cambio de la cámara a modo diurno por culpa de un alumbrado intenso. Generación de alarma de tipo “obstrucción”.....	33
Figura 3-15: Dificultades de detección de verdaderos positivos.....	34
Figura 3-16: Verdadero positivo con características similares a un falso positivo	

debido a un cambio de iluminación.....	35
Figura 4-1: Modelo de fondo e imagen del instante actual.....	36
Figura 4-2: Imágenes de salida del simulador.....	37
Figura 4-3: Gráfico Matlab con imagen original y los tres planos de color.....	38
Figura 4-4: Arquitectura del sistema.....	41
Figura 4-5: Imagen ejemplo del algoritmo base por cuadrantes.....	43
Figura 4-6: Dependencias entre el elemento detectado y el modelo de fondo.....	45
Figura 4-7: Diferencias entre el efecto provocado por cambios de iluminación y la aparición de intrusos en escena.....	46
Figura 4-8: Autoregulación de la cámara tras un cambio brusco de iluminación en escenario indoor.....	47
Figura 4-9: Imagen original e imagen tras la corrección del gradiente de iluminación.....	48
Figura 4-10: Imagen representativa del efecto de la iluminación de vehículos sobre el escenario y su histograma.....	50
Figura 4-11: Escala logarítmica y escala exponencial ($a=3$).....	51
Figura 4-12: Comparativa de masa, dimensiones y proporciones de un blob causante de falsa alarma con uno de alarma verdadera.....	52
Figura 5-1: Ejemplos de curvas ROC.....	55
Figura 5-2: Caso Indoor, curvas ROC P1.....	62
Figura 5-3: Caso Indoor, curvas ROC P2.....	62
Figura 5-4: Caso Indoor, curvas ROC P3.....	63
Figura 5-5: Caso Indoor, curvas ROC P3.....	63
Figura 5-6: Caso Indoor, curvas ROC P4.....	64
Figura 5-7: Caso Indoor, curvas ROC P4.....	64
Figura 5-8: Caso Indoor, curvas ROC comparativas P1 y p2.....	66
Figura 5-9: Caso Indoor, curvas ROC comparativas P3 y p4.....	67
Figura 5-10: Caso Outdoor diurno, curvas ROC P1.....	70
Figura 5-11: Caso Outdoor diurno, curvas ROC P2.....	70
Figura 5-12: Caso Outdoor diurno, curvas ROC P3.....	71
Figura 5-13: Caso Outdoor diurno, curvas ROC P3.....	71
Figura 5-14: Caso Outdoor diurno, curvas ROC P4.....	72
Figura 5-15: Caso Outdoor diurno, curvas ROC P4.....	72
Figura 5-16: Caso Outdoor diurno, curvas ROC comparativas P1 y P2.....	74
Figura 5-17: Caso Indoor, curvas ROC comparativas P3 y P4.....	75
Figura 5-18: Caso Outdoor nocturno. Sin correcciones.....	78
Figura 5-19: Caso Outdoor nocturno. Eliminación del gradiente de iluminación en el bounding box.....	79
Figura 5-20: Caso Outdoor nocturno. Eliminación de píxeles blancos.....	80
Figura 5-21: Caso Outdoor nocturno. Cambio de escala logarítmico.....	81
Figura 5-22: Caso Outdoor nocturno. Cambio de escala exponencial.....	82
Figura 5-23: Caso Outdoor nocturno. Cálculo de la intensidad media de los píxeles vecinos.....	83
Figura 5-24: Caso Outdoor nocturno, curvas ROC P1.....	84
Figura 5-25: Caso Outdoor nocturno, curvas ROC P2.....	85
Figura 5-26: Caso Outdoor nocturno, curvas ROC P3.....	85
Figura 5-27: Caso Outdoor nocturno, curvas ROC P4.....	86
Figura 5-28: Procedimiento para la validación de resultados.....	89
Figura 5-29: Validación de resultados para caso Indoor.....	90
Figura 5-30: Validación de resultados para caso Outdoor Diurno.....	91

Figura 5-31: Validación de resultados para caso Outdoor Nocturno.....	92
ANEXO C.....	XI
ANEXO C.....	XII

Índice de tablas

Tabla 1: Clasificación inicial de secuencias Indoor.....	60
Tabla 2: Clasificación final de secuencias Indoor.....	68
Tabla 3: Clasificación inicial de secuencias Outdoor diurnas.....	69
Tabla 4: Clasificación final de secuencias Outdoor diurnas.....	76
Tabla 5: Clasificación inicial de secuencias Outdoor nocturnas.....	77
Tabla 6: Clasificación final de secuencias Outdoor nocturnas.....	87
Tabla 7: Validación de las tasas de reducción de falsos positivos en secuencias Indoor nocturnas.....	90
Tabla 8: Validación de las tasas de reducción de falsos positivos en secuencias Outdoor diurnas.....	91
Tabla 9: Validación de las tasas de reducción de falsos positivos en secuencias Outdoor nocturnas.....	92

1 Introducción

1.1 Motivación

En los últimos años ha habido un desarrollo creciente en el campo de los sistemas inteligentes de videovigilancia. Estos logran detectar, reconocer y seguir objetos de manera automática a partir de secuencias de imágenes obtenidas por una cámara [1]. Este área tiene una amplia diversidad de potenciales aplicaciones como la seguridad en edificios, vigilancia de tráfico en ciudades y carreteras, control de fronteras, etc [2,3,4 y 5].

Una de las aplicaciones más importantes de estos sistemas es detectar intrusos en zonas de seguridad. Para ello, dichos sistemas suelen utilizar estrategias basadas en la detección de movimiento en la escena [6].



Figura 1-1: Ejemplo de técnicas de *background subtraction* para la detección de intruso en un área videovigilada [1]

Dado que dichas estrategias se basan en detectar cambios en la imagen, estos sistemas pueden generar falsos positivos en situaciones donde el cambio en la imagen se debe a cambios de iluminación y no a movimientos reales en la escena [7].

En la literatura se han propuesto métodos que utilizan modelos adaptativos de fondo [7,8] y otras técnicas de post-procesado de alto coste computacional [9,10], para tener en cuenta estos cambios de iluminación, pero en general fallan ante variaciones muy bruscas de luz como las provocadas por el on/off de luces, coches que pasan cerca de la escena con los faros encendidos, etc.

El problema que ocasionan los cambios de iluminación en los sistemas de videovigilancia, así como el estudio de las aproximaciones existentes a dicho problema ha sido lo que ha motivado la intención de proponer una solución aplicable en tiempo real, innovadora y eficiente, que consiga reducir la tasa de falsas alarmas que se generan en estos sistemas basados en detección de movimiento.

1.2 Objetivos

El principal objetivo de este PFC es la reducción de falsos positivos debidos a cambios de iluminación en los sistemas de videovigilancia automáticos basados en detección de movimiento para la detección de intrusos. Para ello se debe tener en cuenta que esta reducción no debe afectar a la detección de intrusos, siendo una prioridad mantener la tasa de verdaderos positivos al 100%.

En este proyecto se realizará un análisis del comportamiento de dichos sistemas frente a los cambios bruscos de iluminación. Se compararán las características de dichos cambios en función del escenario, las condiciones externas de iluminación y el tipo de fuente de luz que los provoca.

En este PFC se usará como punto de partida algunas de las técnicas y algoritmos implementados previamente en el sistema comercial de la compañía de visión artificial colaboradora. Estos métodos reducen sólo parte de las falsas alarmas debidas a variaciones lumínicas. El objetivo será mejorar las técnicas iniciales, de forma que se puedan eliminar o reducir considerablemente estos falsos positivos.

Para alcanzar los objetivos que se plantean en este proyecto, las tareas que se llevarán a cabo son las siguientes:

- **Estudio del estado del arte:** En primer lugar, se hará un breve resumen de la historia de los sistemas de videovigilancia, desde sus comienzos hasta los sistemas actuales. Se realizará un estudio del estado del arte de las diferentes técnicas de segmentación basadas en modelado de fondo que se utilizan en los sistemas de videovigilancia automáticos. Centraremos nuestra atención en la problemática derivada de esta etapa, que provocará falsas alarmas. Se estudiará el enfoque actual de las soluciones propuestas por diferentes autores para estos problemas; en particular para el problema derivado de los cambios de iluminación.
- **Estudio de la casuística:** Una vez estudiado el funcionamiento de los sistemas de videovigilancia inteligentes, pasaremos al estudio de diferentes situaciones en las que los cambios de iluminación generan falsas alarmas. De esta forma se intentará realizar una clasificación inicial de la casuística y se intentará comprender el funcionamiento del sistema ante estos casos. Ya con una visión general del problema real, lo desglosaremos en problemas de menor complejidad para poder abordarlo de forma sencilla y eficiente.
- **Extracción de características e implementación de algoritmos :** Una vez conocida la problemática, intentaremos extraer características relevantes que nos permitan diferenciar las falsas alarmas de las verdaderas. A partir de la información obtenida se pasará a la implementación de algoritmos que clasifiquen el mayor número de falsas alarmas como tales, reduciendo así la tasa de falsos positivos.

- **Adaptación de algoritmos a la casuística:** Con el conocimiento de ciertos resultados parciales, se intentará adaptar aquellos algoritmos que funcionan relativamente bien a cada uno de los subproblemas. El objetivo de esto será alcanzar tasas de reducción de falsos positivos más favorables.
- **Análisis de resultados y conclusiones:** Se analizarán los resultados obtenidos por medio de curvas ROC. A partir de ellos compararemos las diferentes propuestas. Se utilizará también el método de validación cruzada *Leave one out* para estimar el error de generalización en base a los resultados de los experimentos anteriores. Con estos datos se extraerán las conclusiones de las diferentes soluciones en las que se ha trabajado en este proyecto. En función de la reducción de la tasa de falsos positivos, así como en el comportamiento en tiempo real se decidirá si convendría integrar alguno de los métodos en un sistema comercial.

1.3 Organización de la memoria

La memoria consta de los siguientes capítulos:

- **Capítulo 1:** Introducción, motivación y objetivos del proyecto.
- **Capítulo 2:** Breve resumen de la evolución de los sistemas de videovigilancia. Estudio del estado del arte de los algoritmos de segmentación de video utilizados en sistemas de videovigilancia automáticos, las limitaciones de estos sistemas, y las soluciones propuestas en la literatura para los cambios de iluminación.
- **Capítulo 3:** Análisis del comportamiento de un sistema comercial frente a los cambios de iluminación, estudio de la casuística y desglose del problema.
- **Capítulo 4:** Diseño e implementación de algoritmos que permitan reducir la tasa de falsos positivos debidos a cambios de iluminación. Adición de mejoras a los algoritmos anteriores para alcanzar tasas más favorables.
- **Capítulo 5:** Pruebas, resultados experimentales y análisis comparativo entre los diferentes métodos.
- **Capítulo 6:** Conclusiones y trabajo futuro.

2 Estado del arte

2.1 Análisis de los sistemas de videovigilancia

2.1.1 Introducción

El término videovigilancia proviene de la traducción del inglés de “*video surveillance*”, y según la Real Academia Española de la lengua se define como: “Vigilancia a través de un sistema de cámaras, fijas o móviles”. Los sistemas de vigilancia por vídeo existen desde hace más de tres décadas, pero en los últimos años se ha avanzado mucho en este campo. Esto se debe a tres motivos principales; el desarrollo tecnológico, la demanda de mayores niveles de seguridad y el estudio de técnicas de análisis de vídeo.

Se ha avanzado mucho desde la aparición de las primeras cámaras analógicas con tubo conectadas a VCR (*video cassette recorder*) hasta los nuevos sistemas semiautomáticos. Las nuevas tecnologías y los avances de las últimas décadas han permitido conseguir sistemas con elevadas prestaciones. Las cámaras actuales tienen grandes prestaciones como alta resolución, zoom y una buena cantidad de lentes que nos permite tener una amplia panorámica de lo que estamos vigilando. Además, con suplementos como los infrarrojos, se puede tener visión nocturna [6].

En un principio, las cintas de grabación y los primeros DVR (*digital video recorder*) limitaban la duración de las grabaciones, lo que obligaba a utilizar un *frame rate* bajo. El enorme desarrollo de los sistemas de almacenamiento hace que actualmente el espacio para la grabación ya no suponga un problema. Otra de las prestaciones que se ofrece es el acceso remoto por parte del usuario, haciendo uso de la tecnología IP (*Internet Protocol*).

Los avances citados han venido acompañados de un importante abaratamiento de los sistemas, lo que ha provocado una continua subida de la demanda en seguridad [11]. Actualmente se instalan sistemas de videoseguridad en entornos tradicionalmente protegidos como bases militares, edificios gubernamentales, aeropuertos, estaciones, bancos, etc; pero también en otros entornos como vías públicas, estadios, instalaciones industriales, oficinas, todo tipo de comercios y establecimientos, hogares y otros recintos privados [12].

Este aumento de la variedad de escenarios ha generado la necesidad de mejorar los sistemas y adaptarlos a las necesidades específicas de cada uno de ellos. Esto ha motivado el estudio e investigación de técnicas de análisis de video, visión artificial y automatización de sistemas para que sean capaces de detectar situaciones relevantes de forma automática. Las áreas claves de estudio son la detección y seguimiento de objetos, identificación de personas, identificación de eventos y reconocimiento de patrones de comportamiento. El principal objetivo de este campo de investigación ha sido dotar de inteligencia a estos sistemas, de forma que sean capaces de interpretar automáticamente las escenas [1].

2.1.2 Evolución tecnológica de los sistemas de videovigilancia

Como se ha comentado en el apartado anterior la evolución tecnológica ha sido el principal motor del desarrollo de los sistemas de videovigilancia, siendo las aplicaciones de seguridad las de mayor expansión comercial.

Los primeros sistemas de CCTV (circuito cerrado de televisión) eran 100% analógicos, siendo en la actualidad completamente digitales. Sin embargo, entre los sistemas completamente analógicos y los completamente digitales existen diversas soluciones parcialmente digitales [6].

2.1.2.1 Sistemas completamente analógicos

Sistemas de circuito cerrado de TV analógicos usando VCR

Se trata de sistemas CCTV con cámaras analógicas que se conectan mediante cables coaxiales y utilizan un VCR (grabador de *video cassette*). El video no se comprime y por tanto la duración de la grabación será muy limitada.

Se puede hacer uso de un multiplexor entre la cámara y el VCR, que permite grabar videos de diferentes cámaras. La monitorización del video se hará en un monitor analógico.

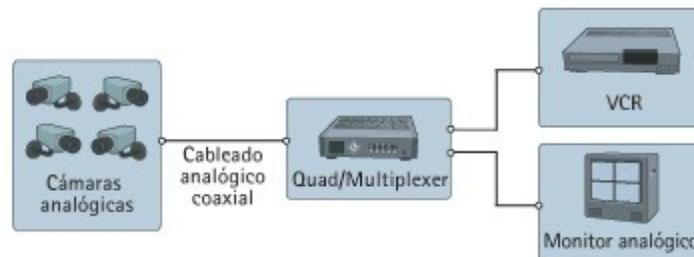


Figura 2-1: Sistema CCTV con VCR [6]

2.1.2.2 Sistemas parcialmente digitales

Sistemas de circuito cerrado de TV analógicos usando DVR

Se trata de un sistema analógico con grabación digital. Para la grabación de vídeo en los discos duros es necesario que el vídeo se digitalice y comprima para almacenar la máxima cantidad de imágenes. El DVR, al tener varias entradas, puede actuar a la vez de multiplexor.

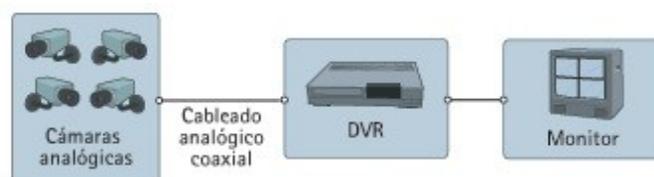


Figura 2-2: Sistema CCTV con DVR [6]

Sistemas de circuito cerrado de TV analógicos usando DVR de red

En este caso además de todos los procesos citados en el apartado anterior, el DVR IP tiene un puerto Ethernet para conectividad de red. Como el vídeo se digitaliza y comprime, se puede transmitir a través de una red y así se puede monitorizar en un PC remoto. En algunos sistemas de este tipo se puede monitorizar tanto video grabado como en directo.

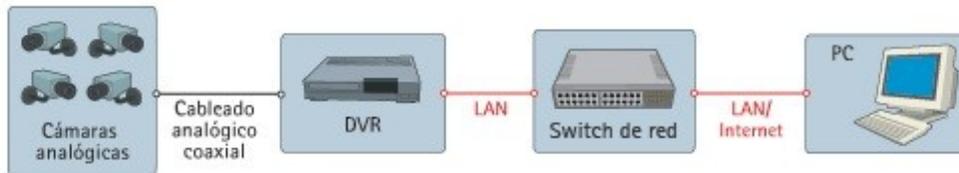


Figura 2-3: Sistema CCTV con DVR de red [6]

2.1.2.3 Sistemas completamente digitales

Sistemas de vídeo IP que utilizan servidores de vídeo

En este caso se utiliza un servidor de vídeo, un conmutador de red y un PC con software de gestión de vídeo. El vídeo se digitaliza y se transmite a través de un conmutador de red a un PC, donde se almacena. Se trata de un sistema escalable y que puede ampliarse con facilidad.

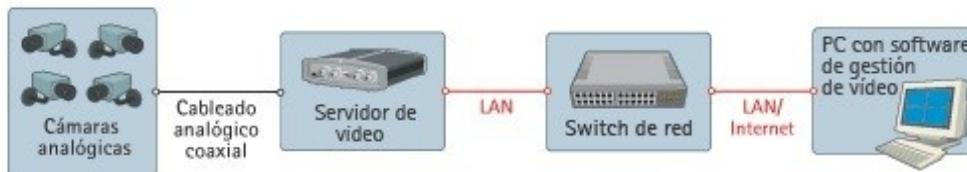


Figura 2-4: Sistemas de vídeo IP con cámaras analógicas [6]

Sistemas de vídeo IP que utilizan cámaras IP

Una cámara IP combina una cámara y un ordenador en una unidad, lo que incluye la digitalización y la compresión del vídeo así como un conector de red. El vídeo se transmite a través de una red IP, mediante los conmutadores de red. Este tipo de sistema tiene numerosas ventajas como alta resolución y calidad de imagen, funcionalidad inalámbrica, funciones de movimiento de la cámara PTZ (*pan-tilt-zoom*), etc.

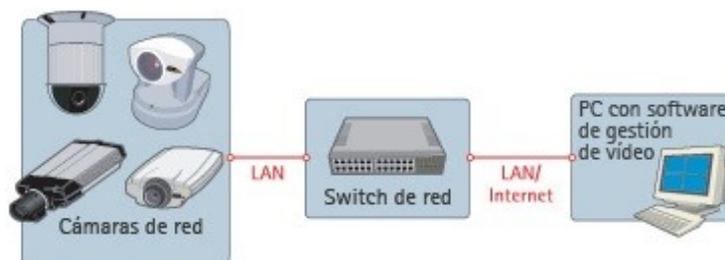


Figura 2-5: Sistemas de vídeo IP con cámaras IP [6]

2.1.3 Clasificación funcional de los sistemas de videovigilancia

Dentro de los sistemas de videovigilancia utilizados actualmente en seguridad se distinguen dos formas claras de llevar a cabo el control de los recintos videovigilados; una de carácter tradicional y otra basada en técnicas de visión artificial. La principal diferencia es el uso de sistemas dotados de inteligencia y autonomía que analizan los video capturados en tiempo real, frente a sistemas donde toda la labor recae en el operador [12].

2.1.3.1 Sistemas controlados por operadores

En los sistemas básicos de CCTV las cámaras cubren una o varias zonas de interés. Las imágenes captadas se graban y se monitorizan en un mural “videowall” del centro de control.

Es un operador el que supervisa los videos en los monitores. Cada controlador puede atender a varias decenas de cámaras, y el tiempo dedicado a cada escena será de 1-3 segundos cada 1-2 minutos; poco tiempo para detectar una situación de riesgo. Además, si el número de cámaras es elevado, se monitorizarán las escenas en turno rotatorio, lo que puede provocar desorientación [11].

En estos sistemas existen numerosas limitaciones derivadas de la intervención humana. Además resulta ineficiente e inadecuado cuando se quiere cubrir grandes áreas.



Figure 2-6: Videowall controlado por operador

2.1.3.2 Sistemas de videovigilancia automáticos

El objetivo de estos sistemas automáticos de videovigilancia no es sólo “vigilar”, sino también extraer de forma automática y en tiempo real la información relevante, facilitando con ello la labor posterior del operador.

En estos sistemas se intenta evitar la subjetividad del ojo humano; de forma que la secuencia que visualiza el operador contiene algún elemento detectado considerado motivo de alarma [11].

Con esta información detallada y en tiempo real, el operador puede actuar de forma eficiente. Dentro de estos sucesos u objetos en movimiento detectados por el sistema, algunos serán falsas alarmas que el operador descartará.

Sin embargo, la evolución de estos sistemas ha permitido que no sólo se eliminen algunas falsas alarmas de forma automática, sino también clasificar, seguir o reconocer la actividad del objeto en movimiento detectado [1].



Figure 2-7: Videowall de un sistema con detección de movimiento

También es muy común el uso de estos sistemas automáticos sin la supervisión de un operador en tiempo real. En esta situación sólo se almacenarán aquellas secuencias donde se detectó algún elemento, evitando así la posterior supervisión de largas horas de grabación [11,5].

2.2 Técnicas de análisis y segmentación de videos

El análisis de video en videovigilancia es un área en continuo desarrollo y evolución. Es precisamente este punto el que dota de inteligencia y automatismo a los sistemas de seguridad más utilizados en la actualidad.

La investigación en el campo de videovigilancia se centra fundamentalmente en tres etapas: detección de objetos en movimiento, su seguimiento (tracking), y análisis de parámetros y de comportamiento del elemento detectado [1]. Los próximos apartados se centrarán en esta primera etapa de detección, ya que es ahí donde comienzan a verse afectados los sistemas de videovigilancia automáticos por los efectos negativos de los cambios de iluminación.

En muchos de los sistemas de seguridad actuales, las técnicas para la detección de elementos móviles más utilizadas son las basadas en el modelado y posterior sustracción de fondo [1,13]. Está es la técnica de segmentación de video que se utilizará en este proyecto, para simular el comportamiento de un sistema real de videovigilancia.

2.2.1 Segmentación mediante técnicas de *Background Subtraction*

Según lo anterior, el primer paso en el análisis de secuencias de videovigilancia es diferenciar los elementos en movimiento (aunque en algunos casos pueden ser de interés objetos estáticos) de lo que serían regiones u objetos que permanecen estáticos en la escena, como por ejemplo el suelo o las paredes del área videovigilada. La extracción de estos objetos en movimiento o regiones de interés de la secuencia analizada se conoce comúnmente como segmentación.

La técnica de segmentación de video basada en la sustracción de fondo también es conocida comúnmente por el término inglés "*Background subtraction*". Debe su nombre al procedimiento de aplicar una sustracción píxel a píxel entre la imagen actual observada y la imagen estimada como fondo, y la umbralización del resultado para obtener los objetos de interés [13] tal que:

$$| \text{Image}_t - \text{Background}_t | > \text{Threshold}$$

En este tipo de propuestas de segmentación de video, en la etapa de detección de objetos en movimiento debemos destacar los tres componentes siguientes:

- *Background*= partes estáticas de las escena que permanecen visibles.
- Objetos= elementos de interés, por ejemplo intrusos.
- Artefactos= cambios en la imagen como sombras, brillos, luces,etc

La combinación de objetos y artefactos, es decir, todo aquello que no es *background*, es lo que se conoce como *foreground* [14]. Para lograr la posterior separación del *foreground* y obtener sólo los objetos será necesaria la aplicación de técnicas de procesamiento de imágenes que eliminen dichos artefactos.

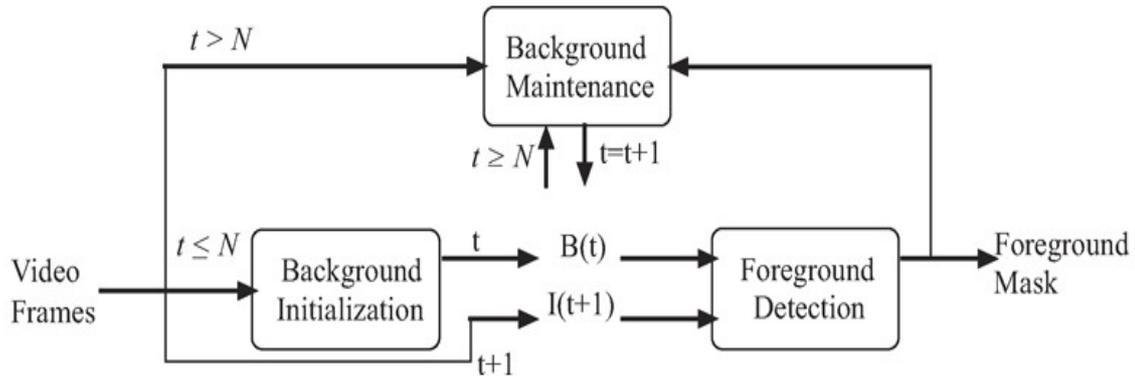


Figura 2-8: Diagrama de bloques de la técnica de sustracción de fondo. t : Tiempo. N : Número de imágenes. B : Imagen de fondo. I : Imagen actual [15]

Las técnicas de “*Background subtraction*”, a pesar de ser muy utilizadas, no están exentas de dificultades. Las más significativas son: la estimación de la imagen o modelo de fondo (*Background Initialization and Maintenance*, ver **Figura 2-8**) y la definición de lo que se considera un cambio importante respecto al fondo (*Foreground Detection*); cuestiones que se describirán en los siguientes apartados.

2.2.2 Modelado de fondo

El modelado de fondo es la etapa central de cualquier algoritmo de sustracción de fondo. Se trata de obtener de forma automática una imagen de fondo que represente las regiones estáticas de la escena.

A lo largo del tiempo se han propuesto diferentes métodos, como vemos en la revisión de Piccardi en [16]. Sin embargo todos ellos parten de un objetivo común: diseñar un modelo de fondo que tenga capacidad de adaptación para hacer frente a algunos cambios propios del entorno.

En sistemas de videovigilancia lo que se pretende es obtener un modelo de fondo robusto y eficiente que evite en esta fase inicial el mayor número de falsas alarmas posibles. En este sentido, un buen modelo debe adaptarse a ciertos cambios propios del entorno como:

- **Perturbaciones o movimientos:** desplazamientos leves de la cámara debidos al viento, movimientos continuados como en las hojas de los árboles, ondulaciones en el agua, etc.
- **Cambios en la escena de fondo:** inclusión de nuevos objetos como coches aparcados, objetos abandonados, que después un tiempo razonable en escena se deben considerar como fondo y no como motivo de alarma.
- **Adaptación a los cambios de iluminación:** lentos y graduales, como la variación lumínica del día; y a ser posible también a cambios rápidos y bruscos.

Además, para lograr mayor robustez hay que tener en cuenta la velocidad o la duración en escena de dichos cambios. Para conseguir esto se introduce un parámetro al modelo de fondo generalmente llamado factor de actualización. En las últimas décadas han aparecido numerosos algoritmos de sustracción de fondo enfocados a diferentes aplicaciones. Por ello, en este resumen del estado del arte comentaremos aquellos que a nivel de tiempo de procesamiento o de coste computacional permiten aplicaciones en tiempo real, requisito fundamental para los sistemas de videovigilancia.

En función de lo anterior se propone la siguiente clasificación de métodos de modelado de fondo [16]:

VELOCIDAD DE PROCESAMIENTO	MODELOS BÁSICOS	MODELOS PARAMÉTRICOS	MODELOS NO PARAMÉTRICOS
Alta	Promedio temporal, mediana, media móvil		
Intermedia		Mezcla de Gaussianas (MoG)	Kernel density estimators (KDS)

2.2.2.1 Modelos Básicos

En este caso el valor de cada píxel del modelo de fondo se calcula en función de los valores recientes de dicho píxel. Se utilizan modelos matemáticos sencillos como la diferencia entre imágenes, valores promedios, máximos y mínimos, etc.

En la mayoría de los modelos propuestos el valor histórico se calcula a partir de los N frames anteriores (**Figura 2-8**), donde N es una constante elegida para el problema. Algunos modelos utilizan una media ponderada del valor de los píxeles, donde las últimas imágenes tienen mayor peso.

➤ Filtro Promedio Temporal (Average)

En este modelo se calcula una imagen de fondo estática hasta que se produce algún cambio. En ese momento el valor del fondo se corresponde con el promedio de N frames consecutivos. El modelo de fondo (B_t) se calcula tal que:

$$|I_t - B_t| > Th \rightarrow B_t = \frac{1}{N} \cdot \sum_{n=t}^{n=N-t} I_n$$

$$|I_t - B_t| \leq Th \rightarrow B_t = I_t$$

donde

$I_t \equiv$ imagen de entrada en el instante t

$B_t \equiv$ imagen de fondo estimada en el instante t

Este método no resulta eficiente cuando hay varios objetos, o estos se mueven lentamente. Esto se debe a que es la información de movimiento la que se utiliza para actualizar el fondo, por lo que si un elemento se mueve lentamente no se detectará. Una posible solución sería tener en cuenta también los píxeles no marcados como objetos móviles a la hora de generar el modelo de fondo.

- Filtro Mediana (*Median*)

El filtro mediana es una de las técnicas de modelado de fondo más utilizadas. El fondo se estima como la mediana de los píxeles almacenados en un *buffer* de N frames anteriores. Uno de los problemas de este método es que cuando se produce algún cambio el fondo se adapta muy lentamente, como se menciona en [17].

La principal ventaja de los dos métodos anteriores es que son bastante rápidos a nivel de coste computacional. Sin embargo los requisitos de memoria son altos, del orden de $N \cdot \text{sizeof}(\text{frame})$

- Filtro de media móvil (*Running Average*)

Este tipo de filtros evita los problemas de memoria de los métodos anteriores, ya que no es necesario un *buffer* de almacenamiento. En este método se calcula una media ponderada entre la última imagen de *background* y la imagen actual (I_t). Se hace uso del parámetro α , denominado factor de actualización, con valores del orden de 10^{-2} . El valor de cada píxel del fondo se calcula tal que:

$$B_t = \alpha * I_{t-1} + (1 - \alpha) \cdot B_{t-1}$$

En numerosas aplicaciones se utiliza el modelo anterior acompañado de la propiedad de selectividad, como se describe en [16]. Para ello se clasifica cada uno de los píxeles como *background* si $|I_t - B_t| < Th$ o *foreground* (F_t) si $|I_t - B_t| \geq Th$ de manera que:

Si $p(x, y)$ es *background*:

$$B_t(x, y) = \alpha * I_{t-1}(x, y) + (1 - \alpha) \cdot B_{t-1}(x, y)$$

Si $p(x, y)$ es *foreground*:

$$B_t(x, y) = B_{t-1}(x, y)$$

Con la modificación anterior lo que se consigue es evitar que el modelo de fondo se corrompa con píxeles que no pertenecen realmente a la escena. Sin embargo, en algunas aplicaciones de videovigilancia puede resultar interesante e incluso necesario incluir información relativa al *foreground*. En este sentido se hace uso de un modelado de fondo con las siguientes características:

Si $p(x, y)$ es *background*:

$$B_t(x, y) = \alpha_{back} \cdot I_{t-1}(x, y) + (1 - \alpha_{back}) \cdot B_{t-1}(x, y)$$

Si $p(x, y)$ es *foreground*:

$$B_t(x, y) = \alpha_{fore} \cdot I_{t-1}(x, y) + (1 - \alpha_{fore}) \cdot B_{t-1}(x, y)$$

Se utilizarán para este método valores típicos de $\alpha_{back} \approx 10^{-2}$ y $\alpha_{fore} \approx 10^{-3}$

Se consigue en este caso una solución intermedia, de forma que el modelo de fondo no se corrompa fácilmente pero que sea adaptable a cambios. De esta forma, aquellos objetos en movimiento detectados que permanezcan un tiempo razonable en escena formarán parte del fondo.

2.2.2.2 Modelos que estiman una distribución de probabilidad

2.2.2.2.1 Modelos Paramétricos

Los modelos paramétricos basan el modelo de fondo en una distribución estadística estándar de la cual hay que estimar los parámetros. Se trata de métodos más complejos con más capacidades adaptativas frente a cambios leves, al ruido, etc. La complejidad computacional se traduce en una menor velocidad de cómputo en comparación con los modelos básicos, sin embargo estos métodos funcionan bien sin la necesidad de almacenar tantas imágenes como el filtro media o mediana [16].

- Mezcla de Gaussianas (*Mixture of Gaussians*)

Estos métodos trabajan con fondos multimodales [16]. Se utilizan generalmente en escenarios donde hay píxeles cuya intensidad fluctúa constantemente; tales como hojas de árboles, cámaras de baja calidad, agua en movimiento, etc.

Debido a estos cambios de intensidad hacer uso de una única gaussiana por píxel (media y varianza de las intensidades) no es suficiente. Lo que pretende este método es caracterizar cada píxel como una mezcla de gaussianas (MoG).

La distribución de cada píxel $f(I_t = u)$ se modela con k gaussianas de la forma:

$$f(I_t=u) = \sum_{i=1}^K w_{i,t} \cdot \eta(u; \mu_{i,t}, \sigma_{i,t})$$

Donde $\eta(u; \mu_{i,t}, \sigma_{i,t})$ es la componente gaussiana i -ésima, con media de intensidad $\mu_{i,t}$, cuya desviación estandar es $\sigma_{i,t}$ y con un peso $w_{i,t}$ que determina la cantidad de distribución utilizada en dicha componente. Los valores típicos de K varían entre 3 y 5. Esta descripción de MoG se basa en el esquema de Kamath y Cheung [17].

Para actualizar el modelo de fondo se compara la intensidad de cada píxel del frame actual con sus posibles distribuciones en la imagen de fondo. Si el valor del píxel no supera el de alguna media $\mu_{i,t}$ en más de la desviación $\sigma_{i,t}$ se considerará un píxel de fondo y se actualizarán los parámetros para ese píxel. Si el píxel no se parece a ninguna de sus distribuciones asociadas se actualizará el modelo sustituyendo la distribución de menor peso.

Al igual que en los métodos básicos, existen diferentes modalidades de MoG en función de la aplicación concreta. En los últimos años se han hecho numerosos estudios relacionados con el modelado de fondo basado en mezcla de gaussianas logrando mayor adaptabilidad a los cambios, mejoras en seguimiento o detección de automóviles. Sin embargo estos modelos tienen alta carga computacional [16] y no son robustos a cambios rápidos de iluminación [18], factores en controversia con los objetivos de este proyecto.

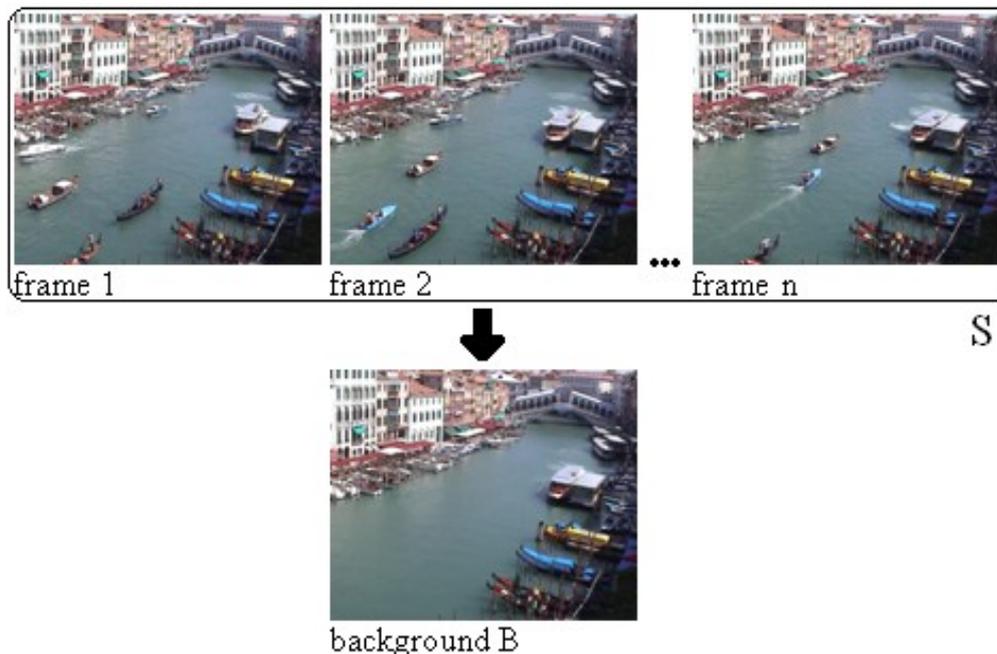


Figura 2-9: Modelado de fondo mediante MoG

(De D. Bloisi, L. Iocchi, G.R. Leone and R. Pigliacampo, "A distributed vision system for boat traffic monitoring in the Venice Grand Canal", ARGOS project)

2.2.2.2 Modelos No Paramétricos

Estos métodos se basan en una estimación no paramétrica de la función densidad de probabilidad de una variable aleatoria. Los modelos no paramétricos son métodos complejos en los que no se asumen distribuciones estándar de probabilidad.

- Kernel Density Estimators (KDE)

Este modelo estima la probabilidad de la intensidad de cada píxel en función de una serie de muestras de intensidad anteriores para dicho píxel. La función densidad de probabilidad viene dada por el histograma de los N últimos valores, que se han almacenado en el buffer.

Según Elgammal *et al.* en [18] la pertenencia al fondo de un píxel concreto se calcula de la forma siguiente: dada la muestra del valor histórico del píxel x_1, x_2, \dots, x_n , la función densidad de probabilidad de que ese píxel tome valor x_t en el instante t puede ser estimada de forma no paramétrica utilizando un *kernel estimator* $K(.)$ tal que:

$$Pr(x_t) = \frac{1}{n} \sum_{i=1}^N K(x_t - x_i)$$

El píxel se considerará fondo si $Pr(x_t) < Th$, en caso contrario será frente.

Este método soporta cambios leves en la escena, movimientos en las hojas de árboles y es robusto al ruido. Sin embargo tiene alta carga computacional, por lo que será más lento que otros métodos de los citados anteriormente.

2.2.3 Detección de frente

Según las técnicas de *Background Subtraction*, la detección de frente o *foreground* se obtiene tras comparar la imagen original con el modelo de fondo generado. Se trata de una técnica basada en diferencia. Es un método sencillo y eficaz en la detección de elementos en movimiento.

Como hemos visto ya en el **apartado 2.2.1**, se realiza una sustracción entre ambas imágenes y los píxeles se clasifican como *foreground* si la diferencia supera un umbral fijado. A partir de esto tendremos una máscara binaria de píxeles candidatos a ser objeto en movimiento.

$$\begin{aligned} |I_t(x, y) - B_t(x, y)| > Th &\rightarrow F_t = 1 \\ |I_t(x, y) - B_t(x, y)| \leq Th &\rightarrow F_t = 0 \end{aligned}$$



Figura 2-10: Detector de frente basado en diferencia.

Para la detección de frente en este caso se utiliza la diferencia absoluta. Esta proporciona objetos de apariencia sólida y bordes gruesos, lo que puede resultar útil en las etapas posteriores de análisis o seguimiento del objeto detectado.

Uno de los principales problemas es la fijación del umbral “ Th ” [19], ya que depende de varios factores como las condiciones de iluminación, la sensibilidad de la cámara o el ruido que ésta introduce.

Este es un punto crítico en la detección de objetos en movimiento. Aunque se puede establecer el umbral de forma empírica, en videovigilancia se suele adaptar a una serie de restricciones como no perder ningún verdadero positivo o reducir las falsas alarmas.

2.3 Limitaciones de los sistemas de videovigilancia automáticos basados en detección de movimiento

2.3.1 Tipología de los problemas derivados de la etapa de segmentación

La etapa de segmentación de video es uno de los procesos más complicados dentro de los sistemas de seguridad automáticos. Un buen sistema debe poder superar ciertos obstáculos, ya que ni las secuencias de video son ideales ni el proceso de sustracción de fondo es perfecto.

Como señala McIvor *et al.* en [14], el frente o foreground está compuesto por objetos de interés y por artefactos que dificultan la siguiente etapa de proceso (detección, tracking, conteo automático de vehículos o personas); o que generan falsas alarmas en los sistemas de vigilancia para la detección de intrusos.

En los últimos años muchos autores han trabajado en este campo, con el objetivo de hacer más robustos los sistemas de videovigilancia. Se citan a continuación los problemas que encontramos con mayor frecuencia en la literatura; algunos de ellos ya mencionados en este resumen del estado del arte.

- **Inicialización y actualización del fondo de la escena:** Como ya se ha mencionado en el apartado anterior, la generación y mantenimiento del modelo de fondo debe adaptarse a la aplicación concreta. Además la inicialización del modelo de fondo no siempre es un proceso sencillo.
- **Fondos multimodales:** Son modelos de fondo que se usan en escenarios donde se producen cambios lentos y periódicos que se deben incluir en dicho fondo; como el movimiento del agua del Gran Canal de Venecia de la **Figura 2-9**.
- **Determinación de los parámetros óptimos de funcionamiento de los algoritmos:** Ajustar de forma adecuada los parámetros de estos sistemas automáticos resulta una tarea complicada. La modificación de umbrales, pesos, factor de actualización u otros parámetros puede alterar significativamente el funcionamiento del sistema.
- **Ruido:** Se considera ruido cuando el valor de un píxel de una imagen está distorsionado respecto del valor representativo de la intensidad real. Es muy típico en imágenes de videoseguridad, debido a que los sistemas de captación introducen bastante ruido.
- **Sombras y reflejos:** Las diferentes fuentes de luz presentes en una escena generan una iluminación poco homogénea. Esto provoca sombras y reflejos que nos impiden distinguir los objetos de interés de los artefactos luminosos. Serán necesarias técnicas de post-procesado para eliminar las sombras o reflejos del foreground y así quedarnos sólo con el elemento real.



Figura 2-11: Ejemplo de Técnicas de post-procesado para la detección de sombras

(De Cucchiara R, Grana C, Piccardi M, Prati A. Detecting moving objects, ghosts and shadows in video streams. IEEE Trans. Patt Anal Mach Intell, PAMI 2003; 25(10):1337-1342.)

- **Cambios de iluminación en la escena:** La variación de la iluminación provoca cambios significativos en la secuencia obtenida. Se puede tratar de cambios graduales a lo largo del día, o rápidos provocados por el paso de una nube cuando estamos en escenarios exteriores (*outdoor*). En el caso de escenarios interiores (*indoor*) los cambios de iluminación se deben principalmente a las diferentes fuentes artificiales de luz. Este tipo de cambios provocan numerosas falsas alarmas en los sistemas de seguridad basados en detección de movimiento [7], ya que se interpretan como objetos en movimiento.



Figura 2-12: Ejemplo de cambio de iluminación debido al paso de nubes

2.3.2 Aproximaciones existentes al problema de la iluminación

Los sistemas automáticos de videovigilancia necesitan detectar de forma precisa los objetos móviles que aparecen en la escena para lograr unos objetivos específicos, como pueden ser: detección de intrusos, conteo de personas, detección de objetos robados u abandonados, seguimiento (*'tracking'*) de objetos en movimiento o análisis de comportamientos, entre otros [1].

Como hemos mencionado, las técnicas de segmentación de video no proporcionan resultados suficientemente eficientes. Por ello será necesario aplicar técnicas de post-procesado que reduzcan los problemas que pueden afectar al funcionamiento deseable del sistema como son las variaciones de luz.

Los problemas citados en el **aparatado 2.3.1** son el principal objetivo de las técnicas de post-procesado que han propuesto algunos autores, destacando por su elevada complejidad la eliminación de sombras o el tratamiento de los cambios de iluminación (**Figuras 2-11 y 2-12**).

Muchos de los investigadores han diseñado algoritmos que aunque no solventan por completo el problema de la iluminación, lo reducen en gran medida. Se trata de reducir los efectos negativos de la variaciones de luz mediante técnicas de *Background Subtraction* robustas a cambios de iluminación.

2.3.3 Propuestas actuales para la reducción de los problemas debidos a cambios de iluminación.

La complejidad en la reducción de los efectos debidos a cambios de iluminación radica en que estos son muy diferentes entre sí y están fuertemente ligados a las condiciones del escenario donde se encuentra el sistema de captación.

Existen cambios lentos, graduales, bruscos, rápidos, debidos a fuentes naturales o artificiales, cambios dinámicos donde la fuente de iluminación se desplaza, variaciones de iluminación tan intensas que saturan la imagen, etc. Debido a esta variedad es difícil encontrar una propuesta única capaz de solucionar el problema en todos los casos.

Los cambios de iluminación graduales que se producen a lo largo del día si la cámara se encuentra al aire libre se pueden solucionar de forma sencilla utilizando técnicas de modelado de fondo. Sin embargo para los demás cambios de iluminación la solución no es tan evidente.

La mayoría de las técnicas propuestas en la literatura para abordar el problema de la iluminación se basan en depurar los algoritmos de sustracción de fondo, de forma que funcionen correctamente ante los cambios de luz [7].

En este sentido se han propuesto numerosos métodos de sustracción de fondo basados en mezcla de gaussianas, como el trabajo de Stauffer and Grimson [20]; o se han añadido algunas variantes y modificaciones a lo anterior como Tian *et al.* [21].

Otros estudios se centran en el análisis de la información extraída del gradiente de la imagen, como describen Klare y Sarkar en [8] o de la textura del fondo como proponen Ding *et al.* en [22].

En el trabajo de Javed *et al.* [9], se hace uso del valor y de la orientación del gradiente en combinación con la intensidad de los planos RGB para crear un algoritmo de mezcla de gaussianas de cinco dimensiones. Con esto se acercan a un mecanismo de *background subtraction* robusto ante cambios de iluminación sin perder de vista otros problemas, como la inicialización del fondo o el efecto de las sombras.

Algunos autores [10, 23] intentan hacer uso del filtro de Kalman para conseguir un modelo de fondo adaptativo que tenga en cuenta los cambios de iluminación.

Existen otros muchos estudios que intentan solventar el problema de los cambios de iluminación por medio de la predicción temporal [24] o mediante cambios de espacio de color [22, 25].

Otra de las propuestas con resultados notables es el uso de varias cámaras para generar un modelo de fondo de referencia donde se tenga información de la geometría de la escena como en el trabajo de Ivanov *et al* [26] o en el de Lim *et al*. [27]. De este último destacamos la imagen siguiente que ilustra tanto el problema, como la solución que proponen.



Figura 2-13: Detección de cambios de iluminación utilizando dos vistas. De izquierda a derecha: vista de referencia, segunda vista, detección con una cámara y detección con dos cámaras. [27]

3 Detección de intrusos y problemas debidos a los cambios de iluminación en sistemas de videovigilancia

En este capítulo se describe en primer lugar el contexto en el que se ha realizado este proyecto. A continuación se detalla el procedimiento general de los sistemas de videovigilancia basados en detección de movimiento para la detección de intrusos y el proceso de generación de alarmas.

Una vez conocida esta actividad, pasaremos a la explicación de los efectos que provocan los cambios de iluminación. De esta forma se intentará realizar una clasificación inicial de la casuística; así como comprender el funcionamiento del sistema ante estos casos.

El fin de este capítulo es tener una visión general del problema real, de forma que podamos desglosarlo en apartados de menor complejidad para poder abordarlo de manera sencilla y eficiente.

3.1 Aplicaciones en las que se enfocará este proyecto

El principal objetivo de los sistemas de videovigilancia inteligentes es extraer de forma automática y en tiempo real la información relevante que está aconteciendo en un escenario videovigilado.

Las posibilidades de los sistemas de videovigilancia automáticos son muy numerosas, tales como control de flujo de personas en aglomeraciones, entrada y salida en recintos, conteo de vehículos, detección de objetos abandonados o robados, localización de incendios, detección de infracciones de tráfico, control de fronteras, vigilancia de instalaciones, etc.

En este proyecto nos centraremos en aquellas aplicaciones más enfocadas a la seguridad, donde principalmente interesa detectar personas o vehículos que entran en el área que se intenta controlar. La motivación del uso de sistemas automáticos en seguridad es evitar la subjetividad del ojo humano; de forma que el mecanismo de alarma sea un proceso inteligente e independiente del operador. Además se pretende evitar la fatiga y distracción que provoca en los operadores la continuidad de falsas alarmas, así como la posibilidad de vigilar zonas más extensas mediante un número de cámaras mayor.

3.2 Sistemas de videovigilancia basados en detección de movimiento

El primer paso en el análisis de secuencias de videovigilancia de estas características es diferenciar los elementos en movimiento de lo que serían regiones u objetos pertenecientes a las escena. La extracción de estos objetos en movimiento o regiones de interés de la secuencia analizada se conoce comúnmente como segmentación [1].

En muchos de los sistemas de seguridad actuales, las técnicas de segmentación de video más utilizadas son las basadas en el modelado y posterior sustracción de fondo (*Background Subtraction*). Está es la técnica de segmentación de video que utiliza la herramienta de procesamiento que se ha utilizado en este proyecto, para simular el comportamiento de un sistema real de videovigilancia.

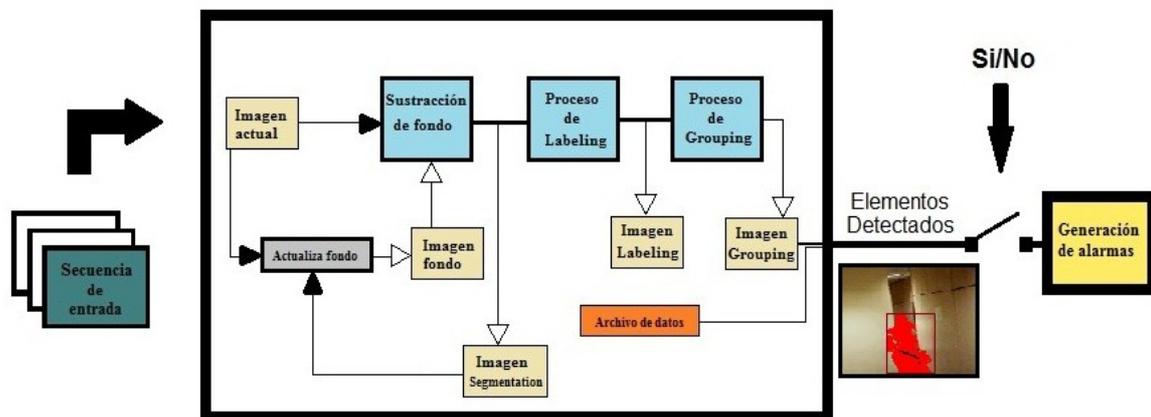


Figura 3-1: Arquitectura básica de un sistema de videovigilancia basado en detección de movimiento.

Se muestra en la **Figura 3-1** el funcionamiento de esta herramienta; donde la detección de cambios se obtiene tras realizar el proceso de sustracción de fondo. La herramienta utilizada también realiza los procesos de *labeling* y *grouping* que son etapas de post-procesado que se utilizan para definir mejor y delimitar los elementos detectados, lo que se conoce como *connected components*. Este proceso se lleva a cabo para facilitar las etapas posteriores de seguimiento de objetos (*tracking*), como proponen Stauffer y Grimson en [20]. Por lo tanto, en esta simulación de un sistema de videovigilancia, cada vez que se detecte algún cambio respecto al modelo de fondo, a la salida tendremos una imagen con los elementos detectados. Se considera que si hay algún elemento detectado se generará una alarma en el sistema.

3.3 Generación de alarmas en sistemas basados en detección de movimiento

Como hemos mencionado antes, a la salida de la etapa de sustracción de fondo tendremos todo aquello que es diferente al background; y si es diferente se considera que se ha movido. Sin embargo, es muy importante tener en cuenta que no todo lo que “se mueve” es necesariamente un objeto de interés.

Por lo tanto, en sistemas basados en detección de movimiento la generación de alarmas se producirá tanto en situaciones de interés como en situaciones erróneas; es decir, habrá alarmas verdaderas y alarmas falsas.

En la mayoría de escenarios donde se instalan este tipo de sistemas de seguridad, las alarmas que se catalogan como verdaderas son las apariciones de intrusos, y a veces también vehículos, en la zona a vigilar. Sin embargo el número de falsas alarmas suele ser grande, reduciendo considerablemente la eficiencia del sistema. En algunos casos este problema se traduce en la necesidad de un operador que supervise las alarmas que arroja el sistema y determine si son falsas o verdaderas (ver apartado 2.1.3.2).

Se citan a continuación algunos casos típicos de falsas alarmas:

- Desplazamientos leves de la cámara debidos al viento.
- Movimientos continuados como en el de las hojas de los árboles, ondulaciones en el agua o la lluvia.
- Ruido excesivo en la imágenes.
- Aparición de sombras pertenecientes a objetos que no están dentro de escena.
- Animales que pasan por la zona de seguridad como gatos u otros.
- Telas de araña que ocultan el objetivo de la cámara generando falsas alarmas cuando ondulan con el viento, ver **Figura 3-2**.
- Insectos que cruzan a corta distancia del objetivo. Este efecto se puede ver el la **Figura 3-3**.
- Cambios de iluminación debidos al paso de nubes, faros de coches o encendido de nuevas fuentes de luz como por ejemplo el alumbrado público.



Figura 3-2: Falsa alarma provocada por el movimiento de una telaraña situada delante del objetivo de la cámara

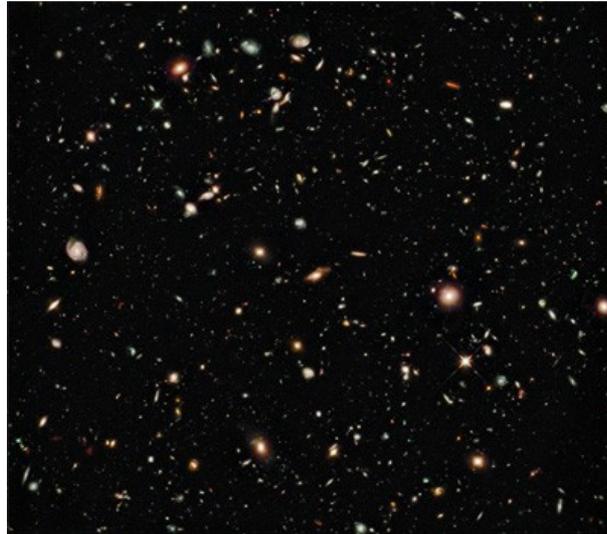


Figura 3-3: Muestra de la falsas alarmas provocadas por insectos

3.4 Efectos de los cambios de iluminación y falsas alarmas debidas a estos efectos negativos

En este proyecto, interesa reducir los falsos positivos debidos a cambios de iluminación. En este apartado se intentará describir e ilustrar de forma detallada los efectos que provoca la alteración de la iluminación en los sistemas de videovigilancia inteligentes, así como las diferencias con la aparición de intrusos.

Existen diversos tipos de variaciones de luz que causan efectos muy distintos entre sí, como pueden ser encendido-apagado de luces, destellos de faros de coches o el movimiento de nubes. Para realizar una primera clasificación consideramos que los cambios de iluminación alteran el escenario de forma diferente según el tipo de fuente de la que proviene la iluminación; en este sentido dividiremos inicialmente el problema en función del escenario.

Se muestran a continuación una serie de imágenes que ilustran el problema que se ha abordado en este proyecto, atendiendo a la separación en escenarios *indoor* o interiores y escenarios *outdoor* o exteriores. Algunas de estas imágenes tienen principalmente carácter ilustrativo, puesto que no están captadas con una cámara de altas prestaciones, como las que se utilizan en videovigilancia. Sin embargo además de estas imágenes, se dispone también de una base de datos, con una amplia casuística para desarrollar algoritmos, realizar cálculos, obtener estadísticas y evaluar resultados con secuencias de video capturadas en sistemas reales.

3.4.1 Efectos en escenarios Indoor

En primer lugar, consideramos escenarios indoor aquellos donde las cámaras están situadas en un recinto interior; donde en general las variaciones luminosas se deben al encendido-apagado de las fuentes de luz eléctricas.

Dentro de este tipo de escenario tendremos verdaderos positivos provocados por la intrusión de personas y falsos positivos originados por el encendido-apagado de luces. Dado que en este caso el espacio captado por la cámara suele ser limitado [28], un on-off de luces causará un cambio global en toda la imagen y una persona será un cambio mucho más localizado. Estos efectos se pueden apreciar en los siguientes ejemplos gráficos.

3.4.1.1 Falso positivo debido a un cambio de iluminación leve:



Figura 3-4: Respuesta del sistema ante un cambio de iluminación leve

3.4.1.2 Falso positivo debido a un cambio de iluminación brusco, on-off de luces:



Figura 3-5: Respuesta del sistema ante un cambio de iluminación brusco

3.4.1.3 Falso positivo debido a un cambio de iluminación muy brusco, on-off de luces:



Figura 3-6: Respuesta del sistema ante un cambio de iluminación muy brusco

3.4.1.4 Verdadero positivo debido a un intruso corriendo:

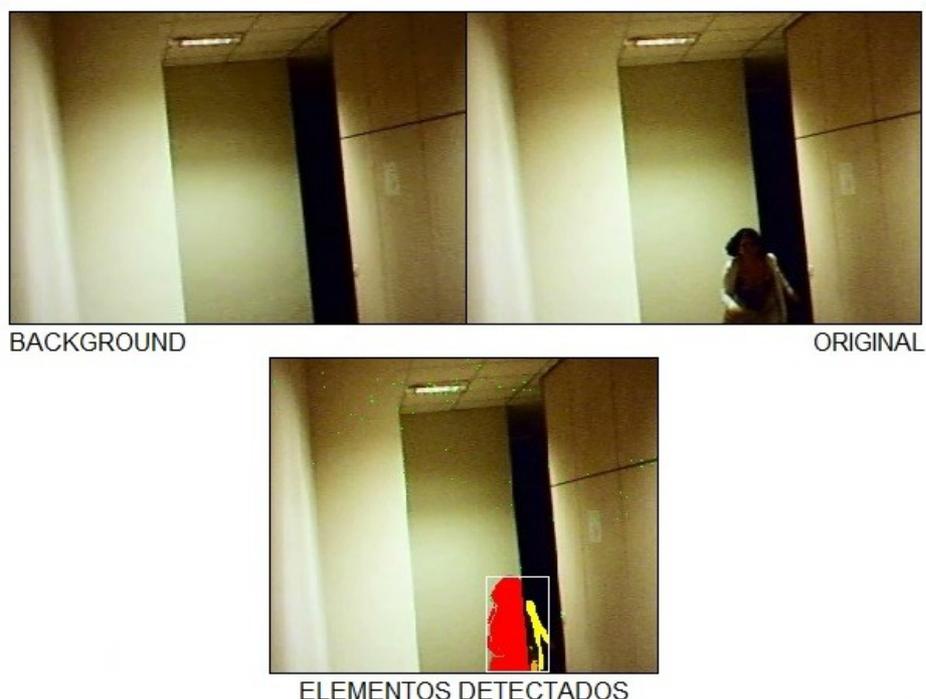


Figura 3-7: Respuesta del sistema ante un caso de alarma real

Como se puede ver en todas las figuras anteriores el impacto que causa el encendido-apagado de luces en un escenario indoor afecta a casi toda la imagen. El sistema generará una falsa alarma cada vez que suceda un evento de este tipo. Sin embargo sólo observando esta muestra ya se ven grandes diferencias respecto al efecto que causa una persona.

3.4.2 Efectos en escenarios Outdoor

Consideramos escenarios outdoor aquellos donde las cámaras están situadas en un recinto exterior tales como carreteras, aparcamientos, fincas privadas o recintos industriales.

En este caso el espacio captado por la cámara suele tener más amplitud y profundidad por lo que los cambios se pueden producir en zonas más próximas a la cámara o en un plano más lejano [28]. Esto aumenta la complejidad del problema frente al caso indoor.

Dentro de este escenario se debe destacar que el efecto provocado por los faros de los vehículos es mucho más intenso que el encendido de una farola o el paso de una nube que oculta el sol. Sin embargo este problema se produce fundamentalmente de noche, donde además la sensibilidad de las cámaras es mucho menor. Es por este motivo que se subdivide el problema en función de si las secuencias se han tomado en condiciones diurnas o nocturnas.

3.4.2.1 Grabaciones diurnas

En condiciones diurnas la cámara captará imágenes con buena calidad y a color, lo que permite que el sistema inteligente detecte cualquier elemento móvil sin dificultad mediante las técnicas de sustracción de fondo.

Dentro de este tipo de escenario se generan alarmas verdaderas provocadas por la intrusión de personas y vehículos, y falsas alarmas provocadas por el encendido-apagado de luces cercanas a la cámara, variaciones bruscas de la iluminación natural o luces provenientes de faros de coche.

Al igual que en el apartado 3.4.1 de este capítulo, intentaremos ilustrar las consecuencias de los cambios de iluminación en escenarios exteriores, así como la respuesta del sistema de detección. Se mostrarán también casos de alarmas verdaderas para apreciar las diferencias.

3.4.2.1.1 Falso positivo debido a un cambio brusco de la luz natural (movimiento de nubes que ocultaban el sol)

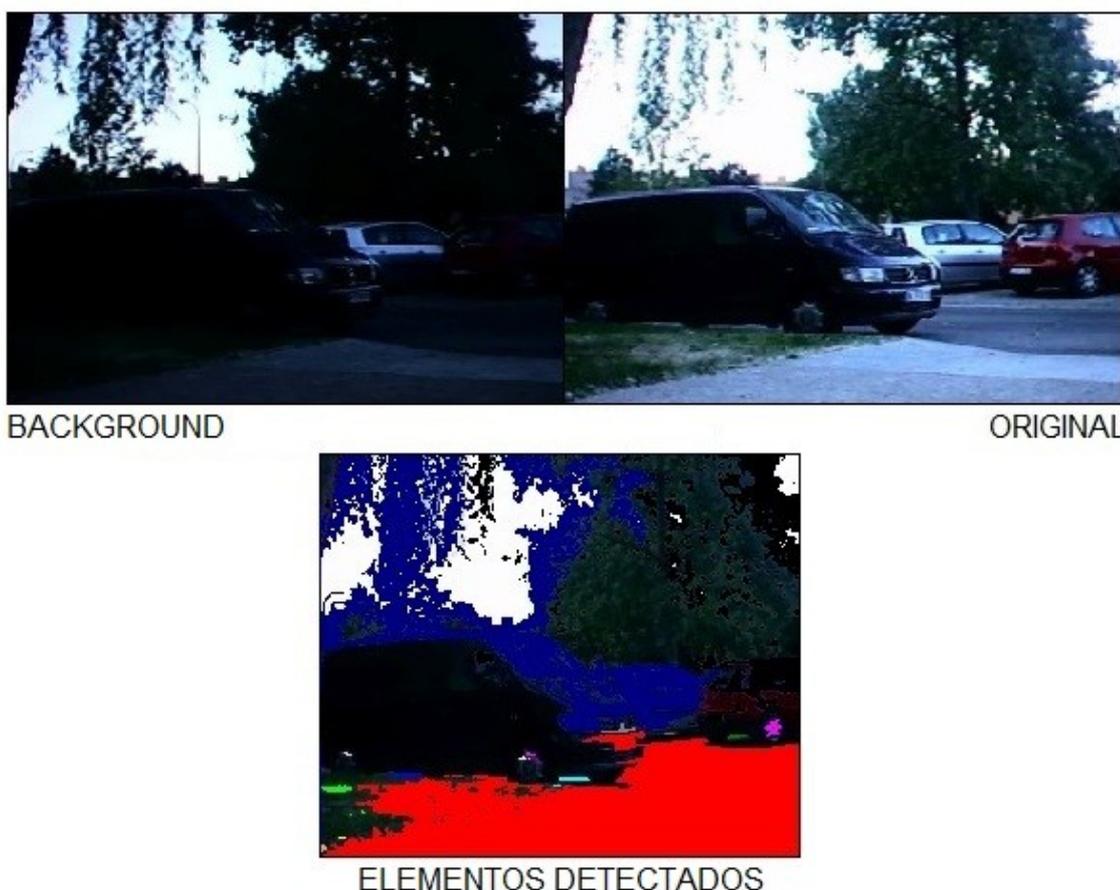


Figura 3-8: Respuesta del sistema ante un cambio de iluminación natural

Otro ejemplo de este tipo de variaciones de luminosidad natural puede verse en la imagen izquierda de la **Figura 4-7**.

3.4.2.1.2 Falso positivo debido al encendido-apagado de una fuente de luz.



Figura 3-9: Respuesta del sistema ante la desaparición de una fuente luminosa

En la **Figura 3-9** se observa que otro de los problemas a tratar es que la aparición o desaparición de algún foco o la iluminación de una ventana próxima a la cámara generará una alarma. En muchos de los recintos donde se instalan sistemas de videovigilancia automáticos resulta imposible evitar que ventanas o farolas estén fuera de la escena; de ahí la necesidad de intentar reducir los falsos positivos de este tipo.

3.4.2.1.3 Verdadero positivo debido a la aparición en escena de una persona.



Figura 3-10: Respuesta del sistema ante un caso de alarma real

3.4.2.2 Grabaciones nocturnas

Las gran mayoría de las cámaras utilizadas en videovigilancia tienen dos modalidades de trabajo, una para el día y otra para la noche o condiciones muy bajas de iluminación. Este tipo de cámaras conocidas como *IR-sensitive cameras* incluyen una iluminación infrarroja (de longitud de onda alrededor de 850nm). En condiciones nocturnas la cámara captará imágenes en niveles de gris con una calidad aceptable que permiten visualizar el área a vigilar a pesar de la poca iluminación externa. La **Figura 3-11** muestra la diferencia entre una imagen captada por una cámara sin iluminación infrarroja y otra con una cámara acompañada con un foco infrarrojo, *IR-illumintor* [6].



Figura 3-11: Escena grabada con una cámara sin un foco IR infrarroja y con iluminación infrarroja [6].

Sin embargo muchas de las cámaras que se colocan en escenarios exteriores no captan imágenes tan nítidas como la de la figura anterior. Esto se debe a que en lugar de un gran foco IR tienen es una corona alrededor del objetivo, por lo que no tienen tanta sensibilidad (**Figura 3-12**) . Esto se traduce en imágenes de peor calidad y más afectadas por el ruido, lo que ocasiona que el sistema inteligente no consiga segmentar los elementos móviles con la misma efectividad que en condiciones diurnas o si se tuviera un sistema de captación de alto nivel.



Figura 3-12: Ejemplo de cámara IR utilizada en escenarios exteriores

Debido a lo comentado en el párrafo anterior, el espacio captado por la cámara suele tener una zona más iluminada en las regiones próximas a la cámara (donde alcanza el IR) y en el plano más lejano, la imagen será mucho más oscura. Por este motivo, el sistema tiene más dificultades ya que los objetos lejanos se mimetizan bien con el entorno, situación que se describirá mejor más adelante. Otro problema añadido se debe a que las grabaciones nocturnas estarán más afectadas por el ruido que introducen los sistemas de captación en condiciones de baja iluminación.

Dentro de los escenarios outdoor también tendremos alarmas verdaderas provocadas por la intrusión de personas y vehículos, y falsas alarmas provocadas por el encendido-apagado de luces contiguas a la cámara, como se ve en la **Figura 3-9**, o luces provenientes de faros de automóviles.

En los próximos apartados se intentará especificar algunas de las consecuencias de las variaciones de iluminación cuando se trabaja de noche. En estas circunstancias se encuentran efectos nuevos que no se observaban en las secuencias indoor, ni en tampoco en las outdoor durante el día.

3.4.2.2.1 Falso positivo debido a los faros de los vehículos.

En estas condiciones de baja iluminación o nocturnidad los efectos provocados por los cambios de luz son mucho más intensos que en las circunstancias anteriormente descritas. Una de las alteraciones más bruscas son las causadas por las ráfagas de luz de automóviles, donde habrá regiones de la imagen saturadas por la iluminación.

Cuando se habla de faros de vehículos otro problema añadido es el movimiento. La iluminación se mueve por toda la escena, se refleja en algunas superficies como en cristales de ventanas o en la chapa de otros coches aparcados, transformando la imagen de tal forma que se dificulta tanto el proceso de segmentación, como el de *tracking*. Se muestra en la **Figura 3-13** una breve secuencia donde aparece sólo la iluminación de un vehículo que se desplaza. Podemos ver en la imágenes de la derecha el efecto mencionado; donde los reflejos hacen que el sistema detecte hasta tres elementos diferentes, cuando el suceso real es simplemente una variación de luz.

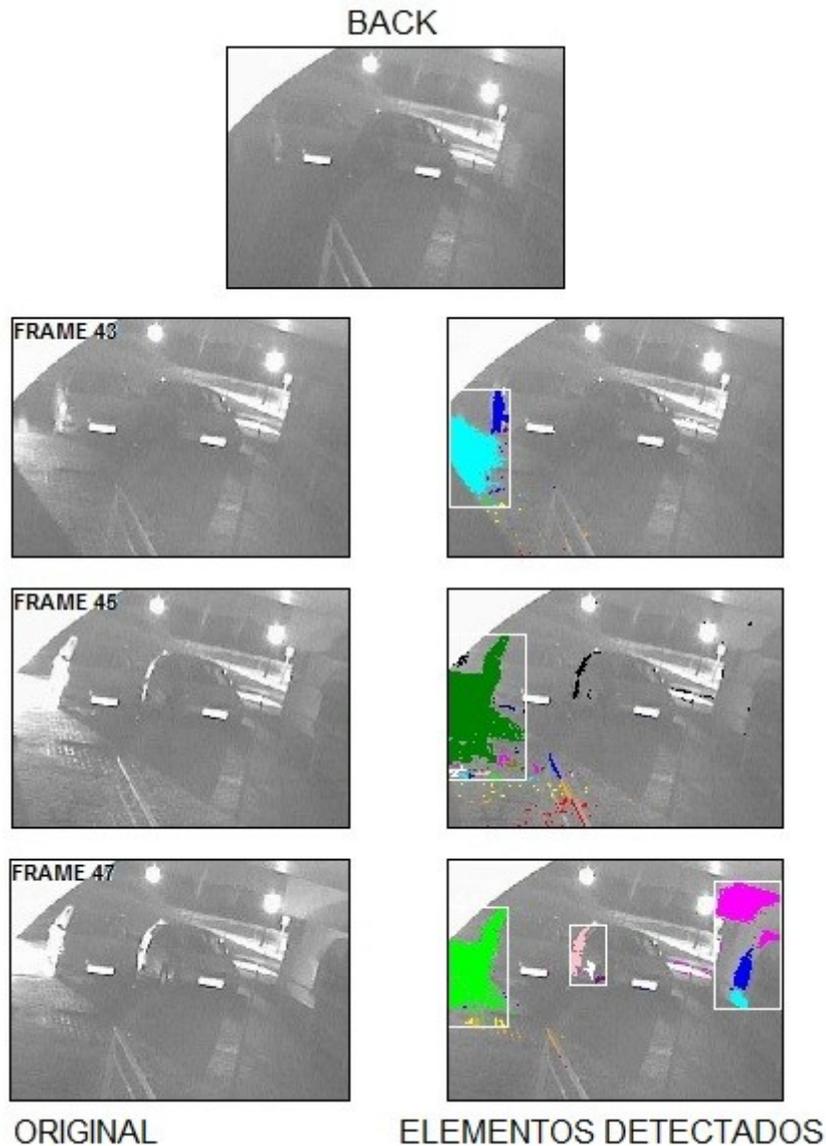


Figura 3-13: Efectos sobre la escena de la iluminación en movimiento

En algunas ocasiones, la gran magnitud de estas variaciones puede llegar a provocar una falsa alarma de tipo “obstrucción”. Una alarma de estas características sucede cuando la iluminación es tan intensa que la cámara cambia a modo diurno; es decir pasa de captar imágenes en niveles de gris a imágenes a color. Se consideran alarmas por obstrucción aquellas donde la máscara de *foreground* tiene marcado un porcentaje de píxeles muy alto; generalmente porque un intruso intenta tapar el objetivo o sabotear la cámara. Sin embargo cuando la cámara cambia de modo por culpa de la nueva iluminación se marcarán grandes áreas como si hubieran cambiado, siendo el efecto confundible con un intento de obstruir la grabación.



Figura 3-14: Cambio de la cámara a modo diurno causado por un alumbrado intenso. Generación de alarma de tipo “obstrucción”

3.4.2.2 Verdadero positivo debido a la aparición en escena de vehículos.

Como se ha mencionado anteriormente, en muchas áreas donde se colocan sistemas de videovigilancia automáticos la aparición de un vehículo en escena se considera un motivo de alarma verdadera. Por lo tanto se debe tener muy en cuenta que una secuencia donde solamente aparece la iluminación de un automóvil será un falso positivo; mientras que si aparece también en la escena el vehículo será un verdadero positivo.

De esta circunstancia surge un nuevo obstáculo en el intento de reducir los falsos positivos debidos a variaciones lumínicas. Si reparamos nuevamente en las imágenes de las **Figuras 3-13** y **3-14** podemos ver la dificultad de distinguir si el vehículo está en escena o es sólo la iluminación. En muchas situaciones el destello de los faros puede ser tan intenso que apenas se aprecia el automóvil; y si se trabaja únicamente con la información del elemento detectado o a nivel de píxel, las características de una u otra situación son prácticamente indistinguibles.

3.4.2.3 Verdadero positivo debido a la aparición en escena de una persona.

Las grabaciones nocturnas conllevan dificultades relacionadas con el propio funcionamiento del sistema y con la baja calidad de las imágenes, que como ya se mencionó dificultan el proceso de detección de verdaderos positivos. En las regiones de la imagen más lejanas, donde no llega el infrarrojo de la cámara, resulta complicado detectar a los intrusos, ya que la sensibilidad del sistema es menor que en el caso diurno.

Además, si el intruso lleva vestimentas oscuras y se mimetiza bien con el fondo es posible que no se genere una alarma verdadera si intentamos realizar cambios en el sistema para reducir las falsas alarmas por iluminación. En la mayoría de los escenarios videovigilados no interesará perder verdaderos positivos a costa de reducir los falsos positivos.

En la **Figura 3-15** se aprecia este problema; donde la imagen superior izquierda se corresponde con la entrada de una persona. La imagen superior derecha muestra los píxeles que el sistema habría detectado como diferentes al fondo tras el proceso de background subtraction, donde vemos regiones con una masa (número de píxeles que conforman el elemento) muy pequeña e inconexas entre sí. Por último, la imagen inferior es una ampliación del área recuadrada en rojo de la primera imagen. En ésta última, intentamos destacar las capacidades del ojo humano que perciben perfectamente al intruso. Sin embargo, también se ve que al trabajar a nivel de píxel es mucho más difícil detectar a la persona, ya que las intensidades de gris son muy similares a las del fondo.

La mayoría de los sistemas comerciales descartarían esta situación en las etapas de análisis posteriores a la de sustracción de fondo, donde es muy común imponer condiciones de dimensiones mínimas, masa o conectividad entre regiones para considerar un elemento como motivo de alarma o considerar que se debe seguir su trayectoria (tracking), ver **Figura 3-1**. Esta práctica se realiza para evitar otras de las falsas alarmas mencionadas en el **apartado 3.3**, como por ejemplo la entrada de un gato, el problema de los insectos o el ruido de adquisición. Sin embargo, en este caso se estaría ignorando un caso de alarma verdadera.

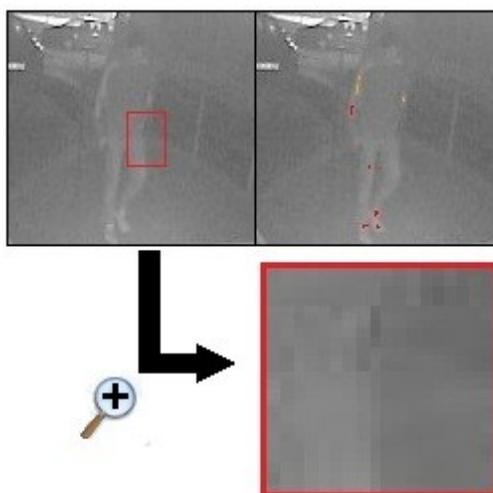


Figura 3-15: Dificultades de detección de verdaderos positivos

Otro efecto indeseable es el falso destello, que se produce cuando un intruso se acerca demasiado al objetivo. En esta situación el sistema detectará un elemento, cuyas intensidades tienen un valor saturado a blancos, debido a la reflexión de la iluminación infrarroja sobre el cuerpo del intruso, ver **Figura 3-16**. Cuando se produce un hecho de estas características, el elemento detectado tendrá muchas similitudes con uno extraído tras un mero cambio de iluminación donde no hay ningún intruso. Nuevamente y junto a los efectos mencionados en los apartados anteriores, esto no hará más que complicar la distinción de falsos positivos debidos a variaciones de luz frente a verdaderos positivos. Además, este efecto limita alguna de las posibles propuestas de solución para el problema que se trata en este proyecto.



Figura 3-16: Verdadero positivo con características similares a un falso positivo debido a un cambio de iluminación

3.5 Conclusiones

A la vista de los ejemplos anteriores ya se empiezan a extraer algunas conclusiones iniciales, como por ejemplo:

- El caso nocturno tiene mayor complejidad.
- Las luces de vehículos en movimiento suelen generar más de un elemento por culpa de los reflejos.
- Algunos cambios de iluminación son tan bruscos que alteran el funcionamiento normal de la cámara.
- Si un intruso se acerca demasiado al objetivo provocará un efecto muy similar a un cambio de iluminación, por lo tanto podemos concluir que no todo lo que parece luz realmente lo es.
- Parecería razonable colocar un foco potente de luz próximo a la cámara para que iluminara mejor la zona a vigilar y de esta forma contrarrestar los efectos de las variaciones de iluminación. Sin embargo, aparte de que esto no siempre es posible, esta solución se traduciría en algunas situaciones en un aumento de los falsos positivos provocados por la nube de insectos que atraería el foco.

A partir de la muestra de situaciones presentadas se puede realizar una clasificación sencilla de eventos que ocurren y cómo se deberían identificar realmente en el sistema:

- Cambio de iluminación natural: Falso Positivo
- On-Off de una fuente de iluminación situada en la escena: Falso Positivo
- Ráfagas de faros de vehículos (sólo la iluminación en escena): Falso Positivo
- Ráfagas de faros de vehículos (el automóvil y la iluminación en escena): Verdadero Positivo
- Intrusos y vehículos: Verdadero Positivo

Una vez ilustrado y entendido el problema que se trata en este proyecto, y haciendo uso de algunos detalles descubiertos en el análisis anterior podremos empezar a proponer algunas soluciones; lo que se realizará en el capítulo siguiente.

4 Diseño

A lo largo de este capítulo se describen los algoritmos para la reducción de falsos positivos debidos a cambios de iluminación implementados en este proyecto clasificándolos en tres categorías: Indoor, Outdoor Día y Outdoor noche.

En las siguientes subsecciones se abordará la cuestión de cómo utilizar la información extraída en la etapa de segmentación. Es decir, haciendo uso de la información que nos proporciona la herramienta de simulación se probarán los algoritmos en las regiones de la imagen donde se encuentran los elementos segmentados.

Posteriormente intentaremos extraer características relevantes que nos permitan diferenciar las falsas alarmas de las verdaderas. A partir de la información obtenida se pasará a la implementación de nuevos algoritmos que reduzcan o eliminen la tasa de falsos positivos.

4.1 Etapa preliminar

4.1.1 Segmentación de vídeos

Antes de comenzar con el diseño de los algoritmos, utilizaremos una herramienta para simular el comportamiento de un sistema de videovigilancia real de detección de movimiento. Así podremos segmentar las secuencias de vídeo de las que se dispone para realizar este trabajo.

En el **apartado 3.2** así como en la **Figura 3-1** se describían las etapas de la arquitectura de un sistema básico de videovigilancia basado en detección de movimiento que nuestra herramienta de procesamiento simula.



Figura 4-1: Modelo de fondo e imagen del instante actual

Por cada imagen de entrada al sistema, éste genera tres imágenes nuevas. Estas son la imagen segmentada, la imagen de labeling y una imagen de grouping. Además se crea un fichero con datos específicos de todos los elementos que se han detectado en cada frame.



Figura 4-2: Imágenes de salida del tras el procesamiento de una imagen

En la imagen de grouping, como se ve en la **Figura 4-2**, habrá una o varias cajas, denominadas *bounding boxes*, que delimitan aquellas regiones donde se encuentran los objetos en movimiento extraídos. El fichero de salida además de otra información, nos proporciona las coordenadas de estas cajas.

El diseño y desarrollo de algoritmos, en su mayor parte, se basará en la utilización de estos elementos detectados y encuadrados en una caja, en lugar de hacer uso de la imagen entera. Se pretende con esto centrarse solo en lo que ha cambiado en la imagen respecto al fondo, eludiendo las regiones estáticas.

Es muy común en la etapa de procesamiento de datos el uso del término *blob* (*binary large object*), para referirse a estas masas de píxeles que se corresponden con los elementos cambiantes respecto al fondo. A partir de este punto se utilizará este término como sinónimo de elemento detectado.

4.1.2 Elección del modelo de color

Si tomamos como punto de partida que un cambio de iluminación es una variación de la intensidad de los píxeles de la imagen actual respecto a la imagen de fondo y entendemos el efecto de la luz como un fenómeno aditivo, podemos enfocar nuestro diseño de algoritmos en el estudio de las dependencias entre dichas imágenes.

Debido a lo anterior escogeremos el modelo RGB, que también es un modelo aditivo, para la representación del color de las imágenes. Sin embargo y debido a que las imágenes nocturnas no tienen color se utilizarán también imágenes en niveles de gris.

4.1.3 Elección de herramientas de desarrollo

Dado que el material de trabajo de este proyecto son imágenes provenientes de secuencias de videovigilancia reales necesitaremos una herramienta para trabajar con ellas y para probar diferentes algoritmos.

Se ha elegido el paquete matemático *MATLAB* como motor para los cálculos, ya que tiene amplias librerías para el procesamiento de imágenes de diversos formatos, así como librerías matemáticas de utilidad en este proyecto. Además es un programa bien conocido en el campo de tratamiento de imágenes.

Como se menciona en el apartado anterior utilizaremos imágenes en RGB e imágenes en niveles de gris. En Matlab tendremos tres matrices R, G y B con las intensidades en rojo, verde y azul de la imagen para el caso RGB; y una matriz I con las intensidades en niveles de gris para el segundo caso. Además se dispone de funciones conversoras entre ambas representaciones y de otras funciones muy prácticas de lectura y escritura de imágenes.

Este entorno nos facilitará en buena medida la etapa de diseño de este proyecto y nos permitirá obtener y representar gráficamente los resultados que se expondrán en el capítulo siguiente.

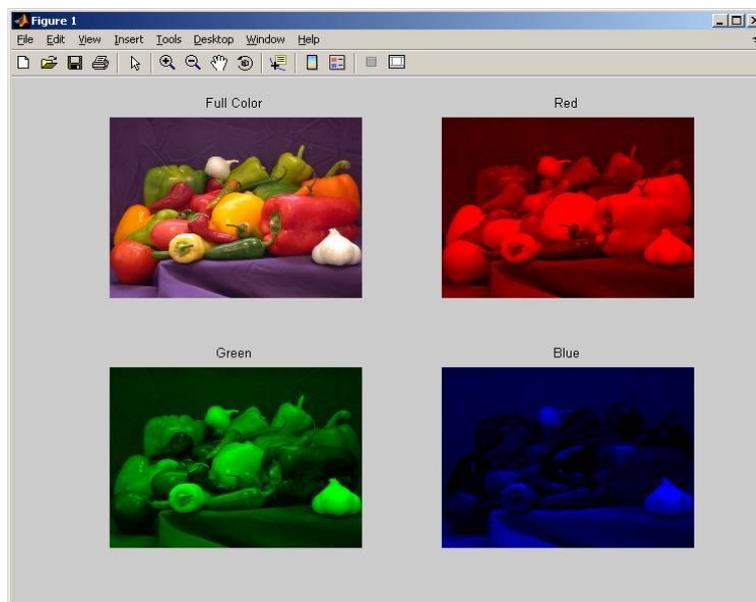


Figura 4-3: Gráfico MATLAB con imagen original y los tres planos de color

4.2 Estrategias de diseño

Como se ha mencionado en el apartado anterior, cuando sucede algún cambio de iluminación se produce una variación en la intensidad de los píxeles de la imagen actual respecto a la imagen de fondo. Al realizar el proceso de sustracción de fondo el sistema confundirá esta variación en la imagen con un objeto móvil. En cambio, al no haber ningún elemento real en la escena, las similitudes y dependencias de la imagen recibida con el fondo deberían ser muy grandes.

Se propone una serie de estrategias basadas en el estudio de la correlación entre la imagen actual y el modelo de fondo para comenzar a diseñar algoritmos que permitan reducir los falsos positivos. Esto se debe a que tras un cambio de iluminación la región segmentada tiene una gran correlación con el fondo, mientras que con un objeto esto no es así.

4.2.1 Medida de Dependencias: Correlación

Según la Real Academia Española de la lengua el término “*correlación*” se define como:

“Correspondencia o relación recíproca entre dos o más cosas o series de cosas.
Mat. Medida de la dependencia existente entre variantes aleatorias.”

En nuestro caso atenderemos a esta segunda definición matemática, ya que la estrategia propuesta se basa en el análisis de las dependencias que existen entre dos imágenes; la de fondo y la imagen del instante actual.

Dado que cuando se produce un cambio de iluminación en una escena la apariencia de la imagen es bastante similar a la imagen anterior al cambio, se intentará determinar cómo de intensa es la relación entre ellas mediante un análisis de correlación [19]. También se parte del supuesto que a priori, la relación de una imagen donde aparece un intruso o un vehículo con la imagen de fondo será menor que en el caso de una variación de luminosidad.

Actualmente, las técnicas de Correlación de Imágenes Digitales (DCI) son muy utilizadas en diversos campos que nada tienen que ver con la videovigilancia. Por ejemplo para comprobar la correcta fabricación de micro y nanosistemas, calcular la deformación de materiales, tratamiento de imágenes cartográficas, tomografías y muchas otras aplicaciones [29,30].

En MATLAB para obtener la correlación entre imágenes utilizaremos la función “corr2”. Esta función implementa la correlación de Pearson en matrices de dos dimensiones. Se expone a continuación la definición del coeficiente de correlación del producto-momento de Pearson ρ [31]:

Siendo X e Y dos variables aleatorias con medias $\mu_x = E(X)$, $\mu_y = E(Y)$
y varianzas finitas $\sigma_x^2 = \text{var}(X)$ $\sigma_y^2 = \text{var}(Y)$:

$$\rho = \text{corr}(X, Y) = \frac{\text{cov}(X, Y)}{\sigma_x \sigma_y}$$

$$\text{donde } \text{cov}(X, Y) = E[(X - \mu_x)(Y - \mu_y)] = E(XY) - \mu_x \mu_y$$

$$\text{y } -1 \leq \rho \leq +1$$

La función “corr2” aplica la definición anterior a matrices bidimensionales previamente convertidas a tipo “double” e indexadas linealmente. Esta función está especialmente desarrollada para procesar imágenes a nivel de píxel generalmente representadas por el tipo de datos “uint8” [32].

4.2.2 Arquitectura del sistema

Antes de centrarnos en una descripción de los algoritmos probados, presentaremos la arquitectura general del sistema propuesto. Esto nos permitirá tener una visión general de la extensión propuesta en este proyecto de la arquitectura básica (ver **Figura 3-1**) y así reducir los falsos positivos. Esta arquitectura se compone de cinco módulos, cuyos procesos se realizan en serie.

Como entrada tenemos una secuencia de video que fue motivo de alarma en el sistema real. Cada una de las secuencias se compone de un número N de frames, en los que se ha producido un evento concreto y único. Dicho evento será un verdadero positivo o un falso positivo.

En el primer módulo se llevará a cabo la etapa preliminar donde se realiza la segmentación de vídeos y extracción de datos, descrita en el apartado **4.1.1** de este capítulo, de forma análoga a como actuaría el sistema real que se está replicando.

El segundo módulo adecuará aquellas imágenes que están tomadas en condiciones de baja iluminación o de noche antes de que pasen al tercer módulo, cuya misión es analizar las dependencias entre las imágenes que contienen algún blob con la imagen de fondo.

En caso de disponer en el sistema real de un tiempo de guarda antes de la generación de alarmas, el módulo cuatro analizará y relacionará las dependencias de los N frames consecutivos que forman dicha ventana temporal. Una vez extraída toda la información, en el módulo 5 se tomará la decisión de si el evento que ha sucedido es una alarma falsa o es verdadera, en cuyo caso se generará la alarma pertinente.

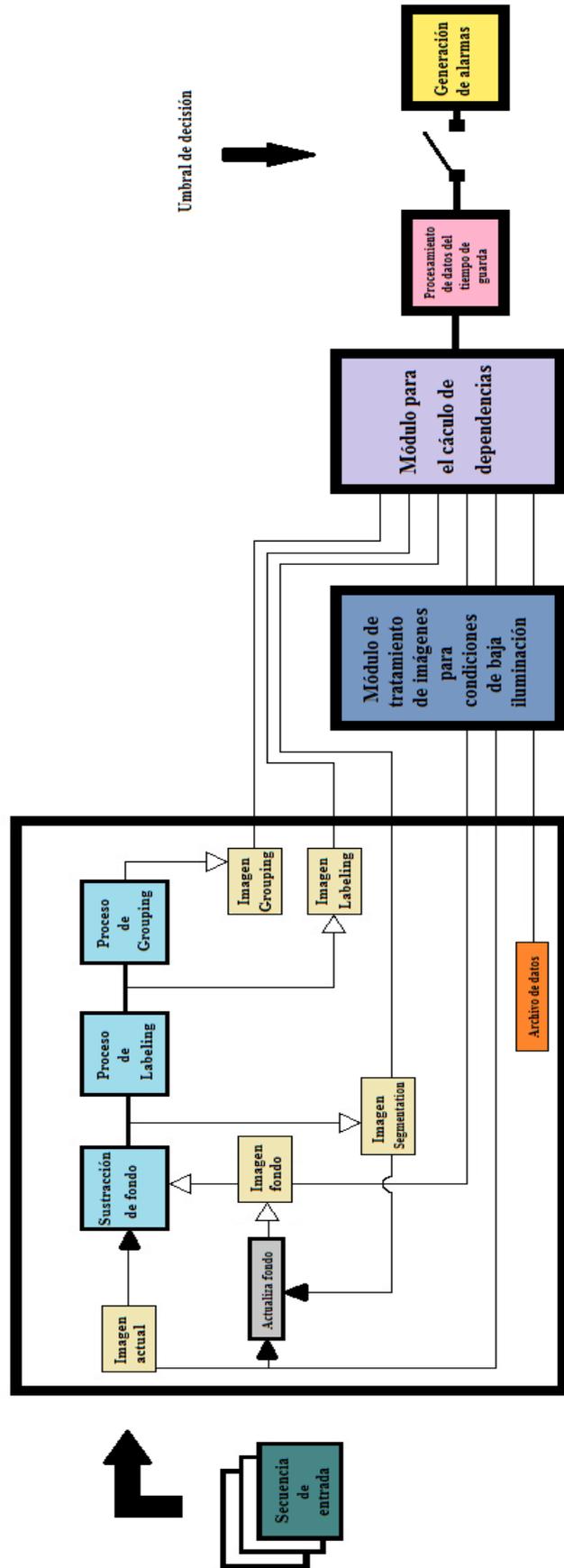


Figura 4-4: Arquitectura del sistema

4.3 Diseño de algoritmos base

En este apartado detallaremos cómo se han diseñado los algoritmos propuestos. Se describirán los procedimientos, las regiones de la imagen utilizadas en los cálculos, la información obtenida del procesamiento de las secuencias de video, descripción de parámetros necesarios, etc.

Estos algoritmos base están diseñados indistintamente para los tres casos en que se ha dividido el problema; indoor, outdoor día y outdoor noche.

4.3.1 Básico

Comenzaremos con un análisis general de los valores de correlación entre la imagen actual y el modelo de fondo que se obtienen para los casos verdaderos, donde aparecen personas o vehículos, y los falsos, donde se producen cambios de iluminación.

Para las primeras pruebas de este método utilizaremos la imagen completa. Se obtendrá un valor único de correlación para cada frame de cada una de las secuencias de prueba.

Analizaremos los primeros resultados con el fin de poder observar una distinción clara, de forma que se pueda utilizar esta característica para separar los casos verdaderos de los falsos y así generar una alarma solamente en el caso correcto.

Este método se aplicará a imágenes en RGB y se repetirá con las mismas imágenes en niveles de gris, para poder hacer una comparativa y determinar si existe una diferencia clara cuando obviamos la información de color.

Este algoritmo es una primera aproximación a las estrategias de partida de este proyecto.

4.3.2 En cuadrantes

En este segundo estudio comenzaremos restringiendo el área de la imagen utilizada para los cálculos de las dependencias. Para ello dividimos la imagen original y la de fondo en cuatro cuadrantes de forma sencilla. Esta división se hará de forma paramétrica haciendo uso de las dimensiones conocidas de las imágenes, es decir filas y columnas de las matrices.

Calcularemos la correlación en cada uno de los cuadrantes por separado. Se obtendrá en este caso cuatro valores de correlación por cada frame analizado por el algoritmo.

De nuevo analizaremos los resultados para ver si un cálculo más localizado resulta mejor que utilizar toda la imagen y se puede utilizar como característica de detección. Se intenta comprobar y apoyar la teoría de que cuando el cambio detectado se encuentra mayoritariamente en uno de los cuadrantes la dependencia de la imagen original con la de fondo será menor en dicho cuadrante y se mantendrá más o menos constante en los otros.

Otro de los fundamentos de esta prueba es la generalización de que cuando una persona o vehículo aparece en la zona videovigilada lo hace por una región de la imagen; que en muchos casos puede corresponder con uno de los cuadrantes como se ve en la **Figura 4-5**. Sin embargo los cambios de iluminación suelen ser cambios más globales afectando a todas las regiones.

Este método se aplicará a imágenes en RGB y se repetirá con las mismas imágenes en niveles de gris, al igual que en el caso básico.

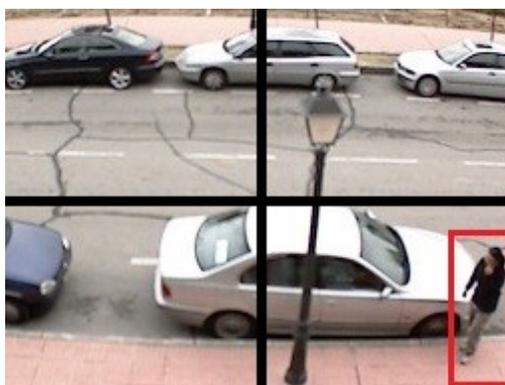


Figura 4-5: Imagen ejemplo del algoritmo base por cuadrantes

4.3.3 En bounding boxes completos

Para desarrollar el siguiente algoritmo será necesario hacer uso de los videos segmentados con la herramienta de simulación que comentamos en el apartado **4.1.1** de este capítulo. Esta herramienta nos proporciona un fichero con información de cada blob detectado en cada uno de los frames.

Continuamos con la estrategia de restringir el área de la imagen utilizada para computar la correlación entre el frame actual y el modelo de fondo que genera el sistema. Aquí lo que se propone es utilizar únicamente aquellas partes de la imagen donde el segmentador dice que se “ha movido” algo, eludiendo todos aquellos píxeles que no han sufrido cambios respecto al background.

En primer lugar, recabaremos del fichero información referente a la situación concreta del elemento o elementos detectados dentro de la imagen; así como del área de la caja o bounding box que los delimita. Con estos datos truncamos la imagen original y la de fondo en función de las posiciones de las cajas resultantes del proceso de grouping. Se procederá de igual forma para cada frame de la secuencia que se está analizando.

En este algoritmo los datos de salida que se generan son tantos como blobs se hayan extraído en el análisis de cada una de las secuencias de prueba.

Siguiendo el mismo procedimiento que en las pruebas anteriores, analizaremos los valores obtenidos para tener una referencia y así describir ciertos resultados parciales. Estos resultados parciales nos servirán para proponer una táctica diversa para abordar el problema o intentar profundizar más en la misma línea de trabajo.

Este método se aplicará a imágenes en RGB y se repetirá con las mismas imágenes en niveles de gris, para poder hacer una comparativa.

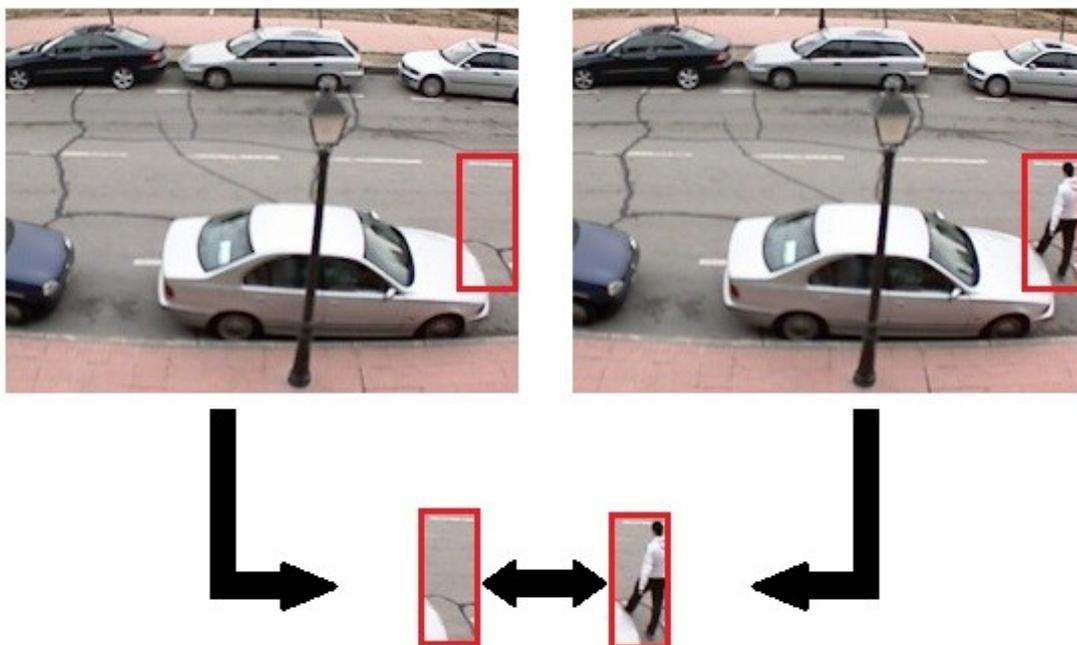


Figura 4-6: Dependencias entre el área del elemento detectado y el mismo área del modelo de fondo

4.3.4 En labeling píxeles dentro del bounding box

Como se indicó en el apartado 4.3.2 el estudio de los resultados parciales nos hace pensar que si nos centramos en evaluar la semejanza de los blobs extraídos con la misma región del fondo, en lugar de hacer uso de toda la imagen, empezamos a ver ya una gran diferencia del efecto que provoca una persona al que provocan los cambios de iluminación. Podemos encontrar diferencias significativas, por lo que se continuará analizando este tipo de características.

Al igual que en el procedimiento anterior para calcular la dependencias del blob con el fondo, truncaremos el fondo y la imagen actual quedándonos sólo con el área del bounding box. Sin embargo, para este método sólo se utilizarán aquellos píxeles que han superado el umbral fijado durante el proceso de detección de frente y que tras los procesos de labeling y grouping conforman el blob.

Para llevar a cabo las pruebas utilizaremos la imagen resultante del proceso de labeling que obtenemos tras pasar la secuencia por el simulador. En estas imágenes, al procesarlas con Matlab, podemos diferenciar de forma muy sencilla los píxeles que pertenecen al foreground de los que se consideran fondo. Esto se debe a que están marcados con valor máximo en uno de los planos de color y valor de cero en los otros dos planos.



Figura 4-7: Diferencias entre el efecto provocado por cambios de iluminación y la aparición de intrusos en escena

La **Figura 4-7** nos ilustra de nuevo cómo se comporta un sistema de videovigilancia inteligente ante casos de alarma verdadera y casos de falsas alarmas. Además, también muestra que restringir el área de cómputo de correlación únicamente a aquellos píxeles que constituyen el blob puede llevarnos a encontrar una característica realmente diferenciadora y que mediante un umbral ayude a reducir los falsos positivos debidos a cambios de iluminación.

Al igual que en los algoritmos anteriores, se analizarán los resultados obtenidos al aplicar este procedimiento antes de proseguir con el diseño de otra táctica. Asimismo, al haber probado con antelación otros dos métodos sobre la misma colección de videos podremos realizar una comparativa viendo las tasas de reducción de falsos positivos de cada uno de ellos.

Este método se aplicará a imágenes en RGB y se repetirá con las mismas imágenes en niveles de gris.

4.3.5 Adición de un tiempo de guarda antes de la generación de alarma

En algunos casos puede resultar beneficioso posponer unos segundos la generación de una alarma en el sistema real. De esta manera disponemos de un tiempo de guarda antes de tomar la decisión de si un cambio se ha debido a variaciones de luz o la entrada de un extraño.

Esto se debe a que en algunas situaciones, sobre todo en escenarios indoor donde las fuentes de luz son artificiales, un cambio de iluminación provoca una alteración muy brusca y tras unos frames la cámara se autoregula en cierta medida a las nuevas condiciones lumínicas.



Figura 4-8: Autoregulación de la cámara tras un cambio brusco de iluminación en escenario indoor

Por este motivo se propone un algoritmo donde se aplican las técnicas de análisis de dependencias en ciertas regiones de la imagen, descritas en los apartados anteriores, pero teniendo en cuenta lo sucedido en el tiempo de guarda. En este proceso se analiza una ventana temporal de W frames posteriores al primer frame donde se detectó un cambio que generó una alarma.

En función de lo que suceda en esas imágenes consecutivas se decidirá si dar alarma u obviarla. En este proyecto se propone un método simple por mayoría; si de los W frames sucesivos $W/2+1$ frames están por debajo del umbral de dependencias que utilizamos como característica diferenciadora se producirá una alarma. El parámetro W se ajustará a la tasa de imágenes por segundo con la que opera el sistema real. Se debe tener en cuenta para este ajuste que si el tiempo de espera es elevado se podría perder algún verdadero positivo, lo que no es asumible.

Siguiendo la misma metodología de trabajo, analizaremos y compararemos los resultados obtenidos al aplicar un tiempo de guarda antes de la generación de alarmas, para ver los posibles beneficios en el sistema real.

4.4 Diseño de algoritmos mejorados para condiciones de baja iluminación o nocturnidad

Después de analizar y estudiar todos los resultados recabados tras la aplicación de los algoritmos propuestos anteriormente, para el caso outdoor noche necesitamos adaptar en cierta medida las imágenes, antes de poder aplicarlos si queremos alcanzar una tasa de reducción de falsos positivos para la noche equiparable al caso diurno o al de indoor.

Las imágenes nocturnas tienen regiones con diferentes grados de iluminación, debidas a farolas, luces de edificios e incluso debido al IR de la propia cámara y esto puede distorsionar los valores de la correlación que se generan tras la comparativa de la imagen actual y el modelo de fondo. Por este motivo se propone eliminar este efecto mediante diferentes técnicas.

4.4.1 Estudio de otras regiones de cálculo

Siguiendo en la línea de pruebas anteriores se ha intentado de forma inversa, ampliar el área de cálculo a un número mayor de píxeles. Por un lado se han realizado los mismos cálculos de dependencias sobre todos los píxeles segmentados de la imagen completa y también sobre los píxeles segmentados dentro del bounding box, aunque no formen parte del blob. Para ello se utiliza la imagen de salida del proceso de segmentación del simulador, en lugar de la imagen resultante del proceso de labeling como hacíamos antes.

La realización de esta prueba se debe a que en condiciones nocturnas quizá interese no concentrar tanto la región de cálculo de dependencias debido a que el efecto de los cambios de iluminación es mucho más brusco afectando a generalmente a toda la imagen. Además las imágenes tomadas de noche proporcionan menos información que cuando tratábamos imágenes diurnas o de interiores.

4.4.2 Eliminar gradiente de iluminación (versión1)

En este nuevo intento trataremos de eliminar el gradiente de iluminación que existe tanto en la imagen actual como en el modelo de fondo que genera el sistema. Se trata de eliminar el efecto que provoca la iluminación existente en adición a la que se debe al nuevo cambio de iluminación.

Para ello se calcula el plano de iluminación de cada imagen, frame actual y fondo, y luego se corrige la iluminación de dichas imágenes de la forma:

Imagen actual – plano iluminación propio

Imagen fondo – plano iluminación propio

Esto se aplicará de manera global sobre toda la imagen de fondo y sobre toda la imagen actual de forma independiente antes de realizar el cómputo de correlación entre ambas.

4.4.3 Eliminar gradiente de iluminación (versión2)

Este método resulta de una simple modificación del método de corrección anterior. En este caso se calcula para la imagen de fondo su plano de iluminación y luego se corregirá la iluminación de la imagen actual y del fondo de la forma:

Imagen actual – plano iluminación fondo

Imagen fondo – plano iluminación fondo

Asimismo, se aplicará de manera global sobre toda la imagen de fondo y sobre toda la imagen actual de forma independiente antes de realizar el cómputo de correlación entre ambas.

4.4.4 Eliminar gradiente de iluminación en el bounding box (versión1)

Los resultados obtenidos tras las dos correcciones anteriores mostrarán cierta mejora en las tasas de reducción de falsos positivos; por lo tanto continuaremos ensayando con variantes de las propuestas anteriores.

Al igual que hiciéramos en el apartado **4.3.3** truncamos la imagen de fondo y la imagen original al área del bounding box de cada uno de los blobs extraídos. Se calcula para cada nueva imagen reducida su plano de iluminación de forma independiente y luego se corrige la iluminación de dicha imagen de la forma:

Imagen actual truncada – plano iluminación propio

Imagen fondo truncada – plano iluminación propio

4.4.5 Eliminar gradiente de iluminación en el bounding box (versión2)

En este nuevo intento modificaremos el método propuesto en el apartado 4.4.3 de la misma manera que en el caso anterior. Para ello truncamos la imagen de fondo y la imagen original al área del bounding box del blob. Se calcula para la imagen reducida de fondo su plano de iluminación y luego se corregirá la iluminación de las respectivas cajas extraídas en cada frame de la forma:

Imagen actual truncada – plano iluminación fondo

Imagen fondo truncada – plano iluminación fondo



Figura 4-9: Imagen original e imagen tras la corrección del gradiente de iluminación

4.4.6 Eliminación de píxeles con valores saturados a blancos dentro del bounding box

Como se describió en el capítulo tres de este proyecto, los cambios bruscos de iluminación, como por ejemplo los fogonazos de los coches generan regiones con blancos saturados debido a que la luz impacta sobre superficies reflectantes.

Por este motivo se propone eliminar del cómputo aquellos píxeles cuyos valores de intensidad están próximos al valor máximo y contenidos en la caja, con el fin de reducir un posible error en el posterior cálculo de la correlación.

En la **Figura 4-10** se expone una imagen que evidencia el efecto que provocan los faros de los vehículos en grabaciones nocturnas y que motiva esta técnica de corrección. Inmediatamente después tenemos el histograma de la imagen, donde se aprecia un pico en los blancos saturados. El número de los píxeles con intensidad 255 es del orden de $5 \cdot 10^4$, mientras que en valores oscuros el número máximo asciende a $1.3 \cdot 10^4$, siendo la imagen original de dimensiones $576 \cdot 720$.

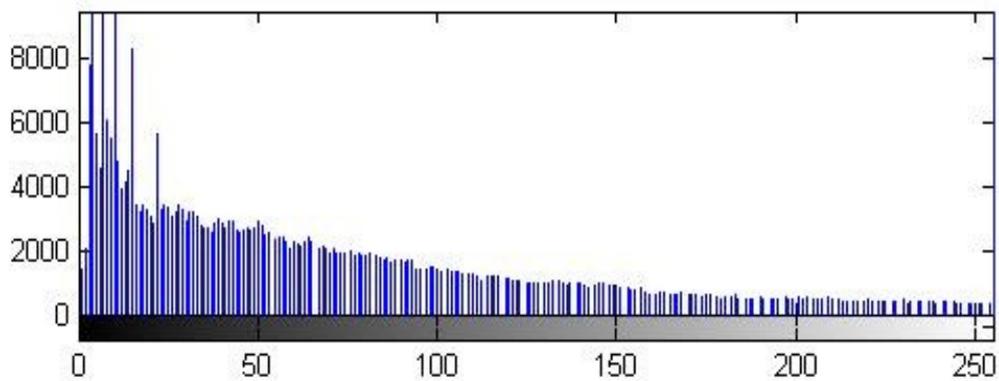


Figura 4-10: Imagen representativa del efecto de la iluminación de vehículos sobre el escenario y su histograma

4.4.7 Cambios de escala

Debido a los problemas que existen en escenarios outdoor de noche, sobre todo por los destellos de los faros de vehículos como ya se ha mencionado, sigue siendo necesario intentar reducir estos efectos antes de comprobar las dependencias entre la imagen actual y la imagen de fondo. Siguiendo la línea del método de adaptación anterior, se propone aplicar un cambio de escala donde se acentúen una parte de los valores de intensidad.

Para lo anterior se realizarán pruebas con dos tipos diferentes de cambios de escala, que resaltan o bien los valores grises, o bien los blancos:

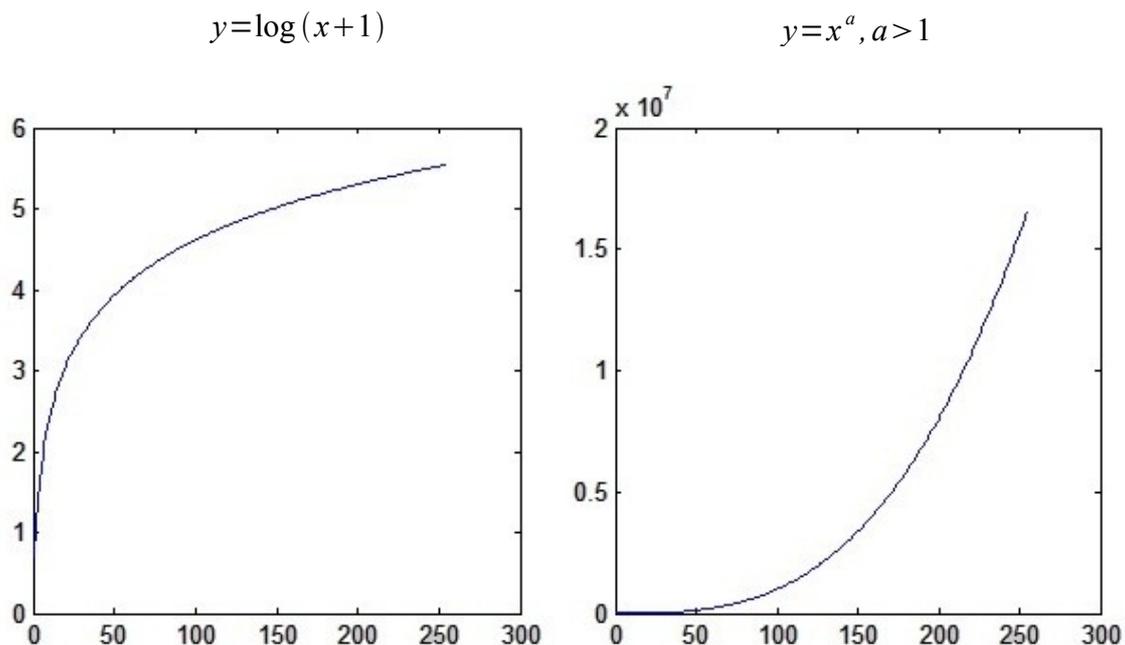


Figura 4-11: Escala logarítmica y escala exponencial (a=3)

4.4.8 Cálculo de intensidad media de los píxeles vecinos

Este método calculará para cada píxel la media de intensidad en función de las intensidades de los píxeles vecinos. Se utilizará una máscara cuadrada de $m \times m$ píxeles, siendo m impar.

Este procedimiento se llevará a cabo sin relleno de ceros (*padding*) en los extremos. Es decir, en las esquinas de la imagen solo computarán los píxeles dentro de la imagen y se dividirá en este caso por ese número de valores en vez de dividir entre el tamaño de la máscara. Con este tipo de cálculos obtendremos cierta información de texturas y bordes. Estos valores medios se restarán tanto al frame actual como al modelo de fondo, Se trata de una medida de corrección de las imágenes antes de aplicar cualquiera de los algoritmos de cálculo de dependencias descritos en el apartado 4.3.

4.4.9 Estudio de otros factores

Debido a que las variaciones de iluminación provocan cambios en la imagen muy característicos, existen una serie de parámetros que en combinación con los algoritmos anteriores pueden ayudar a reducir los falsos positivos.

Muchas de estas características aparecen de forma general en secuencias obtenidas en diferentes lugares, pero no son aplicables en todo tipo de escenarios. Además dependen en gran medida de las necesidades y requisitos de seguridad, es decir, de qué se quiere detectar en ese área concreta. Sin embargo son métodos de bajo coste computacional y resultan muy útiles en aquellos casos donde se pueden emplear.

Algunos de estos parámetros característicos son por ejemplo la masa del blob, las dimensiones del bounding box respecto al área total de la imagen, orientación vertical versus horizontal de la caja, etc.

En la **Figura 4-12**, compuesta por tres secciones del modelo de fondo, la imagen de grouping tras apagar la luz y la imagen de grouping tras la aparición de una persona; se aprecian estas singularidades que tienen los blobs de iluminación respecto a los de intrusos.



Figura 4-12: Comparativa de masa, dimensiones y proporciones de un blob causante de falsa alarma con uno de alarma verdadera

5 Pruebas y resultados experimentales

En este capítulo se muestran los resultados experimentales de los algoritmos que se han desarrollado para mejorar los sistemas de videovigilancia existentes, ante los cambios de iluminación, ya descritos en el **Capítulo 4**.

Se analizarán los resultados obtenidos por medio de curvas ROC. A partir de ellos compararemos las diferentes propuestas. Con estos datos se extraerán las conclusiones de las diferentes soluciones en las que se ha trabajado en este proyecto.

5.1 Curvas ROC

5.1.1 Introducción

Las curvas ROC (*Receiver Operating Characteristics*) son una técnica útil para analizar sistemas de detección y visualizar su comportamiento. Este tipo de gráficos se han utilizado en los últimos años en Teoría de Detección de Señales [33] para representar el equilibrio entre tasas de éxito y tasas de falsas alarmas en clasificadores. Se ha extendido el uso de los análisis ROC a diversas áreas; como para visualizar y analizar el comportamiento del sistema de diagnósticos, donde la comunidad médica ha contribuido notablemente a la literatura acerca del uso de curvas ROC [34].

También en el campo de *Machine Learning* y minería de datos se ha demostrado la valía de las curvas ROC para evaluar y comparar algoritmos [35], donde en los últimos años se ha generalizado su uso como herramienta de representación gráfica. Esto se debe a algunas propiedades especialmente útiles cuando se trabaja con distribuciones sesgadas o cuando los diferentes errores de clasificación no tienen el mismo coste.

Muchos de los trabajos donde se mencionan las Curvas ROC, simplemente hacen una breve descripción de esta técnica, ya que conceptualmente estos gráficos son bastante simples. Sin embargo existen algunos conceptos menos obvios o errores comunes que aparecen durante la práctica, por lo que resultará útil definir algunos conceptos.

5.1.2 Clasificadores binarios

Un modelo de clasificación es una función que permite decidir cuáles de un conjunto de instancias están en un grupo o en otro, donde el límite del clasificador entre cada clase debe determinarse por un valor umbral.

Si consideramos el problema ocasionado por los cambios de iluminación en los sistemas de videovigilancia inteligentes, tendremos un problema de predicción de dos clases o clasificación binaria, en la que los resultados se etiquetarán como “LUZ” o “NO LUZ”.

En función de lo anterior, hay cuatro posibles resultados a partir de un clasificador binario como el que utilizaremos para evaluar los algoritmos propuestos. Si el resultado de una predicción es p' y el valor real es también p , entonces se conoce como un Verdadero Positivo (VP); sin embargo si el valor real es n entonces se conoce como un Falso Positivo (FP). De igual modo, tenemos un Verdadero Negativo (VN) cuando tanto la predicción como el valor real son n , y un Falso Negativo (FN) cuando el resultado de la predicción es n' pero el valor real es p . Estos cuatro posibles resultados se suelen formular mediante una tabla de contingencia o matriz de confusión como la que sigue [36]:

		Valor en la realidad		total
		p	n	
Predicción	p'	Verdaderos Positivos	Falsos Positivos	P'
	n'	Falsos Negativos	Verdaderos Negativos	N'
total		P	N	

Describiremos a continuación algunos de los términos anteriores, así como otros derivados de ellos:

Verdadero Positivo (VP): clasificación con éxito

Verdadero Negativo (VN): rechazo correcto

Falso Positivo (FP): Error Tipo I o falsa alarma

Falso Negativo (FN) : Error Tipo II

Razón de Verdaderos Positivos (VPR): razón de éxitos, también denominado sensibilidad. Viene dado por la siguiente expresión:

$$VPR = \frac{VP}{VP + FN}$$

Razón de Falsos Positivos (FPR): razón de falsas alarmas, cuya expresión es:

$$FPR = \frac{FP}{FP + VN}$$

5.1.3 El espacio ROC

De la tabla de contingencia se pueden obtener varias medidas de evaluación, sin embargo para dibujar una curva ROC sólo serán necesarias las tasas de Verdaderos Positivos (VPR) y de falsos positivos (FPR).

La VPR mide hasta qué punto un clasificador es capaz de detectar los casos positivos correctamente, de entre todos los casos positivos disponibles durante la prueba. En este proyecto se considerarán casos positivos aquellas secuencias de video donde aparezcan intrusos o vehículos y que generaron una alarma verdadera.

La FPR define cuántos resultados positivos son incorrectos de entre todos los casos negativos disponibles durante la prueba. En nuestra evaluación se considerarán casos negativos aquellas secuencias de video donde se ha producido algún tipo de cambio de iluminación y que fue causa de una falsa alarma.

Un espacio ROC se define por VPR y FPR como eje de ordenadas y abscisas respectivamente, y representa los intercambios entre verdaderos positivos y falsos positivos. Cada resultado de la predicción de una instancia de la matriz de confusión representa un punto en el espacio ROC [36].

El método con el mejor resultado posible de predicción se situaría en un punto en la esquina superior izquierda, o coordenada (0,1) del espacio ROC, lo que se traduce en ningún falso negativo y ningún falso positivo. Este punto (0,1) es también llamado una clasificación perfecta. Por el contrario, una clasificación totalmente aleatoria daría un punto a lo largo de la línea diagonal, que se llama también línea de no-discriminación, desde el extremo izquierdo hasta la esquina superior derecha; ver **Figura 5-1**.

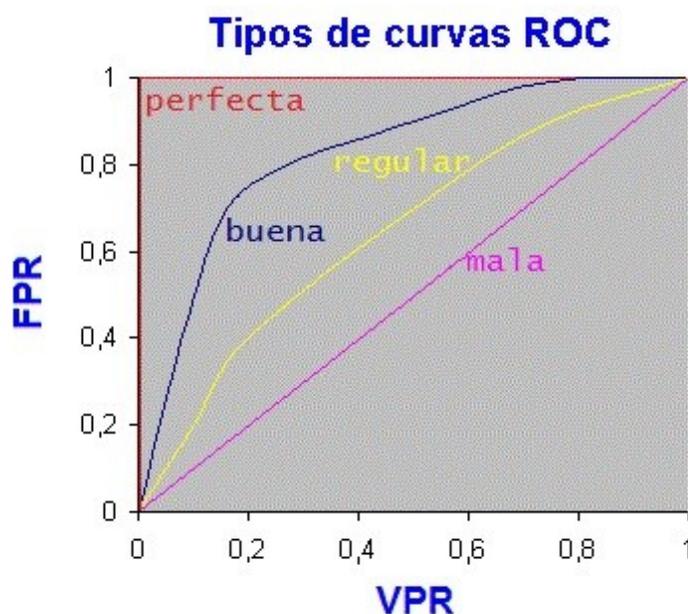


Figura 5-1: Ejemplos de curvas ROC

5.2 Diseño de pruebas experimentales

Con el objetivo de obtener las tasas de reducción de falsos positivos de cada uno de los algoritmos propuestos en el capítulo anterior, se han estudiado y ejecutado diversas pruebas experimentales.

Recordamos que el problema que se trata en este proyecto se dividió en tres categorías, caso Indoor, caso Outdoor diurno y caso Outdoor nocturno. Se han calculado curvas ROC mediante diferentes procedimientos para cada uno de los casos haciendo uso de un número amplio de secuencias reales, cuya muestra se pudo ver en el **Capítulo 3** de este documento.

Se han realizado cuatro tipos diferentes de pruebas cuyo resultado se expresará mediante un gráfico ROC. En cada uno de estos gráficos estarán representadas varias curvas correspondientes a los diferentes algoritmos. De esta forma se podrá comparar el rendimiento de cada uno de los métodos propuestos.

Antes de entrar en especificaciones, aclararemos que las secuencias de la base de datos de la que se dispone se corresponden con eventos independientes, o bien cambios de iluminación (FP) o bien alarmas verdaderas, como la aparición de intrusos (VP). Se parte por tanto de secuencias etiquetadas a priori mediante la variable booleana *INTRUSIÓN*; de esta forma se comparará el resultado de la clasificación de cada algoritmo con dicha etiqueta.

Se detalla a continuación la forma en la que se han calculado, para cada procedimiento, las razones de Verdaderos Positivos (VPR) y de falsos positivos (FPR) que posteriormente se representarán del mismo modo que en la **Figura 5-1**.

5.2.1 Procedimiento 1, clasificación del evento en cada frame de la secuencia (P1)

Para realizar esta prueba se analizará cada uno de los frames de cada una de las secuencias de forma independiente.

Como se aprecia en la muestra de imágenes del **Capítulo 3** el segmentador detectará cualquier cambio que se haya producido en la escena. En cada frame aparecerán uno o varios elementos (blobs), sobre los que se aplicarán los diferentes algoritmos de medida de dependencias con la imagen de fondo. Por lo tanto, cuando aplicamos alguno de los algoritmos desarrollados a una secuencia, lo que obtenemos es una matriz con los valores de correlación de los blobs que aparecen en cada uno de los frames. Se compararán estos valores con un umbral tal que:

$$\text{Si } R_{corr} > Th \rightarrow \text{Cambio de iluminación}$$

$$\text{Si } R_{corr} < Th \rightarrow \text{Intruso o vehículo}$$

Se incrementarán los valores de VP, FP, VN o FN en cada frame. Lo que significa comparar la clasificación realizada en función del umbral con la etiqueta de cada secuencia (*INTRUSIÓN*) tal que:

$$\text{Si } R_{corr} < Th \text{ e } INTRUSION = 1 \Rightarrow VP = VP + 1;$$

$$\text{Si } R_{corr} < Th \text{ e } INTRUSION = 0 \Rightarrow FP = FP + 1;$$

$$\text{Si } R_{corr} > Th \text{ e } INTRUSION = 0 \Rightarrow VN = VN + 1;$$

$$\text{Si } R_{corr} > Th \text{ e } INTRUSION = 1 \Rightarrow FN = FN + 1.$$

Se repetirá este proceso para umbrales (*Th*) comprendidos en el rango:

$$Th \in [-1, 1] \text{ donde } \Delta Th = 0.1$$

A partir de estos datos se calcularán las razones de Verdaderos Positivos (VPR) y de falsos positivos (FPR), para cada valor del umbral y para cada secuencia de pruebas. Por último se calculará el promedio de las razones obtenidas después de procesar todas las secuencias, siendo estos valores los que se representarán en la curva ROC.

Los gráficos de curvas ROC calculadas mediante este procedimiento se denotarán mediante el siguiente título:

“P1 ROC: alarma por frames”

5.2.2 Procedimiento 2, clasificación del evento en cada frame de la secuencia, analizando W frames sucesivos (P2)

Para realizar esta prueba se analizará cada uno de los frames de cada una de las secuencias, al igual que en el procedimiento P1, descrito en el apartado anterior.

Sin embargo debemos recordar el **apartado 4.3.5**, donde se mencionaba el beneficio de posponer unos segundos la generación de una alarma en el sistema real. De esta manera disponíamos de un tiempo extra de análisis antes de tomar la decisión de si un cambio se había debido a variaciones de luz o la entrada de un extraño.

Se incrementarán los valores de VP, FP, VN o FN en cada frame, pero teniendo en cuenta la clasificación en los W frames consecutivos al que se está analizando. En este caso compararemos la clasificación realizada en función del umbral con la etiqueta de cada secuencia (*INTRUSIÓN*) tal que:

Si $R_{corr} < Th$, $R_{corr} < Th$ en $\frac{W}{2}+1$ de los W frames siguientes e $INTRUSION = 1 \Rightarrow$
 $\Rightarrow VP = VP + 1 ;$

Si $R_{corr} < Th$, $R_{corr} < Th$ en $\frac{W}{2}+1$ de los W frames siguientes e $INTRUSION = 0 \Rightarrow$
 $\Rightarrow FP = FP + 1 ;$

Si $R_{corr} < Th$, $R_{corr} > Th$ en $\frac{W}{2}+1$ de los W frames siguientes e $INTRUSION = 0 \Rightarrow$
 $\Rightarrow VN = VN + 1 ;$

Si $R_{corr} > Th$, $R_{corr} > Th$ en $\frac{W}{2}+1$ de los W frames siguientes e $INTRUSION = 1 \Rightarrow$
 $\Rightarrow FN = FN + 1 ;$

Se repetirá este proceso para umbrales (Th) comprendidos en el rango:

$$Th \in [-1, 1] \text{ donde } \Delta Th = 0.1$$

A partir de estos datos se calcularán las razones de Verdaderos Positivos (VPR) y de falsos positivos (FPR), para cada valor del umbral y para cada secuencia de pruebas. Por último se calculará el promedio de las razones obtenidas después de procesar todas las secuencias, siendo estos valores los que se representarán en la curva ROC, de forma análoga al procedimiento 1.

Los gráficos de curvas ROC calculadas mediante este procedimiento se denotarán mediante el siguiente título:

“P2 ROC-W: alarma por frames, comprobando W frames posteriores”

5.2.3 Procedimiento 3, clasificación del evento dada una secuencia completa (P3)

El cálculo de ROC' s mediante este procedimiento se basa en la clasificación de una secuencia completa como un evento que generará alarma o que se obviará.

Al final del procesamiento, cada secuencia tendrá una nueva etiqueta; $Alarma=0$ o $Alarma=1$ en función del umbral. Se parte para todos los casos de $Alarma=0$, situación que se mantendrá mientras no se encuentre ningún elemento en la secuencia cuyo valor R_{corr} sea menor que el umbral con el que se está trabajando. Es decir:

$$\begin{aligned} & Si \ R_{corr}(t) > Th \ \forall t \Rightarrow Alarma = 0 ; \\ & Si \ R_{corr}(t=t_0) < Th \Rightarrow Alarma = 1 ; \text{ siendo } t_0 \text{ un instante cualquiera de la secuencia} \end{aligned}$$

Se comprobará la correspondencia de dicha etiqueta de clasificación de evento, con la etiqueta inicial de cada secuencia a la que denominamos INTRUSIÓN. Según lo anterior, lo que haremos será incrementar los valores de VP, FP, VN o FN tal que:

Si INTRUSION=1 y Alarma=1 $\Rightarrow VP=VP+1$;
Si INTRUSION=0 y Alarma=1 $\Rightarrow FP=FP+1$;
Si INTRUSION=0 y Alarma=0 $\Rightarrow VN=VN+1$;
Si INTRUSION=1 y Alarma=0 $\Rightarrow FN=FN+1$.

Se repetirá este proceso para umbrales (Th) comprendidos en el rango:

$$Th \in [-1, 1] \text{ donde } \Delta Th=0.1$$

A partir de los valores de la matriz construida con los datos anteriores se calculará la razón de Verdaderos Positivos (VPR) y la de falsos positivos (FPR), para cada valor del umbral y para el conjunto de las secuencias de pruebas.

Los gráficos de curvas ROC calculadas mediante este procedimiento se denotarán mediante el siguiente título:

“P3 ROC-Sq: alarma en algún frame de la secuencia”

5.2.4 Procedimiento 4, clasificación del evento dada una secuencia completa, analizando W frames sucesivos (P4)

El cálculo de ROC' s mediante este procedimiento es análogo al procedimiento 3 (P3), salvo que para decidir si un evento se clasifica como alarma o se ignora, se analizará lo sucedido en W frames posteriores de forma similar al procedimiento 2 (P2).

Al igual que antes, se parte para todos los casos de $Alarma=0$, situación que se mantendrá mientras no se encuentre ningún elemento en la secuencia cuyo valor R_{corr} sea menor que el umbral con el que se está trabajando. Si encontramos en un frame un blob cuyo valor R_{corr} no cumple la condición anterior se comprobarán los W frames siguientes antes de establecer que $Alarma=1$. Es decir:

Si $R_{corr}(t) > Th \quad \forall t \Rightarrow Alarma=0$;
Si $R_{corr}(t=t_0) < Th$ y $R_{corr} < Th$ en $\frac{W}{2}+1$ de los W frames siguientes $\Rightarrow Alarma=1$;
Si $R_{corr}(t=t_0) < Th$ y $R_{corr} > Th$ en $\frac{W}{2}+1$ de los W frames siguientes $\Rightarrow Alarma=0$;
siendo t_0 un instante cualquiera de la secuencia

Se comprobará la correspondencia de dicha etiqueta de clasificación de evento, con la etiqueta inicial de cada secuencia (*INTRUSIÓN*) de forma análoga al procedimiento 3 . Se incrementarán los valores de VP, FP, VN o FN tal que:

Si $INTRUSION=1$ y $Alarma=1 \Rightarrow VP=VP+1$;
 Si $INTRUSION=0$ y $Alarma=1 \Rightarrow FP=FP+1$;
 Si $INTRUSION=0$ y $Alarma=0 \Rightarrow VN=VN+1$;
 Si $INTRUSION=1$ y $Alarma=0 \Rightarrow FN=FN+1$.

Se repetirá este proceso para umbrales (Th) comprendidos en el rango:

$$Th \in [-1, 1] \text{ donde } \Delta Th=0.1$$

A partir de los valores de la matriz construida con los datos anteriores se calculará la razón de Verdaderos Positivos (VPR) y la de falsos positivos (FPR), para cada valor del umbral y para el conjunto de las secuencias de pruebas.

Los gráficos de curvas ROC calculadas mediante este procedimiento se denotarán mediante el siguiente título:

“P4 ROC-SqW: alarma en algún frame de la secuencia, comprobando W frames posteriores”

5.3 Resultados caso Indoor

5.3.1 Secuencias utilizadas para las pruebas

Se muestra en la tabla siguiente la clasificación inicial de las secuencias capturadas en escenarios indoor. Dichas secuencias pertenecen a tres escenarios interiores diferentes. Siguiendo la misma nomenclatura que se utilizó en el apartado anterior para la definición de procedimientos se denotarán con $INTRUSIÓN=1$ aquellas secuencias consideradas alarmas verdaderas por la aparición de uno o varios intrusos y con $INTRUSIÓN=0$ las secuencias donde se ha producido un cambio de iluminación.

# Total Secuencias	$INTRUSIÓN=1$	$INTRUSIÓN=0$
32	18	14

Tabla 1: Clasificación inicial de secuencias Indoor

5.3.2 Curvas ROC

Para este caso se evaluará el funcionamiento de los algoritmos “En cuadrantes”, “En bounding boxes completos” y “En labeling píxeles dentro del bounding box”, descritos en los apartados 4.3.2, 4.3.3 y 4.3.4 respectivamente. Se representarán de forma conjunta en un mismo gráfico para obtener los resultados de cada algoritmo y poder compararlos entre sí.

Se describe a continuación la nomenclatura de las diferentes curvas que se representarán en los gráficos ROC.

- **CORR cuadrantes:** curva ROC obtenida a partir de los datos extraídos tras evaluar las dependencias en los cuatro cuadrantes y comparando el valor mínimo con el valor del umbral (ver 4.3.2). En este caso se analizan imágenes en niveles de gris.
- **CORR cuadrantes rgb:** curva análoga a la anterior con la diferencia de que en este caso la correlación se calculará de forma independiente en cada plano de color, siendo el valor final de R_{corr} la media de R_{corr}_r , R_{corr}_g y R_{corr}_b .
- **CORR b-box:** curva ROC obtenida a partir de los datos extraídos tras evaluar la correlación del área de los bounding boxes que delimitan los elementos detectados con la misma área del modelo de fondo y comparándolos con el valor umbral (ver 4.3.3). En este caso se analizan imágenes en niveles de gris.
- **CORR b-box rgb:** curva análoga a la anterior donde el valor de la correlación R_{corr} se calculará como el valor medio de R_{corr}_r , R_{corr}_g y R_{corr}_b .
- **CORR lab-pixel:** curva ROC obtenida a partir de los datos extraídos tras procesar las secuencias y evaluar las dependencias de los elementos detectados con el modelo de fondo en los labeling píxeless (ver 4.3.4). En este caso se analizan imágenes en niveles de gris.
- **CORR lab-pixel rgb:** curva análoga a la anterior donde la correlación se calculará de forma independiente en cada plano de color, siendo el valor final de R_{corr} la media de R_{corr}_r , R_{corr}_g y R_{corr}_b calculados para cada uno de los planos.

Esta notación para las curvas representadas se mantendrá en todas las figuras que se incluyen en este capítulo.

Se muestran a continuación las gráficas del caso Indoor de los procedimientos **P1, P2, P3 y P4**, en el orden indicado. Para los procedimientos 3 y 4 las curvas se desglosarán en dos gráficas diferentes; una para las curvas calculadas con imágenes en niveles de gris y otra para las de RGB para evitar el solape de algunas de ellas.

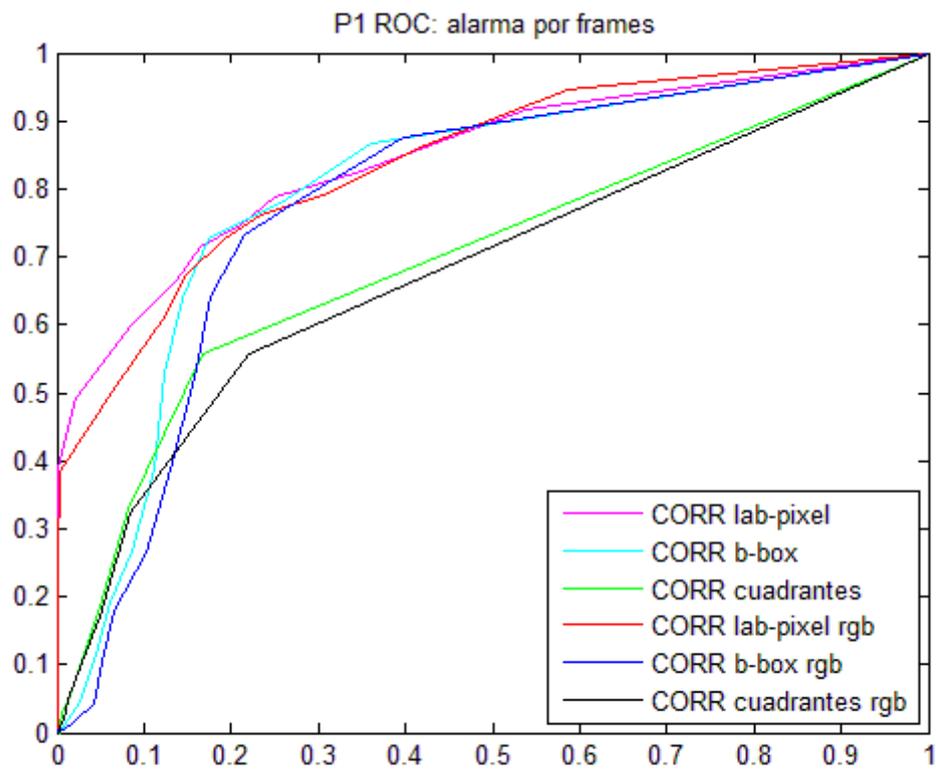


Figura 5-2: Caso Indoor, curvas ROC P1

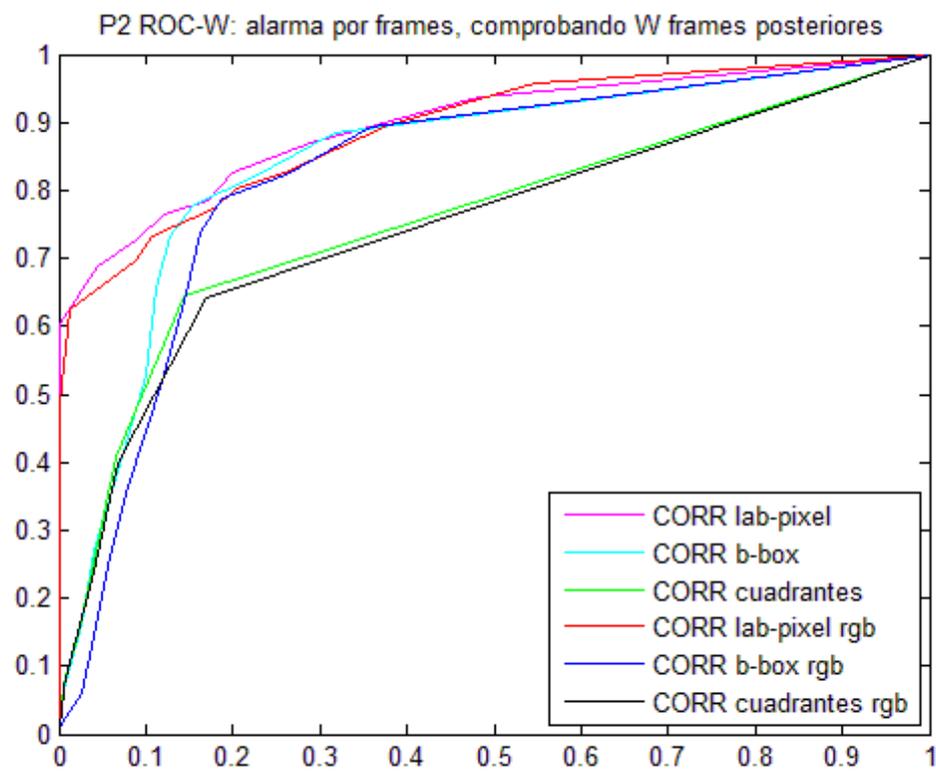


Figura 5-3: Caso Indoor, curvas ROC P2

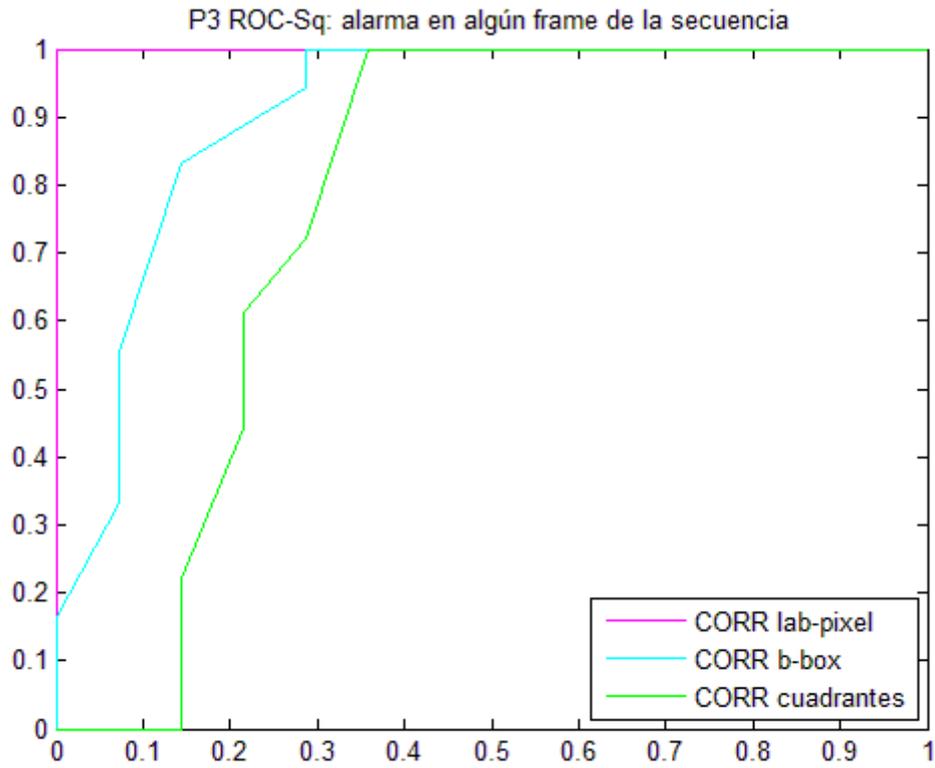


Figura 5-4: Caso Indoor, curvas ROC P3
(La ROC correspondiente a lab-pixel es el cuadrado de área 1)

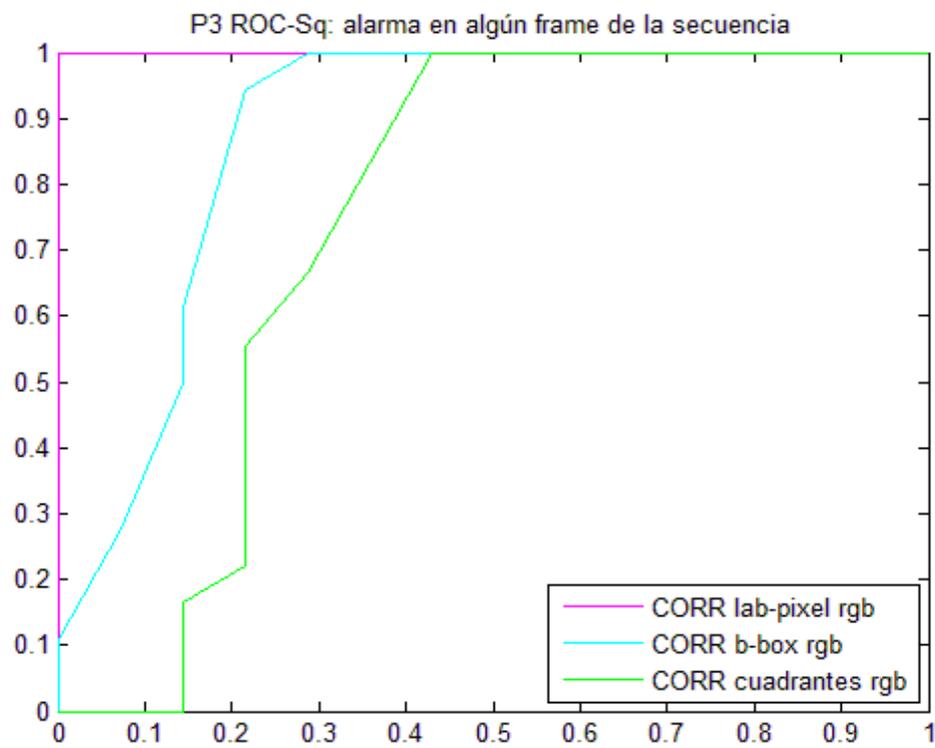


Figura 5-5: Caso Indoor, curvas ROC P3
(La ROC correspondiente a lab-pixel rgb es el cuadrado de área 1)



P4 ROC-SqW: alarma en algún frame de la secuencia, comprobando W frames posteriores

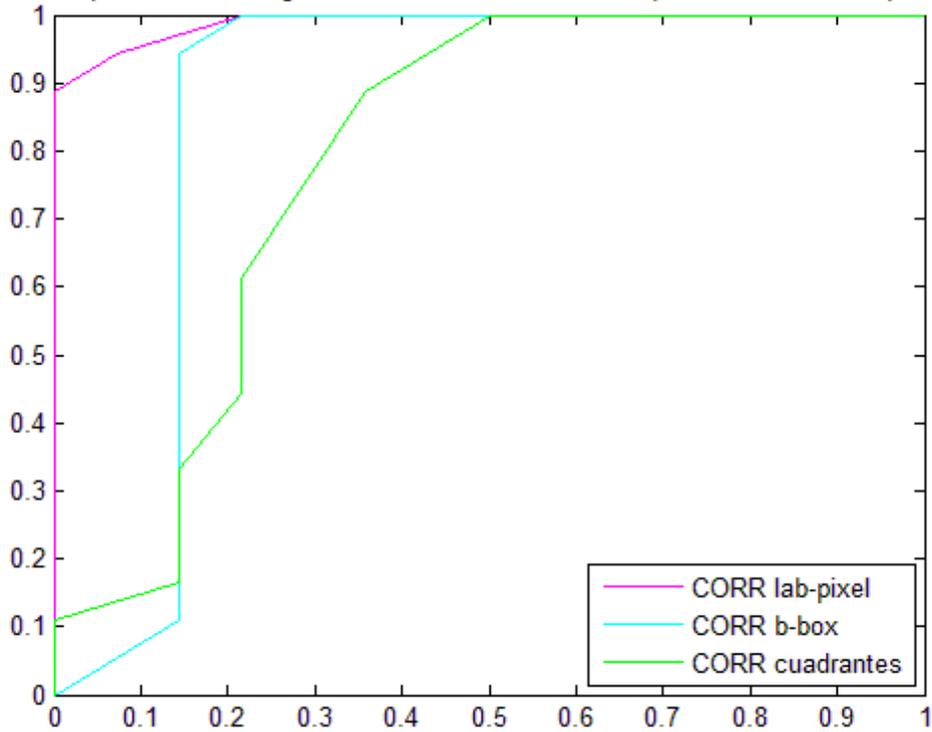


Figura 5-6: Caso Indoor, curvas ROC P4



P4 ROC-SqW: alarma en algún frame de la secuencia, comprobando W frames posteriores

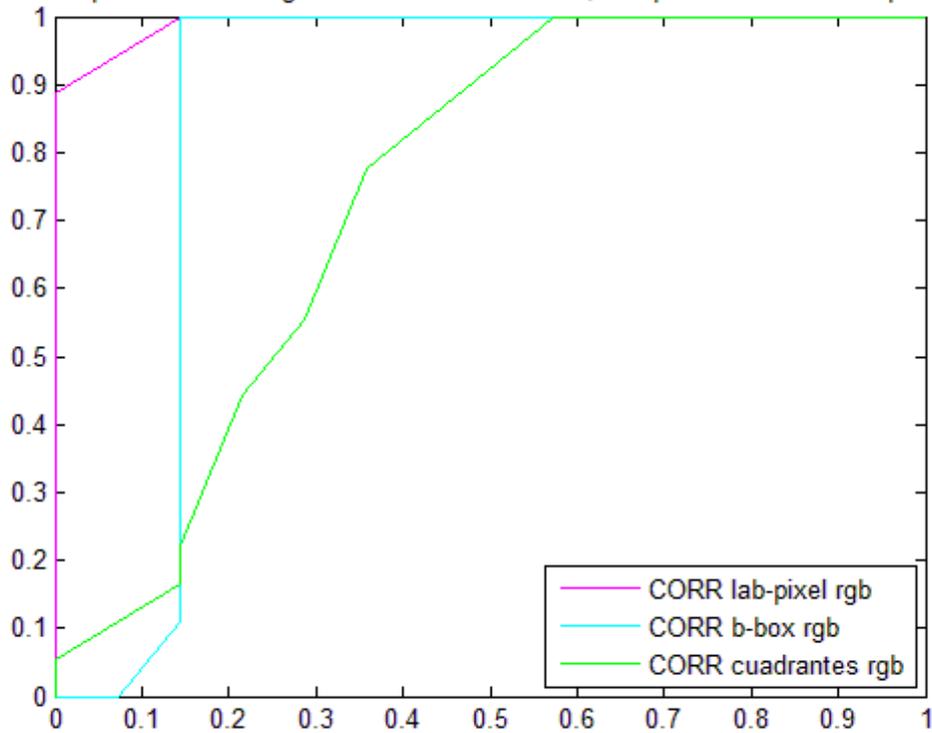


Figura 5-7: Caso Indoor, curvas ROC P4

- **Comentarios**

Como se puede observar en todas las figuras anteriores, independientemente del procedimiento de cálculo, la curva *CORR lab-pixel* correspondiente al algoritmo de cálculo de correlación “En labeling píxeles dentro del bounding box” es la que muestra mejores resultados, seguida de las curvas *CORR b-box* y *CORR cuadrantes*.

Otra de las conclusiones que se extraen a la vista de los gráficos es que las curvas correspondientes a imágenes en niveles de gris obtienen resultados muy similares aunque ligeramente superiores a las curvas RGB.

Si analizamos las gráficas de las **Figura 5-2 y 5-3** , donde recordamos que las tasas FPR y VPR se calculaban como si cada frame de una secuencia fuese independiente de los demás, se obtiene una reducción del 50% de la tasa de falsos positivos para un 90% y un 95% de acierto en casos verdaderos positivos respectivamente.

Sin embargo ningún evento capturado por una cámara de videovigilancia consta de un único frame, se trata de secuencias de imágenes de varios segundos de duración. Incluso en el peor de los casos, donde un intruso intentase cruzar el área videovigilada a gran velocidad y la tasa de imágenes por segundo que se capturan fuera muy reducida se dispone de más de un frame donde aparece el intruso. Según lo anterior podemos concretar que las gráficas calculadas mediante los procedimientos P3 y P4, donde se tenía en cuenta la secuencia completa, representan los resultados que se obtendrían si se implementan estos algoritmos en un sistema real. A partir de este punto y tras la explicación anterior se comentarán con mayor detalle las curvas calculadas mediante dichos procedimientos.

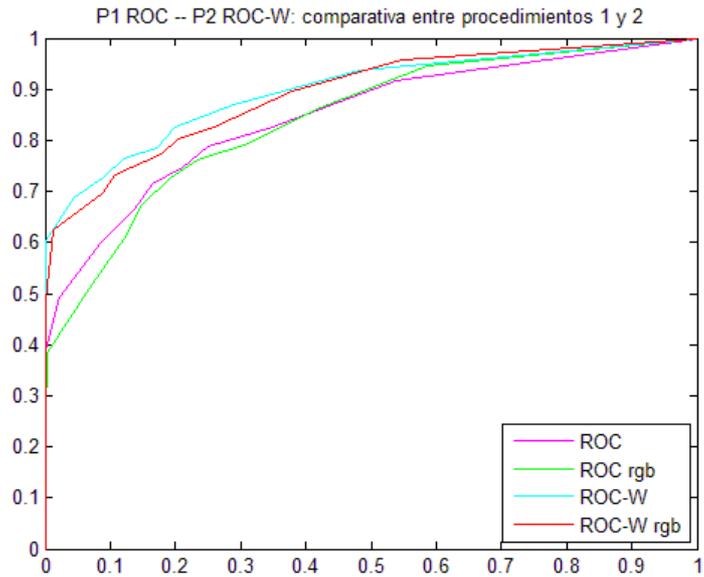
Analizando las gráficas de las **Figura 5-4 y 5-5** , se observa una reducción del 100% de falsos positivos provocados por cambios de iluminación para un 100% de acierto en la generación de alarmas verdaderas; una clasificación perfecta según se vio en el **apartado 5.1.3** . Las curvas *CORR lab-pixel* alcanzan una tasa de reducción del 80-85% para P4 (ver **Figuras 5-6, 5-7**), inferior a la anterior, de lo que se puede concluir que esperar un tiempo para la generación de una alarma resulta menos provechoso que tomar la decisión en cada instante.

5.3.3 Curvas ROC comparativas

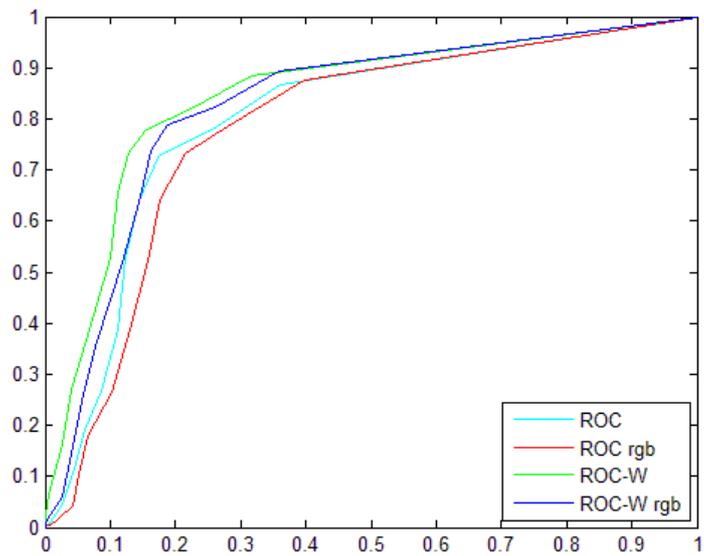
En relación con el comentario anterior se propone mostrar las siguientes gráficas comparativas entre los procedimientos P1-P2 y P3-P4. De esta forma se puede comprobar la diferencia entre curvas y así observar mejor el comportamiento de las dos modalidades del proceso de generación de alarmas.



➔ CORR lab-pixel



➔ CORR b-box



➔ CORR cuadrantes

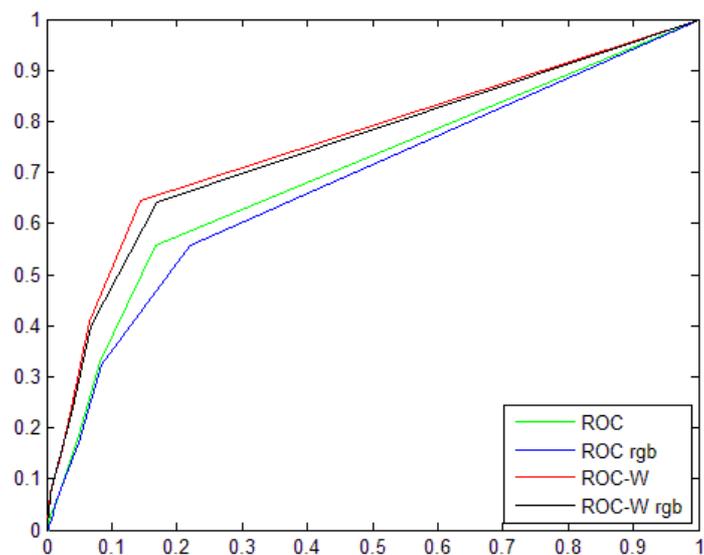


Figura 5-8: Caso Indoor, curvas ROC comparativas P1 y P2

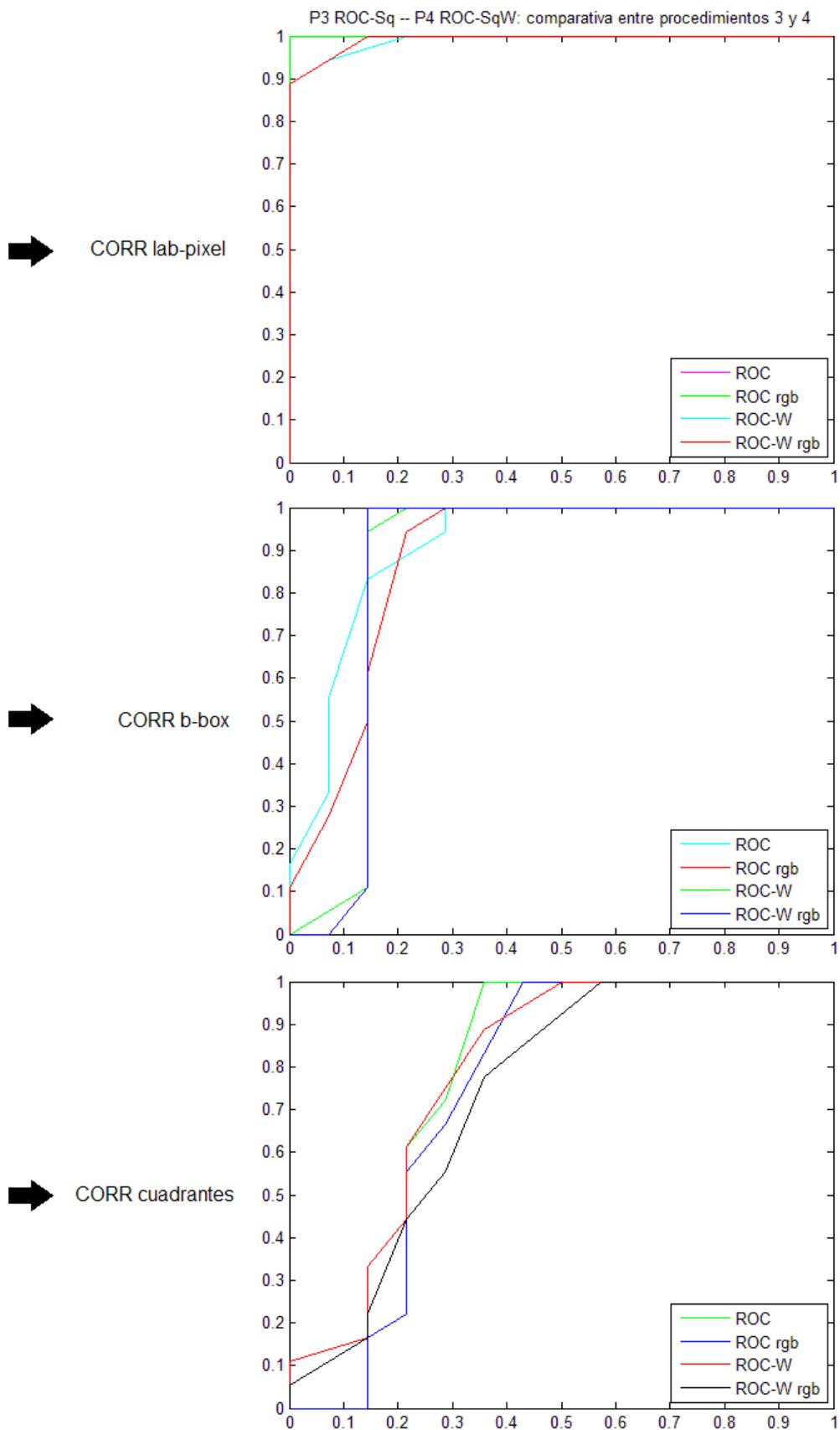


Figura 5-9: Caso Indoor, curvas ROC comparativas P3 y P4
(En el gráfico superior, la curva para ROC y ROC rgb es el cuadrado de área 1)

5.3.4 Tabla de clasificación

La siguiente tabla muestra la clasificación final de las secuencias analizadas para los tres algoritmos evaluados utilizando el procedimiento P3. En función de lo anterior se mostrará el porcentaje de reducción de falsos positivos, donde dicho valor se corresponde con la clasificación obtenida cuando la tasa de verdaderos positivos alcanza el valor máximo, es decir $VPR=1$.

Como se mencionó anteriormente, al final del procesamiento cada una de las secuencias se clasifica como $Alarma=1$ o $Alarma=0$. Se comparará dicha clasificación con la etiqueta inicial para computar las tasas finales de VP y FP.

	# VP	# FP	% Reducción
Inicial	18	14	0
CORR cuadrantes	18	5	64.29
CORR b-box	18	4	71.43
CORR lab-pixel	18	0	100

Tabla 2: Clasificación final de secuencias Indoor

5.4 Resultados caso Outdoor diurno

5.4.1 Secuencias utilizadas para las pruebas

Se muestra en la tabla siguiente la clasificación inicial de las secuencias capturadas en escenarios outdoor en condiciones diurnas. Siguiendo la misma nomenclatura que se utilizó en el apartado anterior para la definición de procedimientos se denotarán con *INTRUSIÓN=1* aquellas secuencias consideradas alarmas verdaderas por la aparición de intrusos y/o vehículos y con *INTRUSIÓN=0* las secuencias donde se ha producido un cambio de iluminación.

# Total Secuencias	<i>INTRUSIÓN=1</i>	<i>INTRUSIÓN=0</i>
18	10	8

Tabla 3: Clasificación inicial de secuencias Outdoor diurnas

5.4.2 Curvas ROC

De igual manera que en el caso indoor se evaluará el funcionamiento de los algoritmos “En cuadrantes”, “En bounding boxes completos” y “En labeling píxeles dentro del bounding box”, descritos en los **apartados 4.3.2, 4.3.3 y 4.3.4** respectivamente. Se representarán de forma conjunta en un mismo gráfico para obtener los resultados de cada algoritmo y poder compararlos entre sí.

Se muestran a continuación las gráficas del caso Outdoor diurno de los procedimientos **P1, P2, P3 y P4**, tal y como se hizo en el **apartado 5.3**. Para los procedimientos 3 y 4 las curvas se desglosarán en dos gráficas diferentes; una para las curvas calculadas con imágenes en niveles de gris y otra para las de RGB para evitar el solape de algunas de ellas.

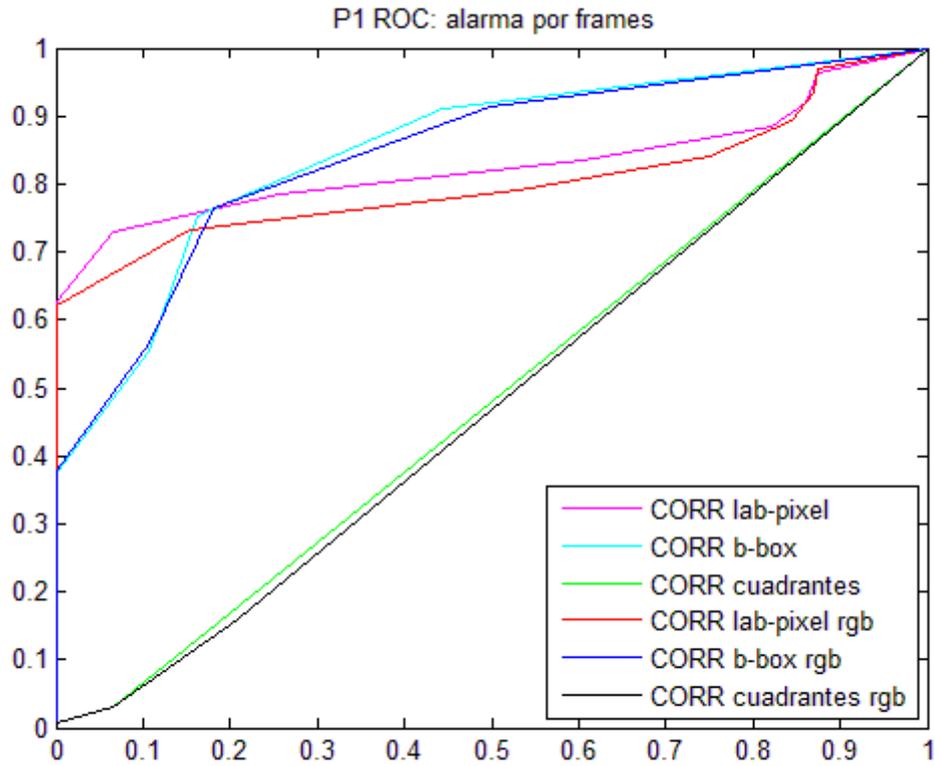


Figura 5-10: Caso Outdoor diurno, curvas ROC P1

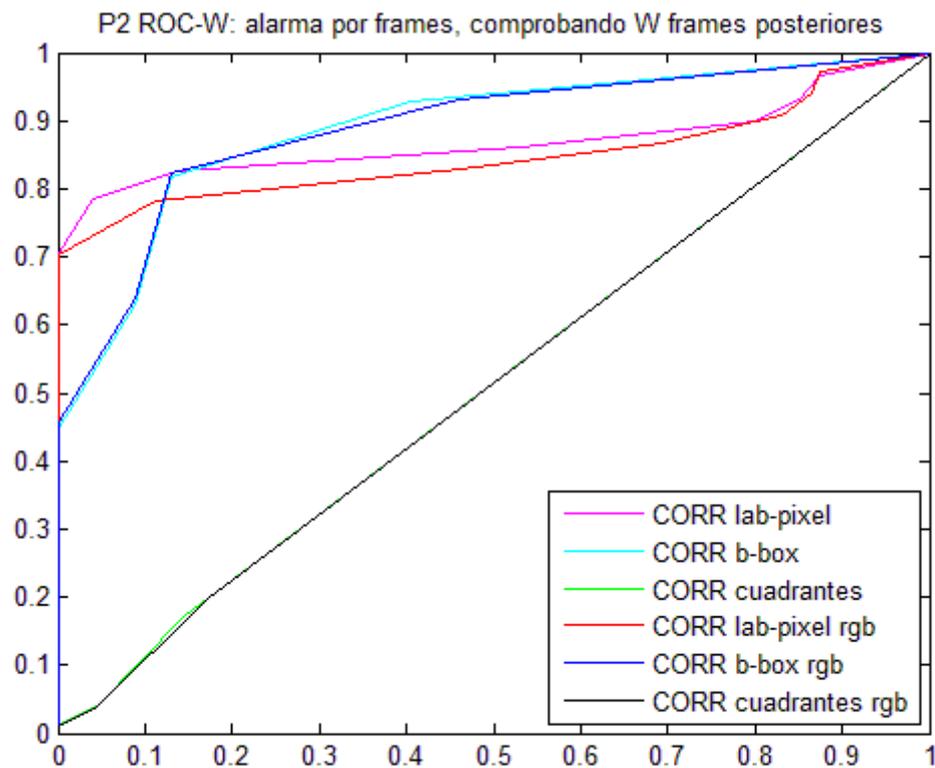


Figura 5-11: Caso Outdoor diurno, curvas ROC P2

(Las curvas para CORR b-box y b-box rgb son prácticamente iguales. Por otra parte, la curva para CORR cuadrantes coincide prácticamente con la de CORR cuadrantes rgb.)

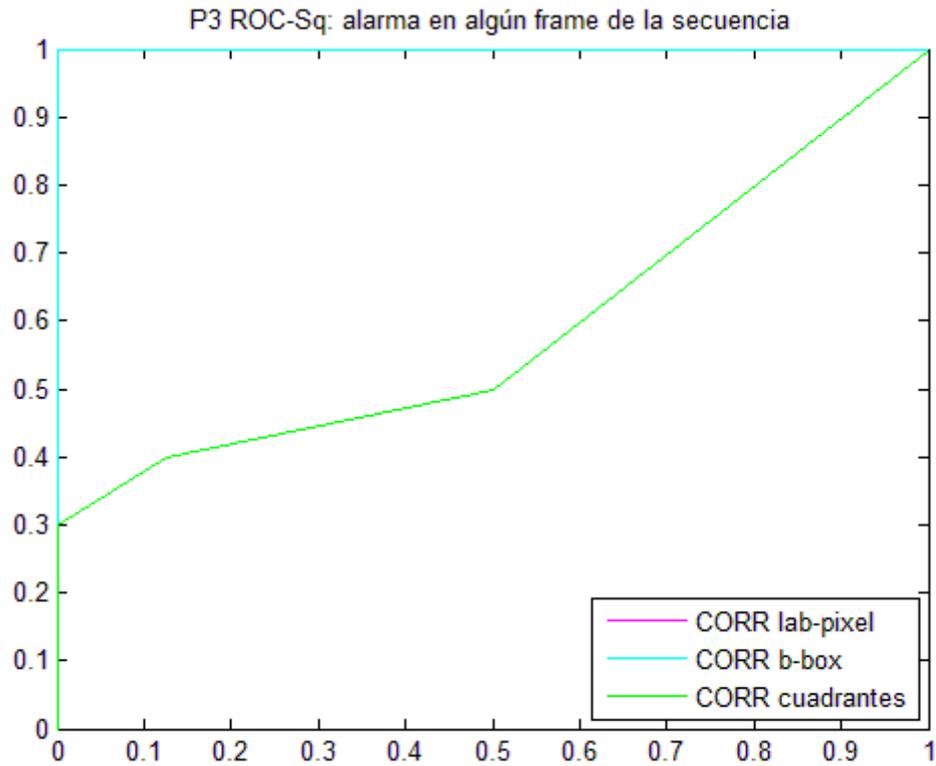


Figura 5-12: Caso Outdoor diurno, curvas ROC P3

(Las ROCs correspondientes a CORR lab-pixel y CORR b-box son el cuadrado de área 1.)

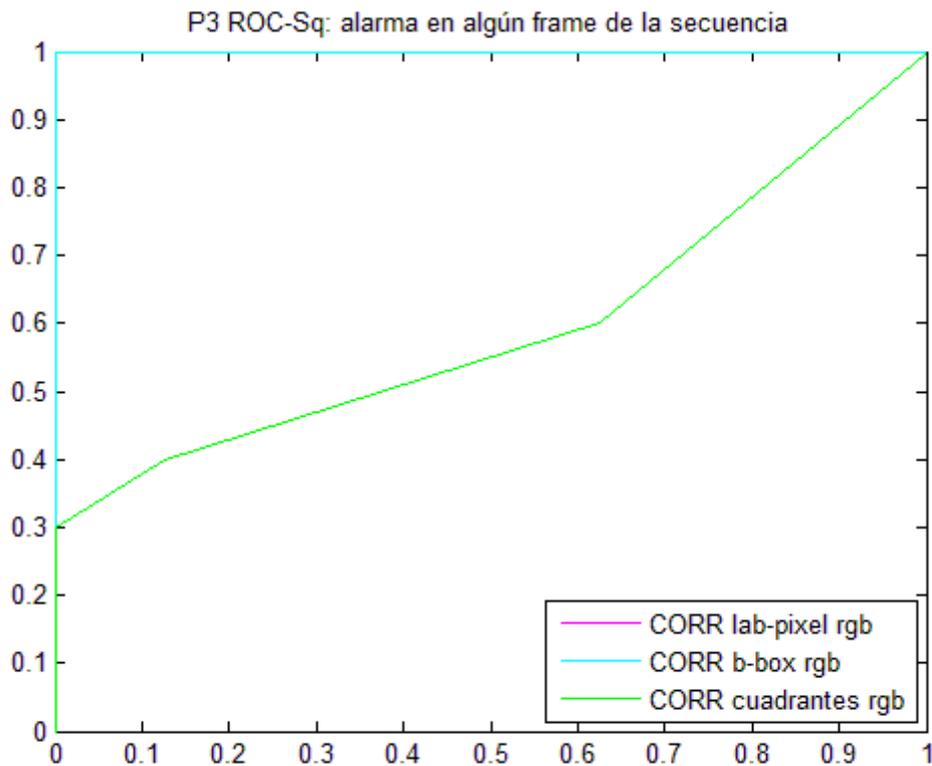


Figura 5-13: Caso Outdoor diurno, curvas ROC P3

(Las ROCs correspondientes a CORR lab-pixel rgb y CORR b-box rgb son el cuadrado de área 1)



P4 ROC-SqW: alarma en algún frame de la secuencia, comprobando W frames posteriores

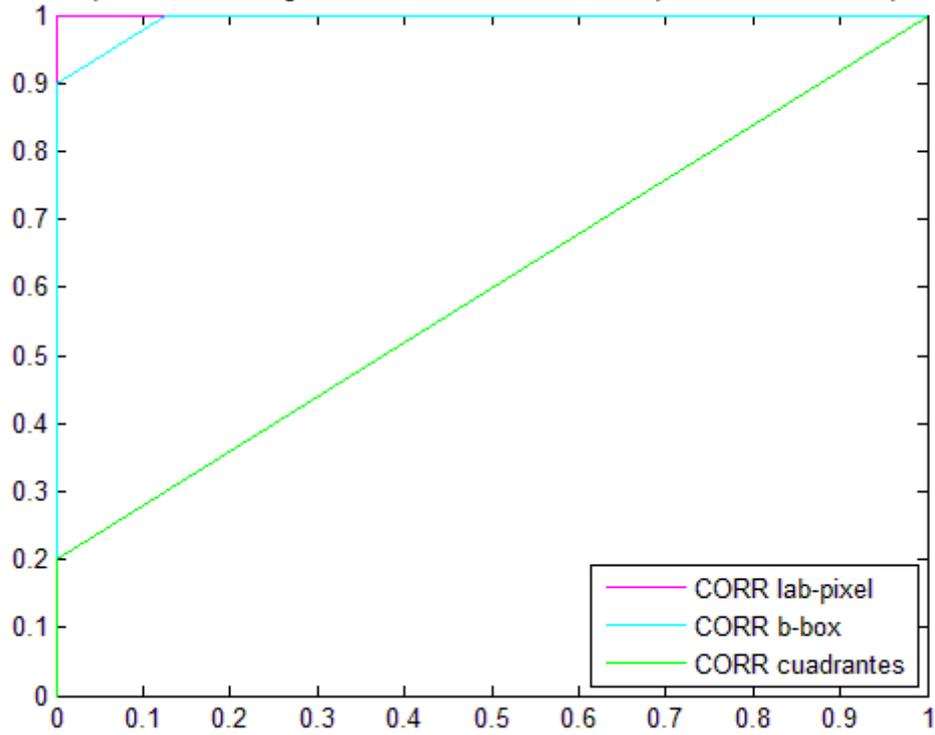


Figura 5-14: Caso Outdoor diurno, curvas ROC P4

(La ROC correspondiente a CORR lab-pixel es el cuadrado de área 1.)



P4 ROC-SqW: alarma en algún frame de la secuencia, comprobando W frames posteriores

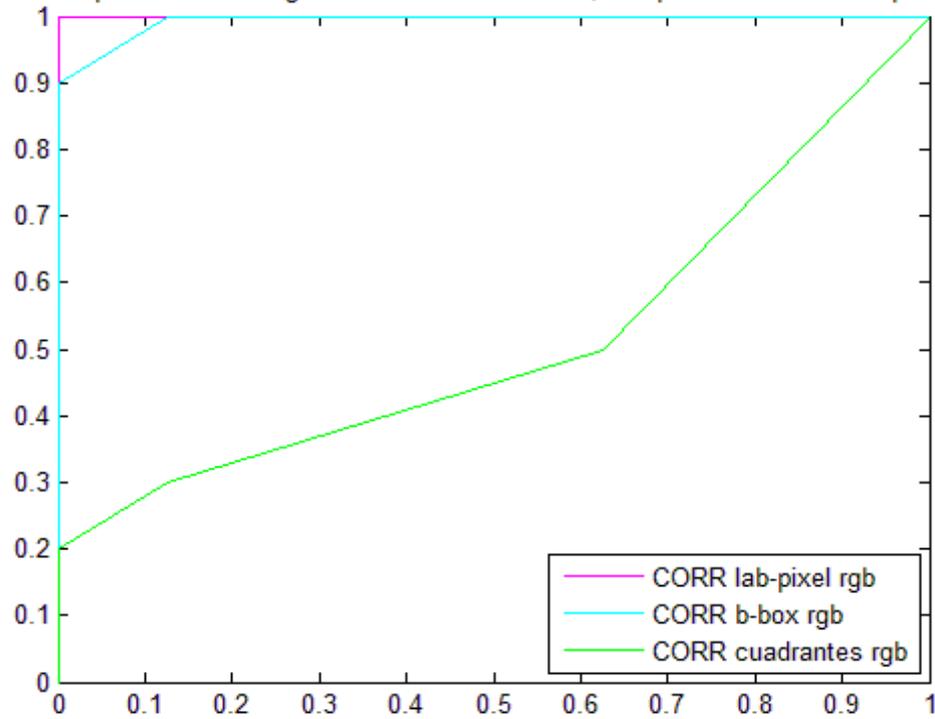


Figura 5-15: Caso Outdoor diurno, curvas ROC P4

(La ROC correspondiente a CORR lab-pixel rgb es el cuadrado de área 1.)

- **Comentarios**

Si analizamos las gráficas de las **Figura 5-10 y 5-11** , donde recordamos que las tasas FPR y VPR se calculaban como si cada frame de una secuencia fuese independiente de los demás, el área bajo las curvas CORR b-box es ligeramente superior al área bajo las curvas *CORR lab-pixel* a diferencia de lo que ocurría en el caso Indoor. Se obtiene una reducción del 50% de la tasa de falsos positivos para un 90% y un 80% de acierto en casos verdaderos positivos respectivamente.

Las curvas *CORR cuadrantes* correspondientes al algoritmo de cálculo de correlación “En Cuadrantes” nos muestran que el comportamiento de este algoritmo es bastante negativo en este escenario ya que $VPR \approx FPR$, lo que no supone ninguna mejora. Se trata de una clasificación aleatoria como se comentó en el **apartado 5.1.3**.

De forma análoga a lo que sucedía en el caso Indoor, las curvas correspondientes a imágenes en niveles de gris obtienen resultados muy similares aunque ligeramente superiores a las curvas RGB.

Analizando las gráficas de las **Figura 5-12 y 5-13**, que como ya se comentó simbolizan los resultados reales, se observa una reducción del 100% de falsos positivos provocados por cambios de iluminación para un 100% de acierto en la generación de alarmas verdaderas para las curvas *CORR lab-pixel* y para las curvas CORR b-box. Sin embargo para P4 (ver **Figuras 5-14, 5-15**) las curvas *CORR lab-pixel* son la únicas que alcanzan ese mismo porcentaje de reducción, siendo para el algoritmo “En bounding boxes completos” del 88% en lugar del 100%. De nuevo se ha obtenido una clasificación perfecta para el algoritmo “En labeling píxeles dentro del bounding box”.

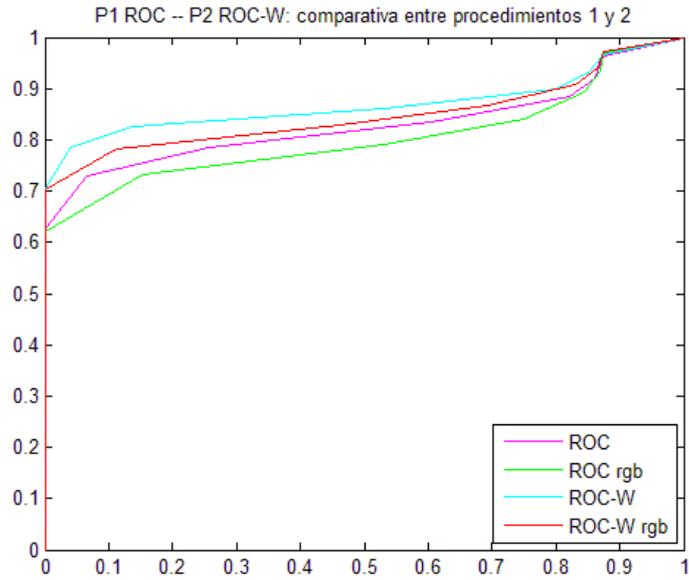
Las curvas *CORR cuadrantes* calculadas mediante P3 y P4 , al igual que sucedía en las primeras figuras, corroboran que el comportamiento de este método no es nada satisfactorio.

5.4.3 Curvas ROC comparativas

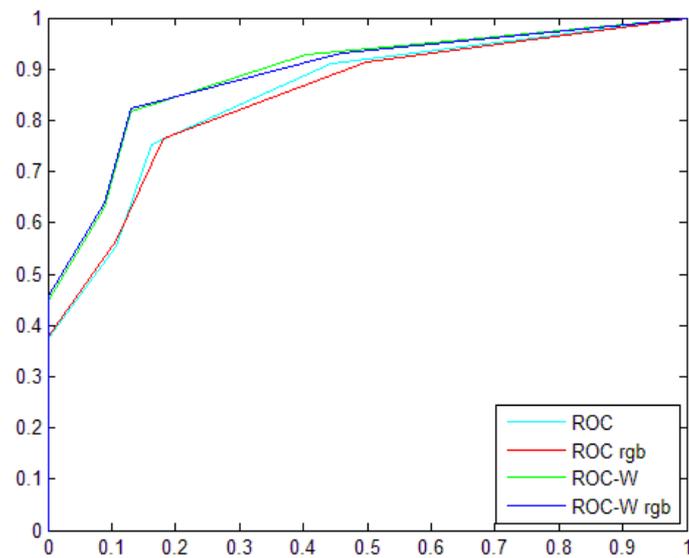
A continuación se muestran las gráficas comparativas entre los procedimientos P1-P2 y P3-P4. De esta forma se puede comprobar la diferencia entre curvas y así observar mejor el comportamiento de las dos modalidades del proceso de generación de alarmas, para concluir si interesa incluir un tiempo de guarda y analizar varios frames consecutivos en este escenario.



CORR lab-pixel



CORR b-box



CORR cuadrantes

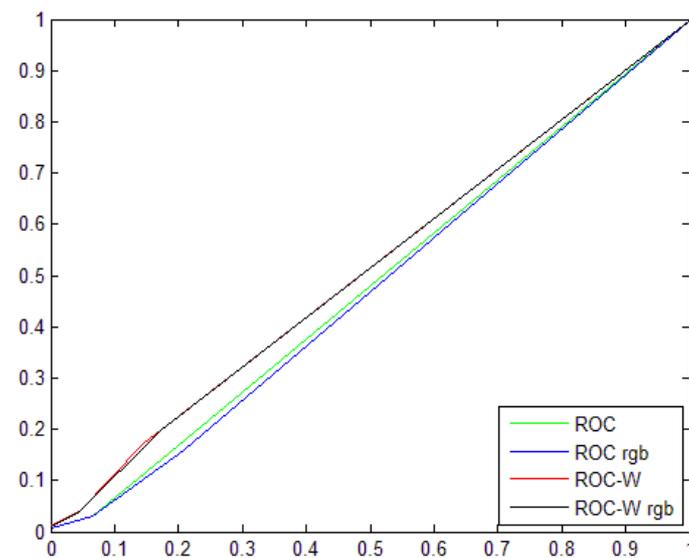
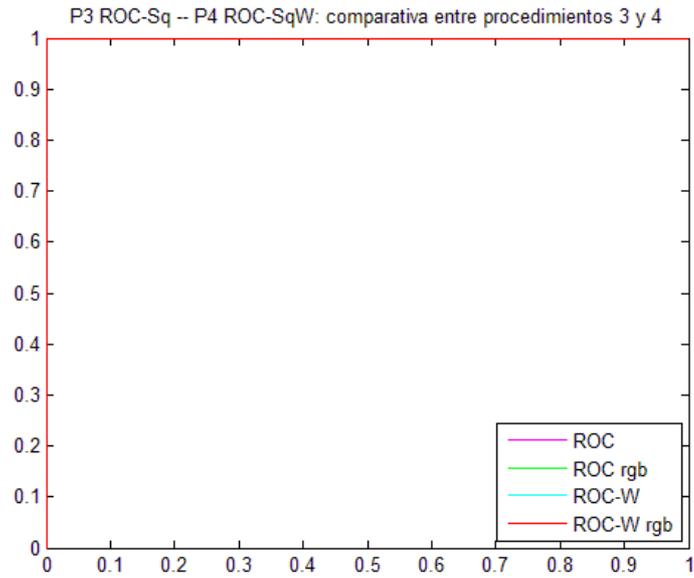


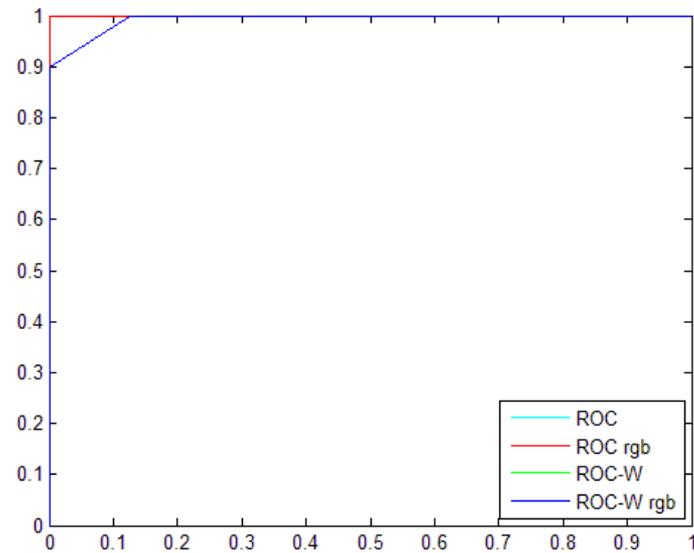
Figura 5-16: Caso Outdoor diurno, curvas ROC comparativas P1 y P2



➔ CORR lab-pixel



➔ CORR b-box



➔ CORR cuadrantes

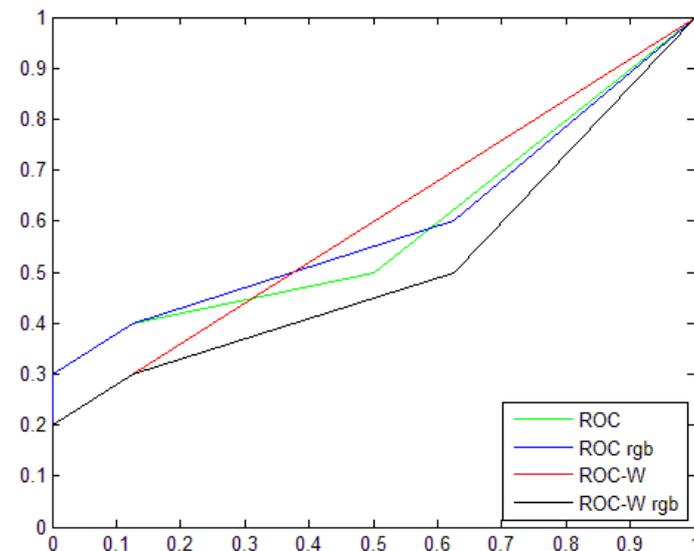


Figura 5-17: Caso Indoor, curvas ROC comparativas P3 y P4

(Gráfico superior: todas las curvas corresponden al cuadrado de área 1. Gráfico intermedio: las curvas correspondientes a ROC y ROC rgb son el cuadrado de área 1.)

5.4.4 Tabla de clasificación

La siguiente tabla muestra la clasificación final de las secuencias analizadas para los tres algoritmos que se han evaluado. En este caso se procesa la secuencia de forma global, tal que al final se clasifica como *Alarma=1* o *Alarma=0*. Las tasas finales de VP y FP se calculan tras comparar dicha clasificación con la etiqueta inicial.

Se mostrará el porcentaje de reducción de falsos positivos correspondiente a la clasificación obtenida cuando la tasa de verdaderos positivos alcanza el valor máximo, es decir $VPR=1$. Para esta clasificación se utiliza el procedimiento P3.

	# VP	# FP	% Reducción
Inicial	10	8	0
CORR cuadrantes	10	8	0
CORR b-box	10	0	100
CORR lab-pixel	10	0	100

Tabla 4: Clasificación final de secuencias Outdoor diurnas

5.5 Resultados caso Outdoor nocturno

5.5.1 Secuencias utilizadas para las pruebas

Se muestra en la tabla siguiente la clasificación inicial de las secuencias capturadas en escenarios outdoor en condiciones nocturnas. Se denotarán con *INTRUSIÓN=1* aquellas secuencias consideradas alarmas verdaderas por la aparición de intrusos y/o vehículos y con *INTRUSIÓN=0* las secuencias donde se ha producido un cambio de iluminación.

# Total Secuencias	<i>INTRUSIÓN=1</i>	<i>INTRUSIÓN=0</i>
41	20	21

Tabla 5: Clasificación inicial de secuencias Outdoor nocturnas

5.5.2 Curvas ROC de diferentes métodos probados para condiciones nocturnas

Para este caso se evaluará el funcionamiento de los algoritmos “En cuadrantes”, “En bounding boxes completos” y “En labeling píxeles dentro del bounding box”, descritos en los **apartados 4.3.2, 4.3.3 y 4.3.4** respectivamente al igual que en los dos casos anteriores. Previamente se realizarán una serie de correcciones y modificaciones a las imágenes, descritas en el **apartado 4.4**.

Una de las propuestas es el “Estudio de otras regiones de cálculo” (ver **apartado 4.4.1**). Se describe a continuación la nomenclatura de las nuevas curvas incluidas en los gráficos ROC donde se ha ampliado la región de cálculo de la correlación.

- ***CORR seg-pixel***: curva ROC obtenida a partir de los datos extraídos tras calcular la correlación en los píxeles segmentados dentro de cada bounding-box .
- ***CORR seg-pixel-imagen***: curva ROC obtenida a partir de los datos extraídos tras procesar las secuencias y evaluar las dependencias de los elementos detectados con el modelo de fondo en todos los píxeles segmentados de la imagen completa.

Esta notación para las curvas representadas se mantendrá en todas las figuras que se incluyen en este apartado.

Se muestran a continuación las gráficas del caso Outdoor en condiciones nocturnas de los procedimientos **P1,P2,P3 y P4** en una misma figura. Cada una de las figuras se corresponde con las diferentes propuestas del **apartado 4.4** para intentar reducir en mayor medida las tasas de falsos positivos.

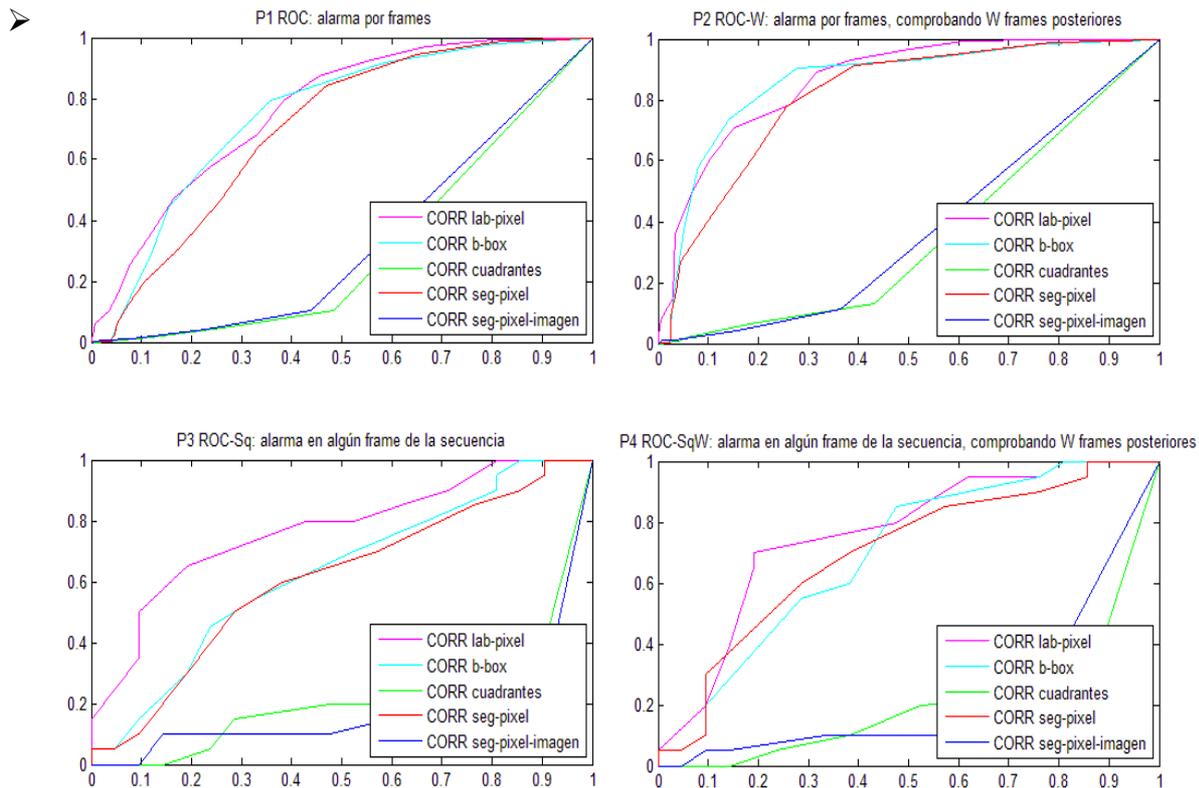


Figura 5-18: Caso Outdoor nocturno. Sin correcciones

- **Comentarios**

Las curvas representadas en la **Figura 5-18** se han obtenido del mismo modo que las de los casos Indoor y Outdoor diurno, sin realizar ningún tipo de modificación a las imágenes de las secuencias utilizadas.

Como se aprecia en la gráfica inferior izquierda, la tasa de reducción de falsos positivos es del 20% frente a la reducción del 100% que teníamos anteriormente. Queda patente a la vista de los resultados la necesidad de introducir modificaciones a los métodos propuestos para intentar reducir de forma más eficiente los falsos positivos debidos a cambios de iluminación (ver **apartado 4.4**).

A pesar de que las tasas no son muy favorables se puede destacar que la curva *CORR lab-pixel* sigue siendo la que muestra mejores resultados. La curva *CORR seg-pixel*, que computaba la correlación en los píxeles marcados en la máscara de segmentación y que pertenecían al área del bounding box del elemento, tiene un comportamiento similar a la anterior aunque peor. Sin embargo la curva *CORR seg-pixel-imagen*, que incluía en el cálculo los píxeles marcados como diferentes del fondo de toda la imagen, muestra una tasa de reducción del 0% y un comportamiento similar al de *CORR cuadrantes* en los cuatro gráficos que componen la figura anterior.

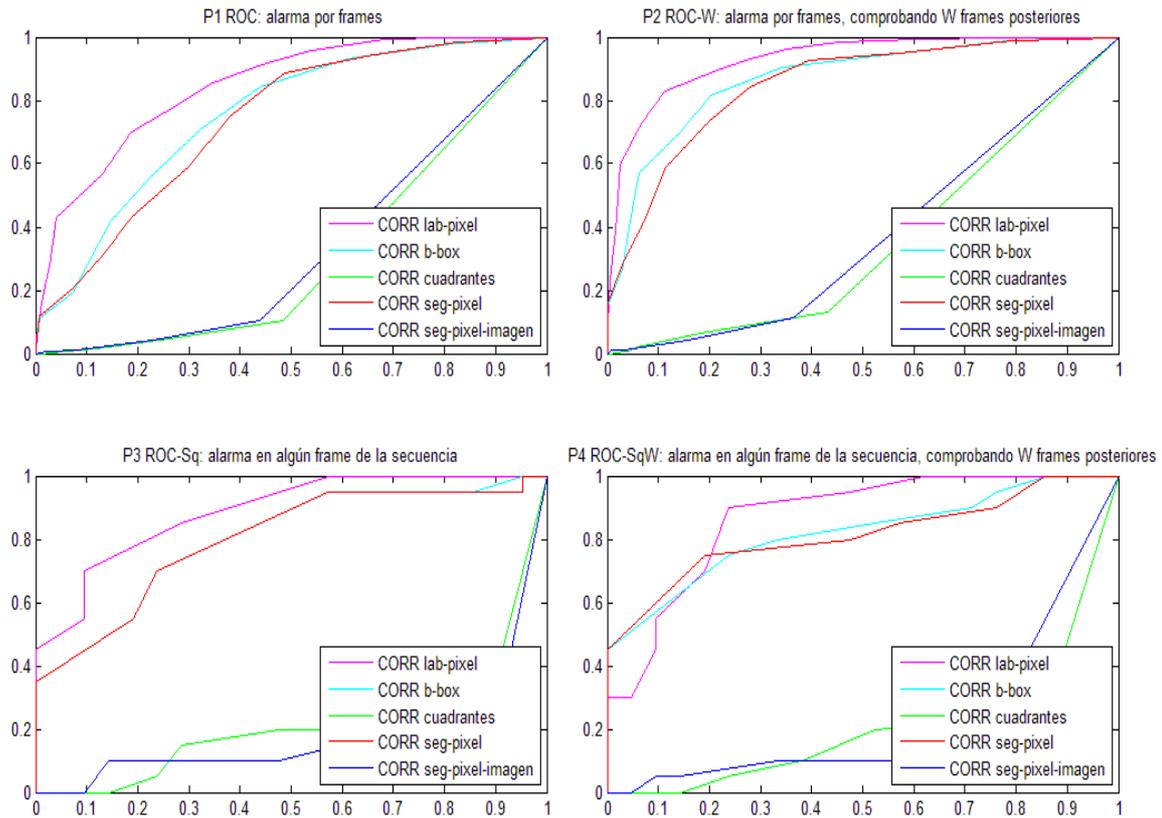


Figura 5-19: Caso Outdoor nocturno. Eliminación del gradiente de iluminación en el bounding box

- **Comentarios**

La **Figura 5-19** muestra los resultados obtenidos tras eliminar el gradiente de iluminación en el área del bounding box de cada elemento detectado a lo largo de la secuencia. Se elimina el gradiente de la imagen original y el de la imagen de fondo según lo descrito en el **apartado 4.4.4** antes de calcular la correlación. No se expondrán los resultados de los **apartados 4.4.2, 4.4.3 y 4.4.5** ya que siendo procedimientos muy similares al de “Eliminar gradiente de iluminación en el bounding box (versión1)”, han arrojado peores resultados.

Las curvas CORR lab-pixel de las cuatro gráficas (P1,P2,P3 y P4) de la figura anterior demuestran que la corrección de la iluminación ha supuesto una mejora considerable. En este caso se obtiene una reducción de la tasa de falsos positivos del 43% y del 40%, para los procedimientos P3 y P4 respectivamente, cuando la tasa de verdaderos positivos es el 100%.

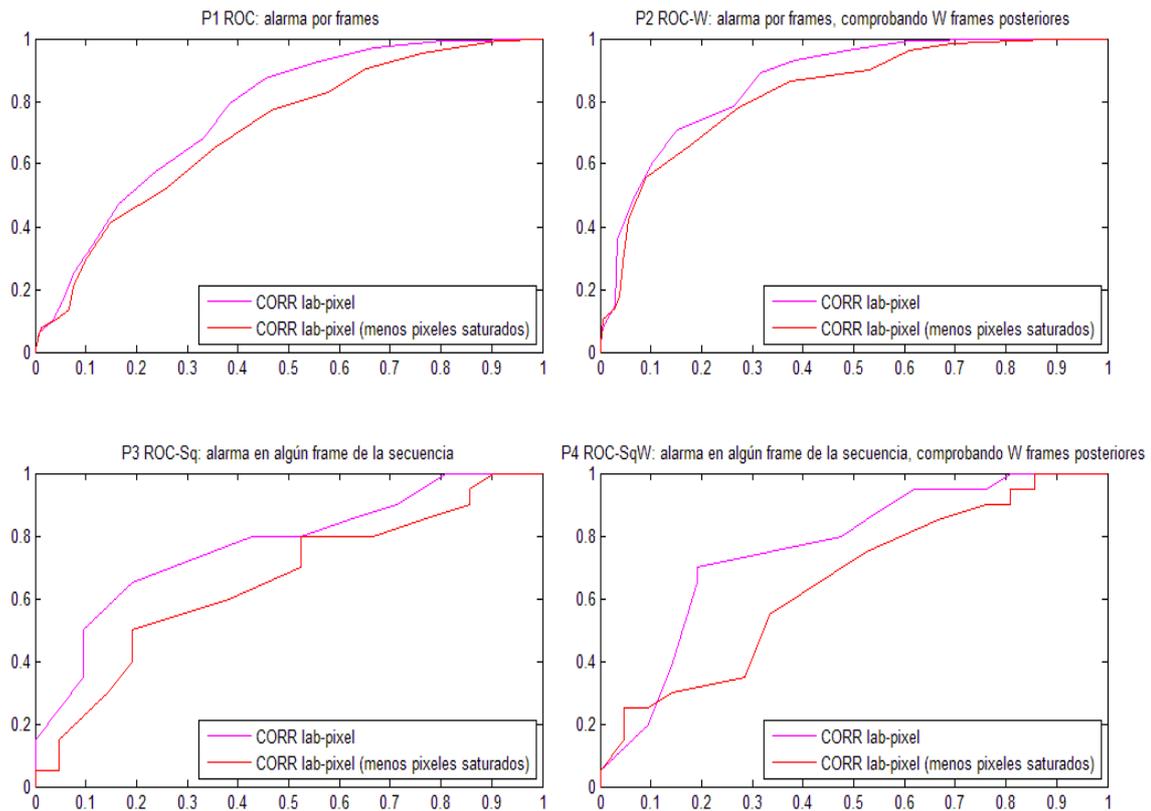


Figura 5-20: Caso Outdoor nocturno. Eliminación de píxeles blancos

- **Comentarios**

En el apartado 4.4.6 se proponía eliminar del cómputo de la correlación aquellos píxeles cuya intensidad fuese superior a un umbral, es decir aquellos píxeles blancos saturados. Para las curvas anteriores dicho umbral se ha fijado en 225.

Se muestra la curva *CORR lab-pixel* y la misma curva tras eliminar los píxeles saturados del blob a modo comparativo. A la vista de los resultados se puede afirmar que esta medida correctiva empeora los resultados frente a no realizar ningún tipo de modificación a las imágenes de las secuencias.

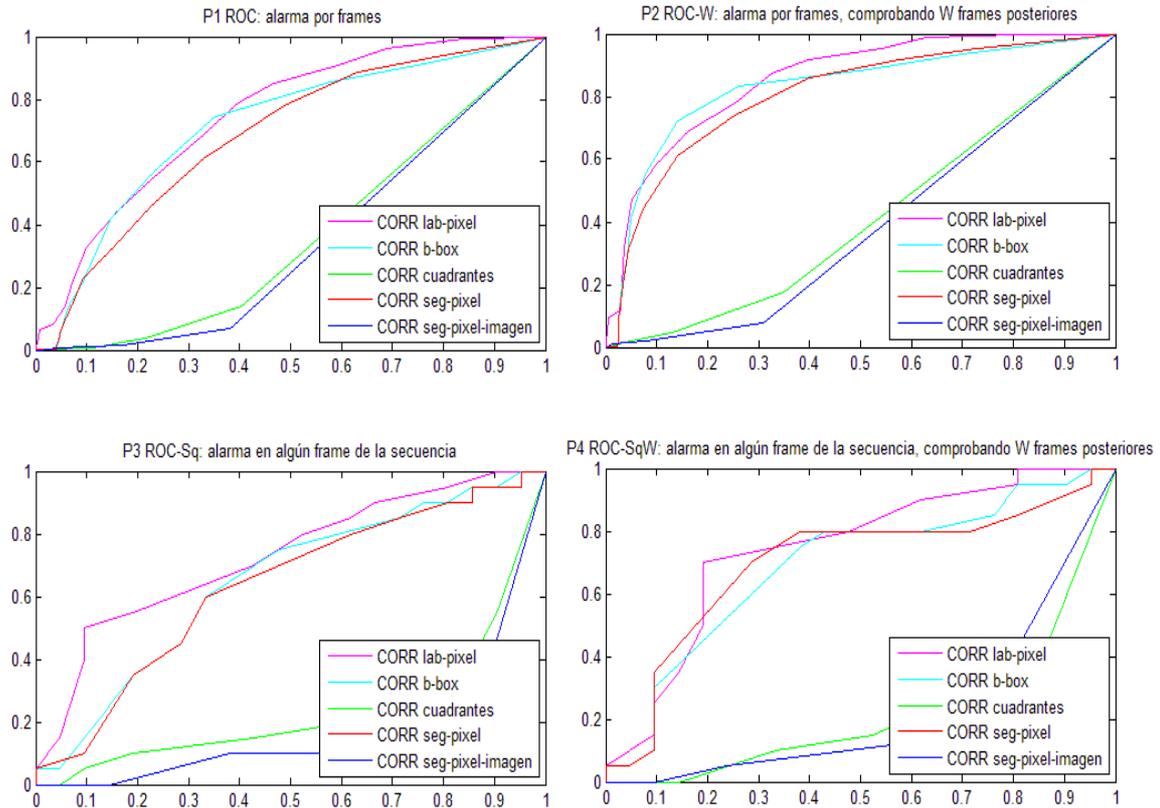


Figura 5-21: Caso Outdoor nocturno. Cambio de escala logarítmico

- **Comentarios**

La **Figura 5-21** muestra los resultados obtenidos tras aplicar un cambio de escala de tipo $\log(x+1)$ a la imagen original y al modelo de fondo previamente al cálculo de dependencias entre dichas imágenes (ver **apartado 4.4.7**).

Fijándose en la gráfica inferior izquierda (P3), la tasa de reducción de falsos positivos es del 10% frente al 20% inicial cuando no se aplicaba ninguna corrección a las imágenes; no supone una mejora sino un empeoramiento.

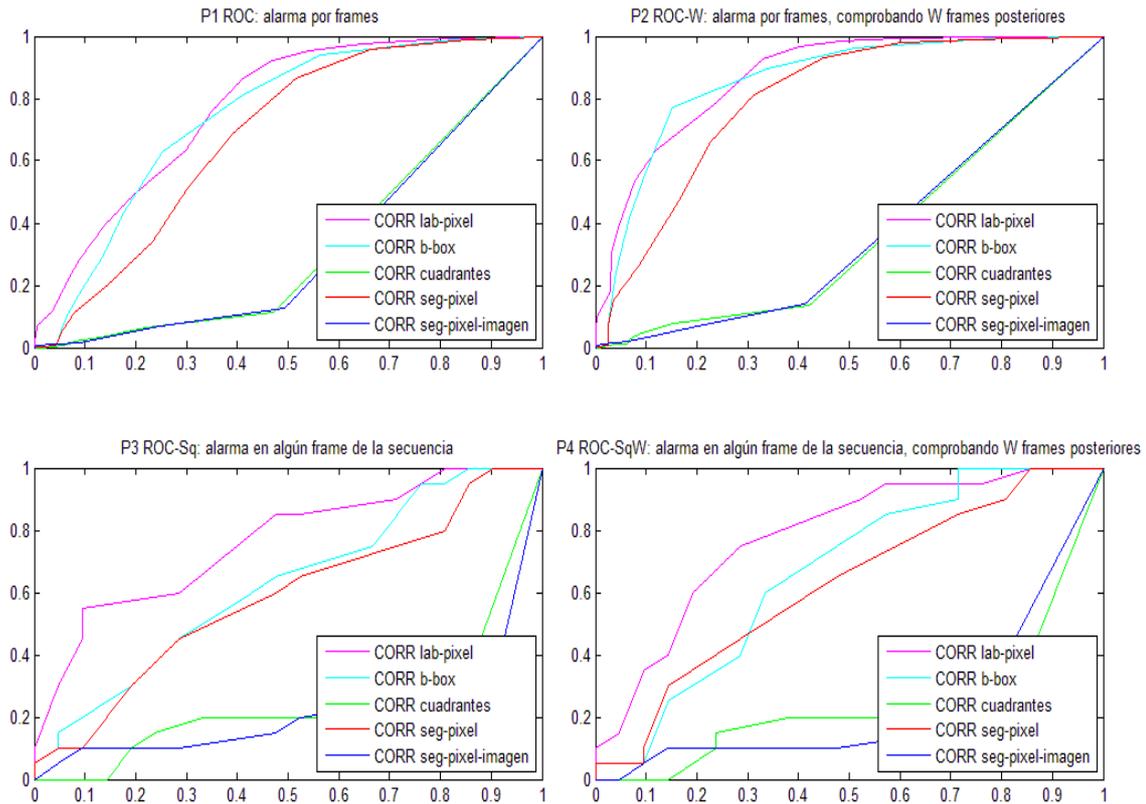


Figura 5-22: Caso Outdoor nocturno. Cambio de escala exponencial

- **Comentarios**

La **Figura 5-22** muestra los resultados obtenidos tras aplicar un cambio de escala de tipo x^a a la imagen original y al modelo de fondo previamente al cálculo de dependencias entre dichas imágenes. Estos resultados se han obtenido para $a=3$ (escala cúbica).

A diferencia de lo mostrado en la **Figura 5-21**, la tasa de reducción de falsos positivos es del 20%, como en el método inicial, al observar la gráfica correspondiente a P3. Por lo tanto no se produce ni mejora ni empeoramiento. En cambio, al observar la gráfica inferior derecha (P4) sí se obtienen mejores resultados que cuando no se trataban las imágenes de partida.

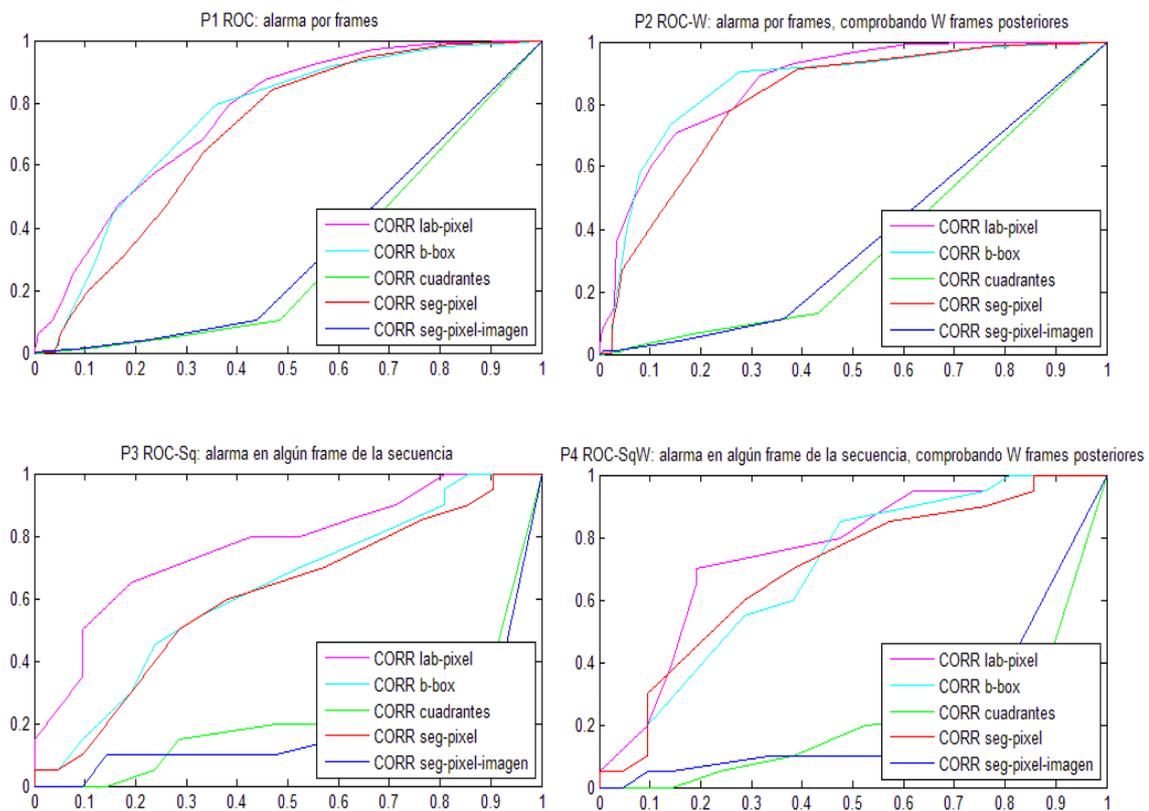


Figura 5-23: Caso Outdoor nocturno. Cálculo de la intensidad media de los píxeles vecinos

- **Comentarios**

Los resultados representados en la **Figura 5-23** se corresponden con la corrección descrita en el **apartado 4.4.8**, donde se calculaba el valor medio de intensidad para cada píxel en función de la intensidad de los píxeles vecinos. Se restará dicho valor medio tanto a la imagen original como a la imagen de fondo antes de evaluar la correlación entre ambas. En este caso se ha utilizado una máscara de 7x7 píxeles.

Las tasas de reducción de falsos positivos son muy similares a las obtenidas sin realizar modificaciones y a otras de las representadas en este apartado. Sin embargo esta medida tiene una gran objeción que no tenían las propuestas anteriores, el tiempo. El coste computacional es elevado y el tiempo de procesamiento de este método no permitiría aplicaciones en tiempo real. Como se ha mencionado en diferentes puntos de este proyecto la solución propuesta se debe ajustar a las necesidades básicas de un sistema de videovigilancia inteligente, por lo tanto es necesario procesar como mínimo 4-5 frames por segundo para que se efectivo.

5.5.3 Curvas ROC del método con resultados más favorables

Se muestran en este apartado las gráficas del procedimiento con el que se han obtenido mejores resultados para el caso Outdoor nocturno. Después de realizar las pruebas expuestas en el apartado 5.5.2 se ha comprobado la eficiencia de combinar los diferentes métodos.

La **Figura 5-19** demostraba que eliminar el gradiente de iluminación de forma localizada en el área donde se encontraban los elementos detectados suponía una mejora sustancial. Debido a lo anterior esta medida de corrección se ha probado con todas las otras propuestas. Otro de los métodos que arrojaba resultados ligeramente superiores a la situación de partida era aplicar un cambio de escala de tipo exponencial.

De lo anterior se puede deducir que la combinación de un cambio de escala seguido de la eliminación del gradiente de iluminación consigue mejorar el comportamiento de los algoritmos de medida de dependencias para imágenes nocturnas o capturadas en condiciones de baja iluminación.

Se muestran a continuación las curvas ROC calculadas mediante los procedimientos P1, P2, P3 y P4, tal y como se hizo en los casos Indoor y Outdoor diurno.

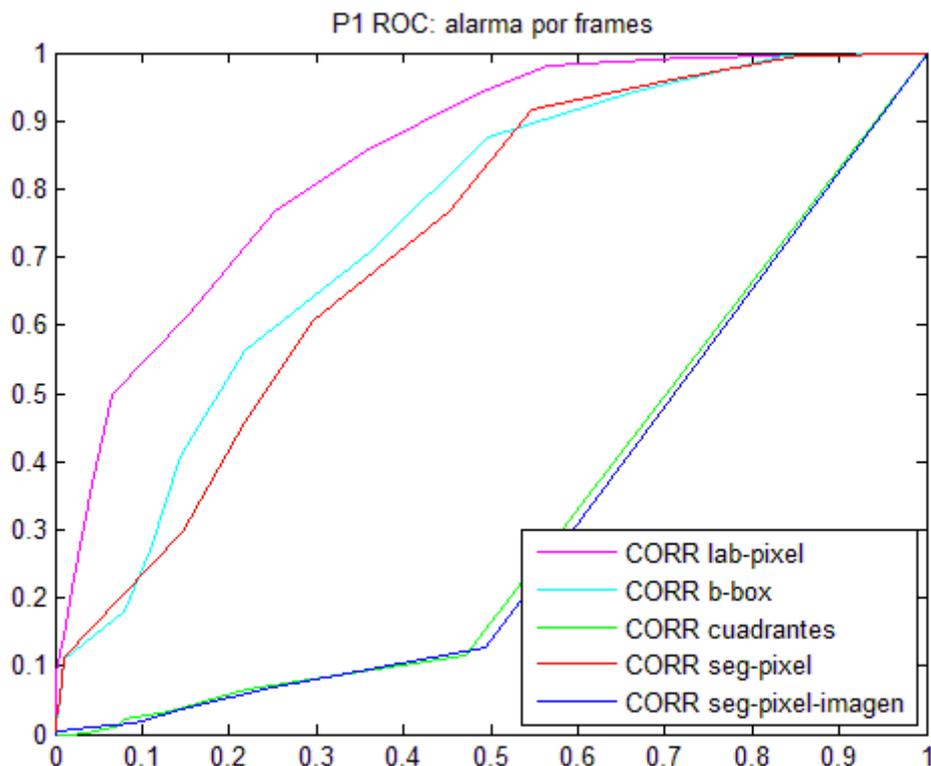


Figura 5-24: Caso Outdoor nocturno, curvas ROC P1

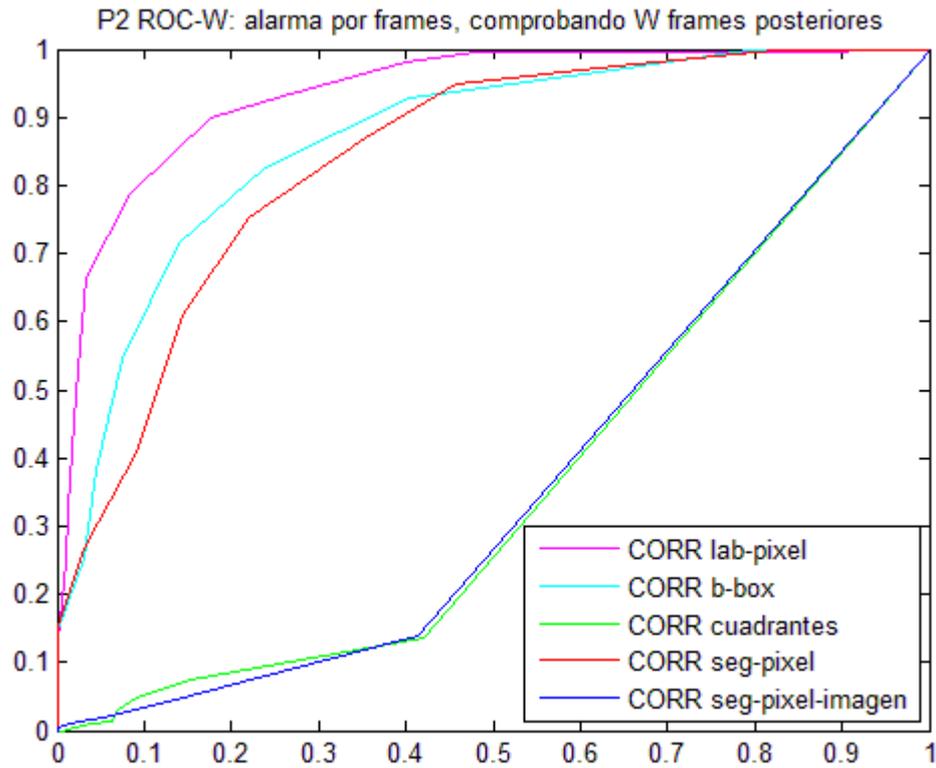


Figura 5-25: Caso Outdoor nocturno, curvas ROC P2

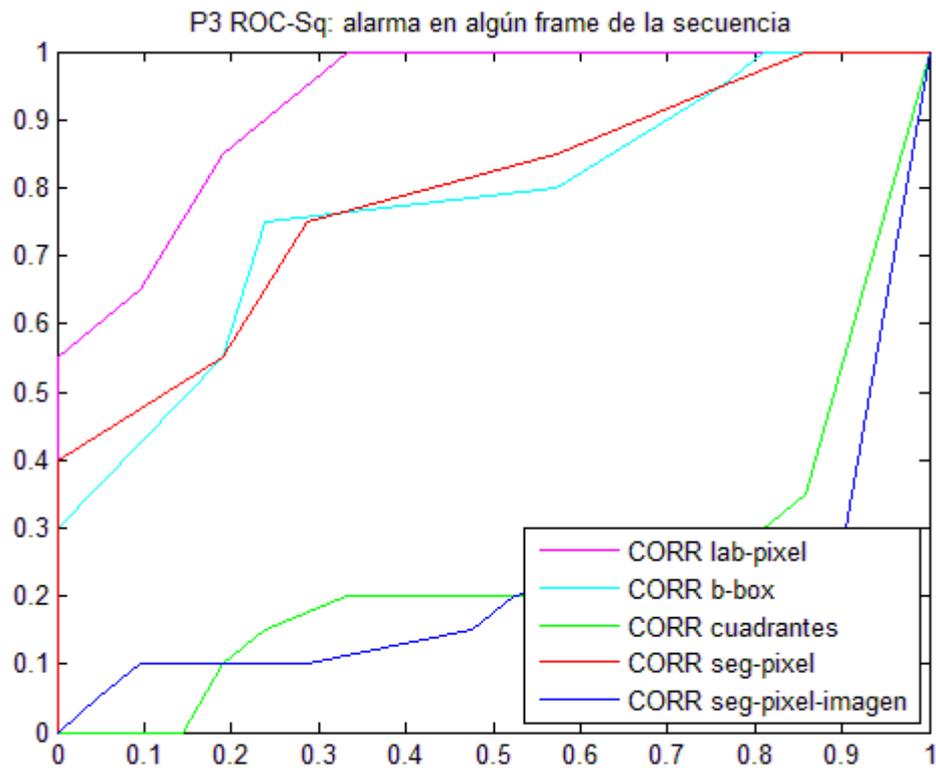


Figura 5-26: Caso Outdoor nocturno, curvas ROC P3

P4 ROC-SqW: alarma en algún frame de la secuencia, comprobando W frames posteriores

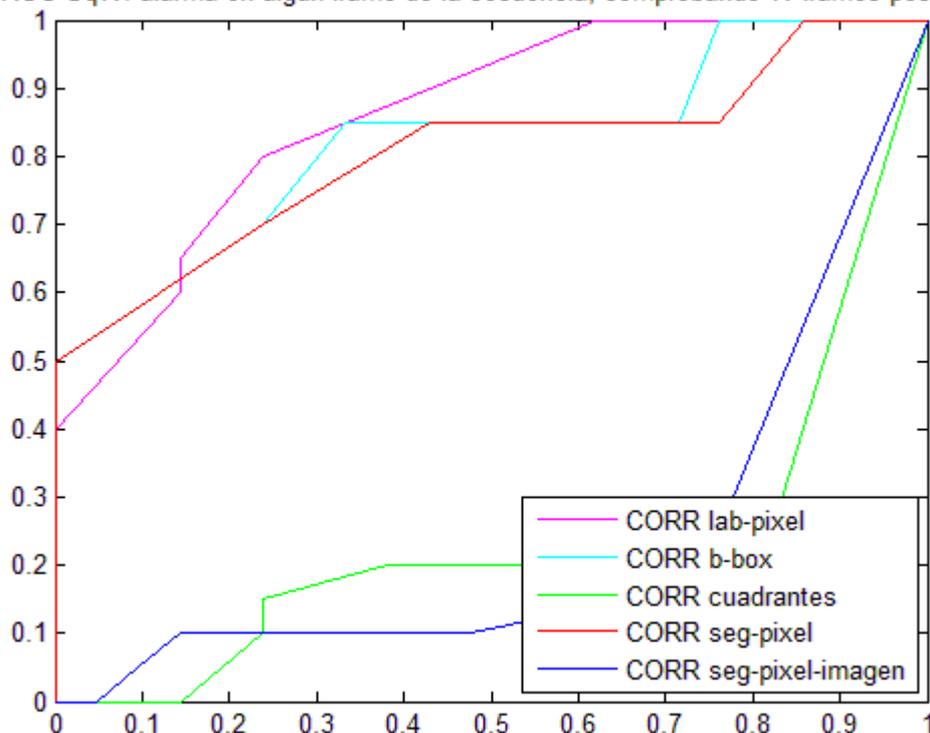


Figura 5-27: Caso Outdoor nocturno, curvas ROC P4

- **Comentarios**

Las gráficas de las **Figura 5-24 y 5-25** , donde recordamos que las tasas FPR y VPR se calculaban como si cada frame de una secuencia fuese independiente de los demás, el área bajo las curvas *CORR lab-pixel* b-box es muy superior al área bajo las demás curvas siguiendo la tendencia de los dos casos previos. Se obtiene una reducción del 50% de la tasa de falsos positivos para un 95% y un 99% de acierto en casos verdaderos positivos respectivamente.

Las curvas *CORR cuadrantes* y *CORR seg-pixel-imagen* correspondientes a los métodos de cálculo de correlación en regiones más extensas muestran un comportamiento muy negativo. Según lo anterior se puede concluir que analizar el área donde se encuentran los elementos detectados resulta más eficiente que trabajar con toda la imagen, por tanto los resultados obtenidos han corroborado las estrategias propuestas inicialmente.

Analizando la gráfica de la **Figura 5-26** se aprecia que esta combinación de métodos permite una reducción de casi un 70% de falsos positivos provocados por cambios de iluminación para un 100% de acierto en la generación de alarmas verdaderas para la curva *CORR lab-pixel* calculada mediante el procedimiento Para P4 (ver **Figura 5-27**) la curva *CORR lab-pixel* muestra una reducción de casi un 40% de los falsos positivos.

5.5.4 Tabla de clasificación

La siguiente tabla muestra la clasificación final de las secuencias analizadas para los cinco métodos evaluados en este caso utilizando el P3. En función de lo anterior se mostrará el porcentaje de reducción de falsos positivos, donde dicho valor se corresponde con la clasificación obtenida cuando la tasa de verdaderos positivos alcanza el valor máximo, es decir $VPR=1$.

Al final del procesamiento cada una de las secuencias se clasifica como $Alarma=1$ o $Alarma=0$. Se comparará dicha clasificación con la etiqueta inicial para computar las tasas finales de VP y FP.

	# VP	# FP	% Reducción
Inicial	20	21	0
CORR cuadrantes	20	21	0
CORR b-box	20	17	19.05
CORR lab-pixel	20	7	66.67
CORR seg-pixel	20	18	14.29
CORR seg-pixel-imagen	20	21	0

Tabla 6: Clasificación final de secuencias Outdoor nocturnas

5.6 Validación de Resultados

5.6.1 Métodos de validación cruzada

Para la validación de resultados se hará uso de técnicas de validación cruzada o *cross-validation*. Se trata de una práctica estadística que consiste en partir una muestra de datos en subconjuntos mutuamente excluyentes de tal modo que el análisis es inicialmente realizado en uno de ellos, mientras los otros subconjuntos son retenidos para su uso posterior en la confirmación y validación del análisis inicial. Esta técnica se utiliza comúnmente en Inteligencia Artificial para validar los modelos generados a partir de un conjunto de datos o muestra y fue introducida por M.Stone en 1974 [37].

La validación cruzada aplica k veces la validación simple. Se divide en k conjuntos la muestra (*k-fold cross-validation*). En cada iteración se construirá y evaluará un modelo, usando uno de los conjuntos como *training set* y el resto como *test set* para medir el ratio de error del modelo construido [38]. Al final obteniendo la media aritmética de los ratios de error obtenidos conseguiremos el ratio de error para la muestra final. La elección del valor k dependerá del tamaño y características de la muestra [39].

- **Leave one out**

En este proyecto se utilizará el método *leave-one-out cross-validation (LOOCV)*, que es un caso especial de *k-fold cross-validation*, donde k es igual al número de muestras. Como su nombre sugiere, “dejando uno fuera”, se deja una muestra para la validación y el resto de muestras componen el set de entrenamiento. Se repite el proceso k veces de forma que cada una de las muestras se utilice una vez como test set. Se trata de un método de validación de alto coste computacional ya que se necesita repetir el proceso de entrenamiento un número elevado de veces [40].

5.6.2 Diseño de pruebas para la validación de resultados

Se representa para cada uno de los tres casos (Indoor, Outdoor diurno y Outdoor nocturno) un figura con los valores de correlación de las secuencias y el umbral resultante del proceso de validación cruzada mediante la técnica *leave-one-out*.

En estos gráficos se representan cada uno de los valores mínimos de correlación obtenidos en las secuencias de pruebas del caso Indoor, donde dichos valores se han calculado mediante el algoritmo “En labeling píxeles dentro del bounding box”. Se representa únicamente un valor mínimo por secuencia ya que se considera que si dicho valor tras compararlo con el umbral de correlación provoca la generación de una alarma será suficiente información para este proceso de validación. De igual manera si el valor mínimo de correlación es superior al umbral dicha secuencia no generará alarma en ningún frame.

Se representan mediante circulo verde los valores mínimos correspondientes a secuencias pertenecientes al conjunto de verdaderos positivos, etiquetadas tal que $INTRUSIÓN=1$ siguiendo la misma nomenclatura que en apartados anteriores. Para las secuencias correspondientes a falsos positivos en la clasificación inicial ($INTRUSIÓN=0$) se utilizan cruces rojas. Se representan también mediante línea continua el valor umbral.

Como se ha mencionado en diferentes puntos de este proyecto es necesario clasificar correctamente todas las secuencias correspondientes a alarmas verdaderas. Debido a lo anterior se utilizarán sólo los valores mínimos de correlación de las secuencias $INTRUSIÓN=1$ para la determinación del umbral en cada iteración. Por lo tanto según el procedimiento *leave-one-out* se utilizarán los valores mínimos de correlación como set de entrenamiento dejando en cada iteración uno fuera como valor de test. Es decir, para cada realización el umbral se determina a partir de los valores mínimos de secuencias verdaderas que restan en el *training set*. Por lo tanto cuando el valor de test pertenece al subconjunto de secuencias $INTRUSIÓN=1$ el valor umbral calculado puede ser diferente al de otras realizaciones, mientras que si el valor de test pertenece al subconjunto $INTRUSIÓN=0$ el umbral será el que clasifica correctamente todos los verdaderos positivos.

Para cada realización el valor de test se compara con con el umbral determinado de forma que si dicho valor es superior al umbral se clasificará con Alarma=0 y si es inferior con Alarma=1. A partir de esta nueva clasificación obtendremos las tasas de verdaderos positivos y falsos negativos de forma equivalente a los procedimientos para la obtención de resultados descritos en el apartado 5.2. Se representarán las tasas de reducción de falsos positivos en una tabla comparativa conjuntamente con las tasas correspondientes al algoritmo “En labeling píxeles dentro del bounding box” extraídas del capítulo de resultados. Esta tabla nos servirá para validar los resultados obtenidos y presentados en apartados anteriores.

Se muestra a continuación la **Figura 5-28** donde aparece un diagrama representativo del procedimiento de validación de resultados realizado. En dicho diagrama se mantiene la nomenclatura utilizada a lo largo de este capítulo.

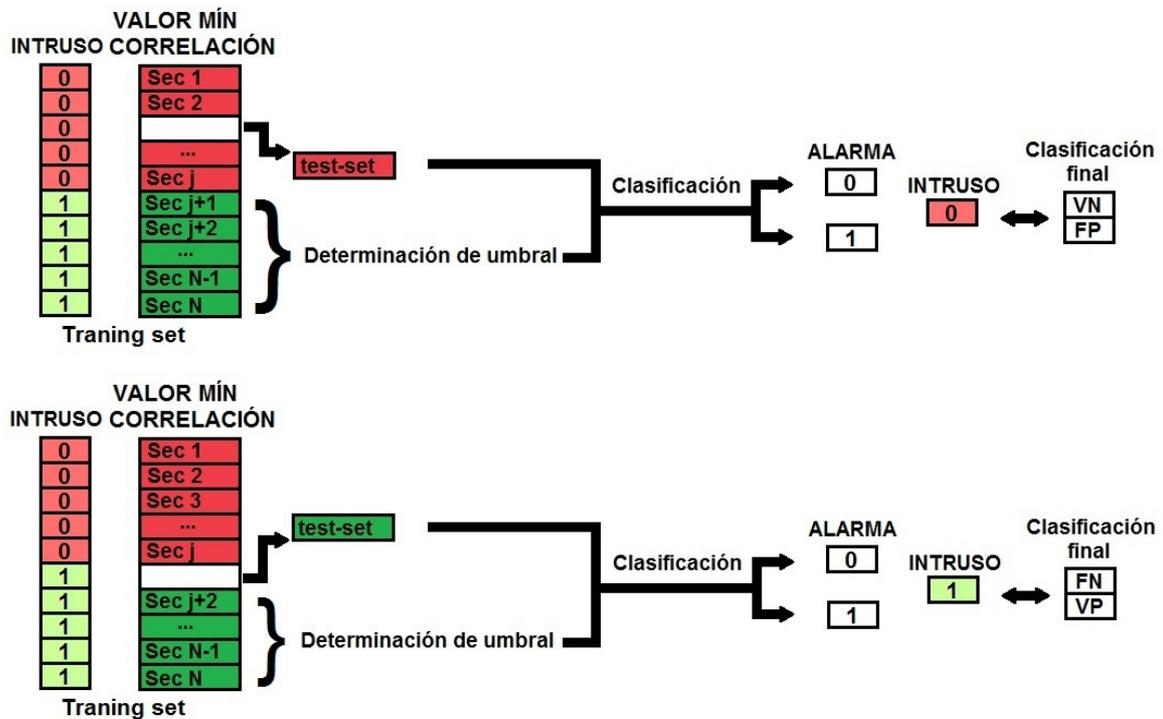


Figura 5-28: Procedimiento para la validación de resultados

5.6.3 Validación de resultados del caso Indoor

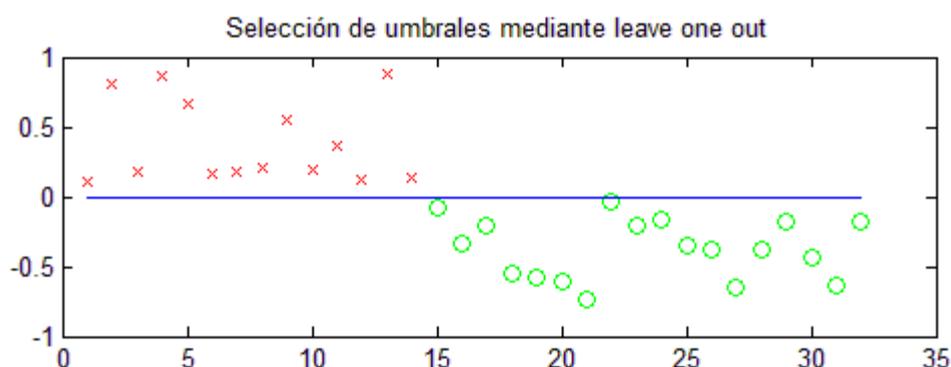


Figura 5-29: Validación de resultados para caso Indoor

	# VP	# FP	% Reducción
CORR lab-pixel	18	0	100
TASA validación	18	0	100

Tabla 7: Validación de las tasas de reducción de falsos positivos en secuencias Indoor

● Comentarios

Como se aprecia en la gráfica superior de la **Figura 5-29** a partir de los valores mínimos de correlación de las secuencias correspondientes con verdaderos positivos se ha determinado un umbral único. También se aprecia que dicho umbral clasifica perfectamente los cambios de iluminación, cuyos valores mínimos están por encima del umbral para todas las secuencias.

La tasa de reducción de falsos positivos obtenida para el umbral calculado mediante *leave-one-out* representa la validación de los resultados que se presentaban en la **Figura 5-4** donde la curva *CORR lab-pixel* representaba una reducción del 100% de falsos positivos para un 100% de detección de verdaderos positivos; quedan por tanto validados los resultados para el caso Indoor.

5.6.4 Validación de resultados del caso Outdoor diurno

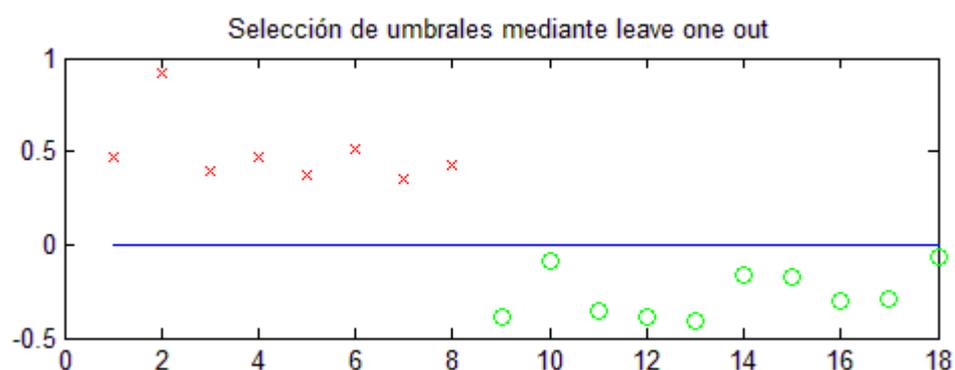


Figura 5-30: Validación de resultados para caso Outdoor diurno

	# VP	# FP	% Reducción
CORR lab-pixel	10	0	100
TASA validación	10	0	100

Tabla 8: Validación de las tasas de reducción de falsos positivos en secuencias Outdoor diurnas

- **Comentarios**

Al igual que en el caso Indoor, tras el proceso de validación cruzada se ha determinado un umbral único como se muestra en la gráfica superior de la **Figura 5-30**. De igual modo, este umbral permite diferenciar los falsos positivos debidos a variaciones de iluminación de los verdaderos positivos debidos a la aparición de intrusos o vehículos en algún instante de la secuencia.

En la tabla anterior la tasa de reducción de falsos positivos obtenida para el umbral calculado mediante *leave-one-out* representa la validación de los resultados que se presentaban en la **Figura 5-12** donde la curva *CORR lab-pixel* mostraba nuevamente una clasificación perfecta, con tasas de reducción de falsos positivos del 100%. Por lo tanto tras realizar el proceso de *leave-one-out* podemos concretar que los resultados expuestos anteriormente para el caso Outdoor diurno son fiables.

5.6.5 Validación de resultados del caso Outdoor nocturno

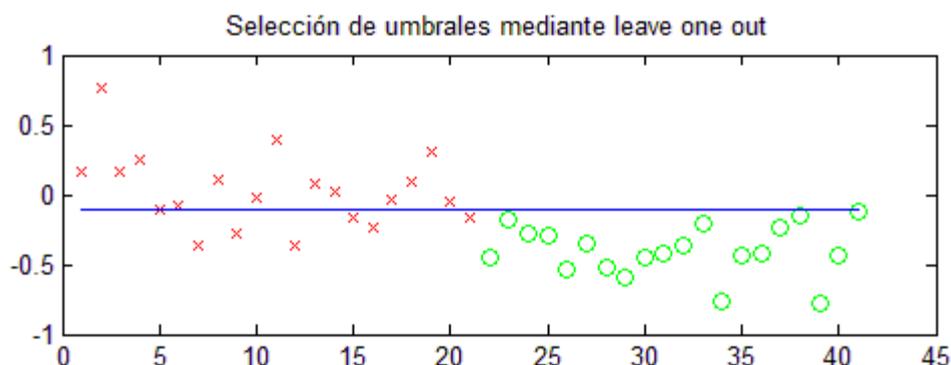


Figura 5-31: Validación de resultados para caso Outdoor nocturno

	# VP	# FP	% Reducción
CORR lab-pixel	20	7	66.67
TASA validación	20	7	66.67

Tabla 9: Validación de las tasas de reducción de falsos positivos en secuencias Outdoor nocturnas

● Comentarios

Para el caso Outdoor nocturno se ha aplicado el proceso de validación cruzada a la combinación de correcciones descritas en el apartado 5.5.3 seguidas del algoritmo de cálculo de correlación “En Labeling píxeles dentro del bounding box”, ya que éste es el procedimiento que arroja mejores resultados cuando se trabajaba con secuencias nocturnas (ver Figura 5-26).

En este proceso de validación también se ha determinado un umbral único como se muestra en la gráfica superior de la Figura 5-31. A diferencia de los dos casos anteriores este umbral no permite clasificar correctamente todos los falsos positivos debidos a variaciones de iluminación. Algunas secuencias con cambios de iluminación activarán el proceso de generación de alarmas como si fueran verdaderos positivos debidos a la aparición de intrusos o vehículos en algún instante de la secuencia. Debido a lo anterior, la tasa de reducción de falsos positivos obtenida para este umbral calculado mediante *leave-one-out* no muestra una clasificación perfecta como sucedía en los casos Indoor y Outdoor diurno. En este caso la tasa de reducción de falsos positivos alcanzada es del 67%, sin embargo dicha tasa es equivalente a la tasa de reducción representada en la ROC de la Figura 5-31 quedando validados los resultados del caso Outdoor nocturno.

6 Conclusiones y trabajo futuro

En este capítulo se extraerán las conclusiones del trabajo realizado en este proyecto, así como proponer algunas directrices para trabajos futuros relacionados con la detección de cambios de iluminación y búsqueda de nuevas aplicaciones en otros campos diferentes al de videovigilancia.

Las conclusiones expresadas a continuación se extraen en su mayor parte de los resultados obtenidos y mostrados en el **capítulo 5**. Es decir, en función de la reducción de la tasa de falsos positivos conseguida, así como en el comportamiento en tiempo real de las soluciones propuestas.

6.1 Conclusiones

A lo largo de este proyecto se han desarrollado varios algoritmos con el fin de reducir los falsos positivos debidos a cambios de iluminación en sistemas de videovigilancia basados en detección de movimiento utilizados para la detección de intrusos. Para ello, se procedió a realizar un exhaustivo análisis del problema, donde se identificaron los diferentes efectos provocados por los cambios de iluminación y se realizó una comparativa con los efectos provocados por intrusos o vehículos que entraban en el área videovigilada. A partir de dicho estudio se realizó una distinción en tres casos : Indoor, Outdoor diurno y Outdoor nocturno.

Para cada uno de estos tres casos se realizaron estudios de dependencias basados en la correlación entre la imagen de fondo y los elementos detectados en las secuencias de video. Para la detección de dichos elementos, se utilizaron técnicas de sustracción de fondo avanzadas e integradas en una herramienta para la segmentación de secuencias de video.

Los algoritmos diseñados para el análisis de dependencias se han basado concretamente en el cálculo del coeficiente de correlación del producto-momento de Pearson. Se observó tras las primeras aproximaciones que restringiendo el área de cálculo, los valores obtenidos para aquellos objetos que provenían de un cambio de iluminación eran superiores a los valores de correlación obtenidos para elementos detectados correspondientes con personas y/o vehículos. Tras esto, se procedió a restringir el área de cálculo a imágenes truncadas que encuadraban los elementos o blobs detectados y a computar únicamente los píxeles que conformaban dichos blobs. Los resultados obtenidos para estos métodos mostraban una mejora sustancial, donde los valores de correlación obtenidos permitían distinguir entre cambios de iluminación y cambios provocados por objetos reales como por ejemplo intrusos.

El método que ha aportado mejores resultados, denominado en este PFC “En labeling píxeles dentro del bounding box”, ha permitido encontrar una característica realmente diferenciadora.

Mediante un umbral se ha conseguido separar correctamente las secuencias donde había cambios de iluminación y las secuencias donde aparecían intrusos o vehículos. Es decir se han podido clasificar las secuencias con cambios de iluminación como verdaderos negativos (reduciendo así la tasa de falsos positivos inicial) y las secuencias con motivos reales para la generación de alarma como verdaderos positivos. Por tanto este sistema de reducción de falsos positivos no reduce los verdaderos positivos.

Para el caso Indoor, cuyos resultados han quedado reflejados en el **apartado 5.3**, se ha alcanzado una tasa de reducción de falsos positivos del 100% manteniendo simultáneamente la tasa de verdaderos positivos al 100% como se aprecia en la **Figura 5-4**. Es decir, el algoritmo “En labeling píxeles dentro del bounding box” permitiría distinguir la intrusión de una persona o vehículo de un cambio de iluminación, de forma que no se generaría ninguna falsa alarma en el sistema. Además se ha ofrecido una comparativa con el resto de métodos implementados, un resumen de dicha comparativa queda patente en la **Tabla 2**.

A la vista de los resultados ofrecidos en el **apartado 5.4** del caso Outdoor diurno, concretamente deteniéndose en la **Figura 5-12**, podemos concluir que la tasa de reducción de falsos positivos también es del 100% sin perder ningún caso de verdadero positivo. El método “En labeling píxeles dentro del bounding box” ha funcionado de forma igualmente satisfactoria que en el caso Indoor, consiguiendo que el sistema no genere falsas alarmas debidas a cambios de iluminación. En la **Tabla 4** se mostraba una comparativa de los diversos algoritmos, donde se ha podido observar que en este caso el método “En bounding boxes completos” también alcanzaba una tasa de reducción de falsos positivos del 100%.

El caso Outdoor nocturno ha supuesto mayor complejidad y se han realizado un número mayor de pruebas, ya que los métodos que ofrecían muy buenos resultados para los casos Indoor y Outdoor diurno no conseguían reducir apenas la tasa de falsos positivos en condiciones nocturnas. Para este caso se han llevado a cabo diversas técnicas de corrección de las imágenes de las secuencias utilizadas. En el **apartado 5.5.3** se ofrecían las gráficas del método con resultados más favorables, donde se combinaba un cambio de escala seguido de la eliminación del gradiente de iluminación antes de aplicar el cálculo de la correlación. Con esto se consiguió mejorar el comportamiento de los algoritmos de medida de dependencias para imágenes nocturnas o capturadas en condiciones de baja iluminación. Como se aprecia en la **Figura 5-26** para una tasa del 100% en la detección de verdaderos positivos se ha alcanzado una tasa de reducción de falsos positivos del 66,67% mediante el algoritmo “En labeling píxeles dentro del bounding box”.

Tras la obtención de resultados se decidió utilizar el método de validación cruzada *leave-one-out* para estimar las tasas de reducción de falsos positivos obtenidas en las pruebas realizadas. Este proceso de validación se ha llevado a cabo para el método “En labeling píxeles dentro del bounding box”, que es el que ofrecía mejores resultados, aplicándolo a los tres casos diferentes Indoor, Outdoor diurno y Outdoor nocturno.

En las **Tablas 7,8 y 9** se mostraba la tasa de reducción obtenida inicialmente y la tasa de reducción obtenida tras utilizar el umbral calculado mediante *leave-one-out* para cada uno de los tres casos siendo coincidente en todos ellos. Por tanto el proceso de *cross-validation* ha permitido validar los resultados.

Finalmente toda la labor realizada nos ha permitido obtener resultados muy favorables que han supuesto el alcance del objetivo inicialmente planteado, que daba nombre y motivaba el desarrollo de este proyecto: "Detección de falsos positivos en sistemas de vídeovigilancia debidos a cambios bruscos en la iluminación de la escena"

Se ha logrado además de detectar cambios de iluminación eliminar las falsas alarmas que provocan en los sistemas de videovigilancia, concretamente en sistemas inteligentes basados en detección de movimiento y empleados para la detección de intrusos o vehículos con fines de seguridad. La solución propuesta permite eliminar en su totalidad las falsas alarmas en condiciones de iluminación alta o media y reducir considerablemente estas alarmas en condiciones de baja iluminación o en condiciones nocturnas, todo ello en tiempo real.

El trabajo realizado ha derivado en un proceso de transferencia tecnológica ya que se integrará convenientemente en un sistema comercial. Así mismo, esta integración ha suscitado la solicitud de una patente con el título "METODO Y DISPOSITIVO DE DETECCIÓN DE CAMBIO DE ILUMINACIÓN PARA SISTEMA DE VISIÓN" con número de solicitud *EP 10382281.3*.

6.2 Trabajo futuro

A lo largo de este proyecto se ha hecho hincapié en que los cambios de iluminación dependen de varios factores como las diferentes fuentes de luz existentes (naturales y artificiales), el escenario donde se colocan los sistemas de captación, la calidad de las imágenes, etc. Esta variabilidad hace que los casos aquí tratados y los efectos descritos en el **capítulo 3** sean representativos de este problema pero no únicos. Esto conlleva la necesidad de continuar con esta línea de trabajo para asegurar su validez en otro tipo de escenarios y comprobar su robustez frente a situaciones más extremas.

Se sugiere también continuar la investigación utilizando diferentes sistemas de detección de movimiento para la segmentación de las secuencias, sobre todo aquellos cuyo modelado de fondo incluye el tratamiento de los cambios de iluminación como los mencionados en el **apartado 2.3.3**. Esto podría mejorar el comportamiento del sistema en una etapa anterior a la que se propone en este proyecto.

A su vez, el estudio de más métodos para el caso nocturno puede aportar nuevas ideas y mejorar las tasas de reducción de falsas alarmas a las aportadas en este documento.

Otra futura línea de investigación es la extensión del sistema propuesto con técnicas de *machine learning*. Esto se puede hacer de dos formas diferentes pero complementarias:

- 1) usar las técnicas para descubrir características diferenciadoras de los intrusos frente a cambio de iluminación, como por ejemplo aspectos de la geometría de los blobs (propuesto en el **apartado 4.4.9** de este proyecto), y entrenar un sistema de decisión usando dichas características;
- 2) para dotar a un sistema de la capacidad de aprender de errores pasados (falsas alarmas que han sido marcadas a posteriori por el usuario como tales). De tal forma el usuario tendría un sistema cuya eficiencia va mejorando a lo largo del tiempo.

Por último, se propone hacer estudios más extensos de los cambios de iluminación atendiendo a sus características no sólo espaciales sino temporales, lo cual puede ser llevado a cabo también por técnicas de machine learning que extraigan características de las secuencias de vídeo.

Referencias

- [1] Intelligent distributed surveillance systems: a review
M. Valera and S.A. Velastin. In *IEE Proc.-Vis. Image Signal Process.*, Vol. 152, No. 2, April 2005
- [2] Ronetti, N., and Dambra, C.: 'Railway station surveillance: the Italian case'. In *Foresti, G.L., Mahonen, P., and Regazzoni, C.S. (Eds.): 'Multimedia Video Based Surveillance Systems' (Kluwer Academic Publishers, Boston, 2000)*, pp. 13–20
- [3] Pellegrini, M., and Tonani, P.: 'Highway traffic monitoring'. In *Regazzoni, C.S., Fabri, G., and Vernazza, G. (Eds.): 'Advanced Video-based Surveillance Systems' (Kluwer Academic Publishers, Boston, 1998)*
- [4] Pozzobon, A., Sciutto, G., and Recagno, V.: 'Security in ports: the user requirements for surveillance system'. In *Regazzoni, C.S., Fabri, G., and Vernazza, G. (Eds.): 'Advanced Video-based Surveillance Systems' (Kluwer Academic Publishers, Boston, 1998)*
- [5] R J Evans, E L Brassington, "Video Motion Processing For Event Detection And Other Applications". In *IEE Annual Conference on Visual Image Engineering, VIE2003, University of Surrey*.
- [6] Intelligent network video. Understanding modern video surveillance systems. Fredrik Nilsson, Axis Communications. CRC Press, 2008
- [7] Xu, M., and Ellis, T.: 'Colour-Invariant Motion Detection under fast Illumination Changes'. In *Remagnino, P., Jones, G.A., Paragios, N., and Regazzoni, C.S. (Eds.): 'Video-based Surveillance Systems' (Kluwer Academic Publishers, Boston, 2002)*, pp. 101–120
- [8] B. Klare and S. Sarkar. Background subtraction in varying illuminations using an ensemble based on an enlarged feature set. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2009*
- [9] O. Javed, K. Shafique, and M. Shah. A hierarchical approach to robust background subtraction using color and gradient information. In *Motion and Video Computing, Proceedings of Workshop on*, pages 22–27, 2002.
- [10] S. Messelodi, C. Modena, N. Segata, and M. Zanin. A Kalman filter based background updating algorithm robust to sharp illumination changes. In *ICIAP 2005, 13th International Conference on Image Analysis and Processing*, pages 163–170, 2005.
- [11] CCTV for security professionals. Alan R. Matchett. Butterworth-Heinemann.

- [12] Brodsky, T., Cohen, R., Cohen-Solal, E., Gutta, S., Lyons, D., Philomin, V., and Trajkovic, M.: 'Visual surveillance in retail stores and in the home'. In: '*Advanced Video-based Surveillance Systems*' (Kluwer Academic Publishers, Boston, 2001), Chapter 4, pp. 50–61
- [13] McIvor, A. 2000. 'Background subtraction techniques'. In *Proceedings of Image & Vision Computing New Zealand 2000 IVCNZ'00*, Reveal Limited, Auckland, New Zealand.
- [14] A. McIvor, V. Zang, R. Klette. 'The Background Subtraction Problem for Video Surveillance Systems'. *International Workshop Robot Vision 2001, Auckland, New Zealand, February 2001*.
- [15] Fida El Baf and Thierry Bouwmans. 'Robust fuzzy statistical modeling of dynamic backgrounds'. In *IR videos.13 November 2009, SPIE Newsroom*. DOI: 10.1117/2.1200911.1842
- [16] M. Piccardi. 'Background subtraction techniques: a review'. In *Proc. of IEEE SMC 2004 International Conference on Systems, Man and Cybernetics, The Hague, The Netherlands, October 2004*.
- [17] Cheung S.C. and Kamath C. 'Robust techniques for background subtraction in urban traffic video'. *Video Communications and Image Processing, SPIE Electronic Imaging, San Jose, 2004*.
- [18] Ahmed M. Elgammal, David Harwood, and Larry S. Davis. 'Non-parametric model for background subtraction'. In *Proceedings of the 6th European Conference on Computer Vision-Part II, pages 751–767. Springer-Verlag, 2000*.
- [19] Computer Vision and Image processing. S. Nagabhushana. New Age International Publishers
- [20] C. Stauffer and W. Grimson. 'Adaptive background mixture models for real-time tracking'. In *IEEE Conference on Computer Vision and Pattern Recognition*, volume 2, 1999.
- [21] Ying-Li Tian, Max Lu, and Arun Hampapur. 'Robust and Efficient Foreground Analysis for Real-time Video Surveillance'. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01*
- [22] Ding Ying, Li Wen-hui, Fan Jing-tao and Yang Hua-min. 'Robust moving object detection under complex background'. In *Computer Science and Information Systems / ComSIS , 2010, vol. 7, br. 1, str. 201-210*
- [23] C. Ridder, O. Munkelt, and H. Kirchner. 'Adaptive background estimation and foreground detection using kalman filtering'. In *Proc. ICAM*, pages 193–199, 1995.

- [24] Cvetkovic, S.D., Bakker, Peter, Schirris, J., With, P.H.N. de (2009). 'Background estimation and adaptation model with light-change removal for heavily down-sampled video surveillance signals'. In *Proceedings of the IEEE International Conference on Image Processing (ICIP 2006), October 8-11, 2006, Atlanta, Georgia. (pp. 1829-1832).*
- [25] G. Gordon, T. Darrell, M. Harville, J. Woodfill. 'Background estimation and removal based on range and color'. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (Fort Collins, CO), June 1999.
- [26] Yuri Ivanov, Aaron Bobick and John Liu . 'Fast Lighting Independent Background Subtraction'. In *International Journal of Computer Vision, volume 37, number 2, (pp. 199-207), June 2000*
- [27] Ser-nam Lim, Anurag Mittal, Larry S. Davis, and Nikos Paragios. 'Fast illumination-invariant background subtraction using two views: Error analysis, sensor placement and applications'. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'05), 1:1071–1078, 2005.*
- [28] Paulidis, I., and Morellas, V.: 'Two examples of indoor and outdoor surveillance systems'. In *Remagnino, P., Jones, G.A., Paragios, N., and Regazzoni, C.S. (Eds.): 'Video-based Surveillance Systems' (Kluwer Academic Publishers, Boston, 2002), pp. 39–51*
- [29] Wisetphanichkij, S., Dejhan, K., Cheevasuvit, F., Mitatha S., Hanpipatpongsa, S. 'An improvement of geometric correlation of satellite image', *19th Asian Conference on Remote Sensing*, 1998.
- [30] H. Lu and P.D. Cary. 'Deformation measurements by digital image correlation: Implementation of a second-order displacement gradient'. In *Experimental Mechanics Volume 40, Number 4, 393-400*
- [31] K. Pearson, 'Mathematical contributions to the theory of evolution. III. Regression, heredity and panmixia.' *Philos. Trans. Royal Soc. London Ser. A* , 187 (1896) pp. 253–318
- [32] <http://www.mathworks.com/>
- [33.] Egan, J. P. (1975). *Signal Detection Theory and ROC Analysis*. Series in *Cognition and Perception*. Academic Press, New York.
- [34.] Swets, J. (1988). 'Measuring the accuracy of diagnostic systems.' *Science*, 240, 1285-1293.
- [35.] Spackman, K. A. (1989). 'Signal detection theory: Valuable tools for evaluating inductive learning' . In *Proceedings of the Sixth International Workshop on Machine Learning*, pp. 160-163 San Mateo, CA. Morgan Kaufman.

- [36.] Fawcett, T. (2004): 'ROC Graphs: Notes and Practical Considerations for Data Mining Researchers' . *Technical report. Palo Alto, USA: HP Laboratories.*
- [37] Stone, M. (1974). 'Cross-validated choice and assessment of statistical prediction.' *J.Royal Statistical Society, B*, 36, 111-147.
- [38] Stone M. (1978) 'Cross-validation: a review', *Math. Operationforsch. Statist., Ser. Statistics*, 9, 127–139.
- [39] Kohavi, Ron (1995): 'A study of cross-validation and bootstrap for accuracy estimation and model selection'. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence 2 (12)*. pp. 1137-1143. (San Mateo, Morgan Kaufman)
- [40] Andrew W. Moore. 'Cross-validation for detecting and preventing overfitting.' *School of Computer Science. Carnegie Mellon University.*
<http://www.cs.cmu.edu/~awm/tutorials>

ANEXO A: Presupuesto

1. Material

- Compra de ordenador personal 2000 €
- Material de oficina 150 €
- Total de ejecución material 2150 €

2. Gastos generales

- 18 % sobre Ejecución Material 387 €

3. Beneficio Industrial

- 6 % sobre Ejecución Material 129 €

4. Honorarios Proyecto

- 700 horas a 15 € / hora 10.500 €

5. Material fungible

- Gastos de impresión 160 €
- Encuadernación 20 €
- Total de Material fungible 180 €

6. Subtotal del presupuesto

- Subtotal Presupuesto 13.346 €

7. Total presupuesto

- 18% Subtotal Presupuesto 2402,28 €
- Total Presupuesto 15748,28 €

Madrid, Noviembre de 2010
La Ingeniera Jefa del Proyecto

Fdo.: Sara Rodríguez González
Ingeniera Superior de Telecomunicación

ANEXO B: Pliego de Condiciones

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de *Detección de falsos positivos en sistemas de videovigilancia debidos a cambios bruscos en la iluminación de la escena*. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho sistema. Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

Supuesto que la utilización industrial de los métodos recogidos en el presente proyecto ha sido decidida por parte de la empresa cliente o de otras, la obra a realizar se regulará por las siguientes:

Condiciones generales

1. La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.
2. El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.
3. En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.
4. La obra se realizará bajo la dirección técnica de un Ingeniero Superior de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.
5. Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.
6. El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.
7. Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha obra se haya ajustado a los preceptos de los pliegos de

condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.

8. Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.

9. Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.

10. Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.

11. Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

12. Las cantidades calculadas para obras accesorias, aunque figuren por partidaalzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.

13. El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.

14. Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.

15. La garantía definitiva será del 4% del presupuesto y la provisional del 2%.

16. La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.

17. La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.

18. Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.

19. El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.

20. Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.

21. El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.

22. Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.

23. Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad "Presupuesto de Ejecución de Contrata" y

anteriormente llamado "Presupuesto de Ejecución Material" que hoy designa otro concepto.

Condiciones particulares

La empresa consultora, que ha desarrollado el presente proyecto, lo entregará a la empresa cliente bajo las condiciones generales ya formuladas, debiendo añadirse las siguientes condiciones particulares:

1. La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.
2. La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.
3. Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.
4. En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.
5. En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.
6. Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.
7. Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.
8. Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.
9. Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.
10. La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta

objeto del presente proyecto para la realización de otras aplicaciones.

11. La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.

12. El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.

ANEXO C: Acreditación de méritos



A quien pueda interesar

Proyecto Fin de Carrera

"Detección de falsos positivos en sistemas de vídeo vigilancia debidos a cambios bruscos en la iluminación de la escena"

Por la presente certifico que Sara Rodríguez González ha realizado el Proyecto Fin de Carrera con título "Detección de falsos positivos en sistemas de vídeo vigilancia debidos a cambios bruscos en la iluminación de la escena" bajo mi supervisión en la compañía, desde Octubre 2009 a Octubre 2010.

Algunos de los resultados obtenidos en este proyecto serán convenientemente integrados en alguno de los sistemas comerciales de los que dispone la empresa. Con el fin de proteger la investigación realizada se ha solicitado una patente con el título "MÉTODO Y DISPOSITIVO DE DETECCIÓN DE CAMBIO DE ILUMINACIÓN PARA SISTEMAS DE VISIÓN"

Se incluye como Anexo el justificante de la presentación telemática de la solicitud de patente con fecha 27 de Octubre de 2010 y con Application number EP 10382281.3.

Firmado: Eduardo Cemeño
Director General

VAELSYS FORMACION Y DESARROLLO, S.L.
Crt. CANILLAS, 99 9º 3 MADRID
C.I.F.: B-84071976
TLE: +34 497 34 82



MINISTERIO
DE INDUSTRIA, TURISMO
Y COMERCIO



Oficina Española
de Patentes y Marcas

Acknowledgement of receipt

We hereby acknowledge receipt of your request for grant of a European patent as follows:

Submission number	300018297	
Application number	EP10382281.3	
File No. to be used for priority declarations	EP10382281	
Date of receipt	27 October 2010	
Your reference	P6395EP00	
Applicant	VAELSYS FORMACIÓN Y DESARROLLO, S.L.	
Country	ES	
Title		
Documents submitted	package-data.xml application-body.xml SPECNONEPO.pdf#20101026 P6395EP00 DET CAMBIOS DE LUZ - VF_TS.pdf (26 p.) f1002-2.pdf (1 p.) f1002-4.pdf (1 p.) f1002-6.pdf (1 p.)	ep-request.xml ep-request.pdf (4 p.) f1002-1.pdf (1 p.) f1002-3.pdf (1 p.) f1002-5.pdf (1 p.)
Submitted by	CN=F. Prieto Moran 8258,O=ABG Patentes, S.L.,C=ES	
Method of submission	Online	
Date and time receipt generated	27 October 2010, 16:22:41 (CEST)	
Official Digest of Submission	C2:BB:AC:45:46:0C:3C:E7:30:7B:D2:D6:97:2A:20:B2:B8:A1:78:89	

