

UNIVERSIDAD AUTONOMA DE MADRID

ESCUELA POLITECNICA SUPERIOR



PROYECTO FIN DE CARRERA

**ANÁLISIS COMPARATIVO DE TÉCNICAS
DE SEGMENTACIÓN DE SECUENCIAS DE
VÍDEO BASADAS EN EL MODELADO
DEL FONDO**

-PROYECTO FIN DE CARRERA-

**Sonsoles Herrero Martín
Abril de 2009**

Análisis comparativo de técnicas de segmentación de secuencias de video basadas en el modelado del fondo.

AUTOR: Sonsoles Herrero Martín

TUTOR: Jesús Bescós Cano



Video Processing and Understanding Lab

Dpto. de Ingeniería Informática

Escuela Politécnica Superior

Universidad Autónoma de Madrid

Abril de 2009

Palabras clave

Segmentación, vídeo, modelado de fondo, sustracción de fondo, objetos en movimiento, frente, fondo, FG ('Foreground'), BG ('Background'), Diferencia de Frentes, Mediana, Generación de Histogramas, Media móvil, Gaussiana Simple (SG), Mezcla de Gaussianas (MoG), Estimación de la densidad de núcleo (KDE), Gamma, HMM, precisión, 'recall', 'fscore'.

Resumen

El principal objetivo de este PFC es el análisis comparativo de técnicas de segmentación basadas en el modelado de fondo (denominadas también técnicas de sustracción de fondo), que permitan, a partir de una secuencia de video captada por una cámara estática, detectar las áreas pertenecientes al primer plano y al fondo de la escena.

Con este fin, se han estudiado, implementado y evaluado los algoritmos más relevantes en cuanto a modelado de fondo se refiere, para ofrecer tanto un análisis de los parámetros de cada técnica como una comparativa entre ellos y, de esta manera, contribuir al estado del arte.

Abstract

The main objective of this master thesis (PFC) is the comparative analysis of segmentation techniques based on background removal (sometimes referred as background subtraction). These techniques aim to detect foreground and background areas of the video sequences captured by a static camera.

For this purpose, we have studied, implemented and evaluated the most relevant segmentation techniques based on background subtraction providing an analysis of the parameters of each technique and a comparison between them and thus contributing to the state of art.

Agradecimientos

En primer lugar, quiero agradecer a mi tutor, Jesús Bescós, por permitir que realizase este Proyecto Fin de Carrera en el VPULab, y por su apoyo y dedicación durante todos estos años de estudios.

También quiero agradecer a José María Martínez por su confianza, ánimos y enseñanzas a lo largo de la Carrera.

Gracias a mi familia, a mis padres y especialmente a mis hermanas, por sus consejos y todo el amor y apoyo que me han dado durante toda mi vida que me ha permitido sacar energía para seguir estudiando, aún con más ganas cuando las fuerzas y los ánimos decaían.

Muchas gracias a Juan Carlos por todo su cariño y por saber darme en cada momento todo lo que necesitaba desde el día que le conocí. Espero estar a su lado en muchos otros proyectos que emprendamos.

Gracias a los miembros del VPULab por aconsejarme y ayudarme con el proyecto. Y a Miguel Ángel por estar siempre dispuesto a responder a mis dudas.

A todos mis amigos y amigas de Tres Cantos, en especial a Ana, Sonia, Curro y Albertillo porque siempre me han escuchado y animado cuando lo necesitaba.

También quiero darles las gracias a todos los amigos que he tenido la gran oportunidad de conocer en la Facultad y que siempre llevaré en el corazón y recordaré con la más enorme de las sonrisas: gracias Lucas por ser un apoyo constante y el mejor compañero de prácticas, gracias Moni por todos los momentos que hemos pasado juntas, gracias Vero porque tus sorpresas nos llenaban de alegría, gracias Ele por hacerme recordar lo importante que es estar cerca de los amigos, gracias Esther por ser tan buena chica y confiar en mí. Y gracias Gus, Kiko, Chus y Pablo por haber compartido con todas nosotras estos años inolvidables.

Muchas gracias a todos los profesores que he tenido y que me transmitido su pasión por la enseñanza que me han hecho llegar cada día a clase con ilusión y llenarme de ganas por aprender.

Y por último quiero dedicar este proyecto a la gente que me ha hecho sonreír y los que he hecho sonreír, porque esa felicidad es la que hace posible enfrentarse a los retos de cada día con alegría.

INDICE DE CONTENIDOS

1	Introducción	1
1.1	Motivación.....	1
1.2	Objetivos.....	3
1.3	Organización de la memoria.....	4
2	Antecedentes y estado del arte.....	5
2.1	Introducción a la segmentación de objetos en movimiento	5
2.2	Técnicas de pre-procesado en segmentación de video.....	8
2.3	Modelado de fondo	10
2.3.1	Modelo de representación del fondo	14
2.3.1.1	Modelos básicos	15
2.3.1.2	Modelos paramétricos.....	16
2.3.1.3	Modelos no paramétricos.....	19
2.3.2	Modelo de inicialización del fondo	21
2.3.3	Modelo de adaptación del fondo.....	22
2.4	Detección de frente	24
2.4.1	Técnicas de detección de frente.....	24
2.4.1.1	Basadas en diferencia	24
2.4.1.2	Basadas en estadísticos.....	25
2.4.2	Selección del umbral	27
2.5	Técnicas de post-procesado.....	28
2.6	Medidas de evaluación de algoritmos	29
3	Diseño	33
3.1	Modelos básicos.....	34
3.1.1	Diferencia con el cuadro anterior	34
3.1.2	Mediana de los ‘N’ cuadros previos	36
3.1.3	Generación del fondo basada en histogramas.....	37
3.2	Métodos paramétricos.....	40
3.2.1	Gaussiana Simple (SG).....	40
3.2.1.1	Filtro de análisis espacial.....	43
3.2.2	Mezcla de Gaussianas (MoG).....	46
3.2.3	Método Gamma	52
3.3	Métodos no paramétricos.....	55
3.3.1	Estimación de Densidad del Núcleo (KDE)	55
3.3.1.1	Mejora del KDE: Filtro de análisis espacial y temporal.....	59
3.3.1.2	Mejora del método KDE: Combinación de modelos a corto y largo plazo.....	62
3.4	Método combinado: Histogramas y Gaussiana	63
4	Integración en una arquitectura para su evaluación	69
5	Resultados: pruebas y análisis comparativo	71
5.1	Descripción del dataset utilizado.....	71
5.1.1	Factores críticos del objeto	72
5.1.2	Factores críticos del fondo.....	73
5.2	Diseño de pruebas experimentales	76
5.3	Métricas utilizadas.....	78
5.4	Selección de parámetros.....	80

5.5 Análisis comparativo	85
5.5.1 Tiempos de ejecución	85
5.5.2 Calidad de la segmentación	87
5.5.2.1 <i>Comparativa de métodos básicos</i>	87
5.5.2.1.1 Análisis cuantitativo	87
5.5.2.1.2 Análisis cualitativo	105
5.5.2.1.3 Conclusiones	107
5.5.2.2 <i>SG vs SG con mejoras</i>	109
5.5.2.3 <i>KDE vs KDE con mejoras</i>	116
6 Conclusiones y trabajo futuro	123
6.1 Trabajo futuro	124
Referencias	127
Glosario	133
Anexos	I
Anexo A	I

INDICE DE FIGURAS

FIGURA 1-1: SEGMENTACIÓN EN UNA SECUENCIA DE VIDEO EN FRENTE Y FONDO.	1
FIGURA 2-1: DIAGRAMA DE BLOQUES DE UN ALGORITMO DE BGS.	8
FIGURA 2-2: SECUENCIA UNIMODAL (SUPERIOR) Y MULTIMODAL (INFERIOR) Y LOS HISTOGRAMAS DEL PÍXEL SITUADO EN LA POSICIÓN (80, 313) A LA DERECHA.	11
FIGURA 2-3: SECUENCIA CON DISTINTA ILUMINACIÓN DEBIDO A: (A) NIEBLA Y (B) LLUVIA. SECUENCIA (D) CON UN OBJETO APARECIDO (FOLIO) Y DESAPARECIDO (RATÓN) CON RESPECTO A LA (C).	12
FIGURA 2-4: SECUENCIA (A) Y FONDO (B) INICIALIZADO ERRÓNEAMENTE A LA PRIMERA IMAGEN CAPTURANDO EN ÉL OBJETOS EN MOVIMIENTO. SECUENCIA (C) Y FONDO (D) ACTUALIZADO CORRECTAMENTE.	13
FIGURA 2-5: EJEMPLO DE LA ACTUALIZACIÓN SELECTIVA SOBRE UNA SECUENCIA DE VIDEO EN EL INSTANTE T.	23
FIGURA 2-6: EJEMPLO DE LA ACTUALIZACIÓN A CIEGAS SOBRE UNA SECUENCIA DE VIDEO EN EL INSTANTE T.	24
FIGURA 2-7: MÁSCARA DEL FRENTE.	27
FIGURA 2-8: DIAGRAMA DE UN SISTEMA DE SEGMENTACIÓN DE OBJETOS CON TÉCNICAS DE POST-PROCESADO.	28
FIGURA 2-9: EJEMPLO DE LAS CURVAS ROC OBTENIDAS CON DISTINTOS ALGORITMOS DE SEGMENTACIÓN DE OBJETOS SOBRE DIFERENTES SECUENCIAS DE UN BANCO DE PRUEBAS. ...	32
FIGURA 3-1: DIAGRAMA DE FLUJO DE LOS ALGORITMOS DE SEGMENTACIÓN.	33
FIGURA 3-2: DIFERENCIA ENTRE LA IMAGEN ACTUAL Y LA DE FONDO Y CÁLCULO DE LA IMAGEN BINARIA DEL FRENTE.	35
FIGURA 3-3: HISTOGRAMA DEL PÍXEL LOCALIZADO EN LA POSICIÓN (162,176) DE LA IMAGEN 46 DE LA SECUENCIA TENIS1 DE LOS VIDEOS DE PRUEBA.	38
FIGURA 3-4: DISTRIBUCIÓN GAUSSIANA DEL PÍXEL (188, 321) DE LA SECUENCIA HOMBRE HAMBRIENTO EN EL INSTANTE 184.	41
FIGURA 3-5: DISTRIBUCIÓN GAUSSIANA DE UN PÍXEL.	42
FIGURA 3-6: IMAGEN, MODELO SG DE FONDO Y FRENTE CAPTURADOS EN EL INSTANTE 130 DE LA SECUENCIA. EL CÍRCULO AMARILLO INDICA EL PÍXEL SELECCIONADO Y EL RECTÁNGULO ROJO A SU ALREDEDOR ES EL VECINDARIO ESCOGIDO PARA REALIZAR EL FILTRADO ESPACIAL.	44
FIGURA 3-7: PARÁMETROS DEL MODELO DE FONDO GAUSSIANO DEL PÍXEL SITUADO EN LA POSICIÓN (96, 226) EN EL INSTANTE 130 DE LA SECUENCIA.	45
FIGURA 3-8: DISTRIBUCIÓN DE LA MEZCLA DE GAUSSIANAS DE UN PÍXEL.	46

FIGURA 3-9: IMAGEN, MODELO MOG DEL FONDO Y FRENTE EN LA EVALUACIÓN DEL PÍXEL (88, 233) DE LA SECUENCIA ‘BAILE 2’ EN LA IMAGEN 83.....	50
FIGURA 3-10 : EJEMPLO DEL MODELO MOG DEL PÍXEL (88, 233) DE LA SECUENCIA BAILE MULTIMODAL EN LA IMAGEN 83.....	51
FIGURA 3-11: REPRESENTACIÓN DE LAS FUNCIONES DE NÚCLEO GAUSSIANAS CENTRADAS EN VALORES ANTERIORES Y EVALUADAS EN EL VALOR ACTUAL DEL PÍXEL LOCALIZADO EN LA POSICIÓN (234, 338).....	56
FIGURA 3-12: REPRESENTACIÓN DE LA IMAGEN PROCESADA Y EL MODELO DE FONDO EN EL INSTANTE 138.....	58
FIGURA 3-13: BUFFER DEL PÍXEL (2, 3) QUE CONTIENE LOS L VALORES ANTERIORES AL INSTANTE DE PROCESAMIENTO 138.....	59
FIGURA 3-14: MODELO DE KDE CON FILTRADO ESPACIO-TEMPORAL EN LA SECUENCIA HOMBRE HAMBRIENTO.....	60
FIGURA 3-15: VECINDARIO 5 X 5 DEL PÍXEL (233, 326) EN EL INSTANTE 58.....	61
FIGURA 3-16: BUFFER DEL PÍXEL (X, Y) QUE CONTIENE LOS L VALORES ANTERIORES AL INSTANTE DE PROCESAMIENTO.....	63
FIGURA 3-17: BUFFER DEL PÍXEL (x, y) QUE CONTIENE LOS L VALORES ANTERIORES AL INSTANTE DE PROCESAMIENTO EN UNA VENTANA W (W >> L).....	63
FIGURA 3-18: EJEMPLOS DE LOS HISTOGRAMAS DE UN PÍXEL: A) FONDO ESTÁTICO, B) FONDO PEQUEÑAS VARIACIONES DE INTENSIDAD. C) FONDO DINÁMICO.....	64
FIGURA 4-1: ARQUITECTURA DEL SISTEMA DE MODELADO DE FONDO Y EVALUACIÓN.....	70
FIGURA 5-1: TIEMPOS MEDIOS DE PROCESAMIENTO EXPRESADOS EN SEGUNDOS POR IMAGEN LOS MÉTODOS MÁS RÁPIDOS DE MODELADO DE FONDO.....	86
FIGURA 5-2: TIEMPOS MEDIOS DE PROCESAMIENTO EXPRESADOS EN SEGUNDOS POR IMAGEN LOS MÉTODOS MÁS LENTOS DE MODELADO DE FONDO.....	86
FIGURA 5-3: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL FOREGROUND.....	88
FIGURA 5-4: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	89
FIGURA 5-5: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S1.....	90
FIGURA 5-6: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL FOREGROUND.....	91
FIGURA 5-7: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	91
FIGURA 5-8: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S2.....	93
FIGURA 5-9: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL FOREGROUND.....	94

FIGURA 5-10: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	94
FIGURA 5-11: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S3.	95
FIGURA 5-12: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL FOREGROUND.....	96
FIGURA 5-13: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	97
FIGURA 5-14: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S4.	98
FIGURA 5-15: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL FOREGROUND.....	99
FIGURA 5-16: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	99
FIGURA 5-17: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S5.	100
FIGURA 5-18: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL FOREGROUND.....	101
FIGURA 5-19: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	102
FIGURA 5-20: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S6.	103
FIGURA 5-21: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	104
FIGURA 5-22: GRÁFICA DE ERRORES EN LA DETECCIÓN DEL BACKGROUND.....	104
FIGURA 5-23: DIAGRAMAS DE BARRAS (DE IZQUIERDA A DERECHA) DE: PRECISIÓN, RECALL, FSCORE EN LA SECUENCIA S7.	105
FIGURA 5-24: IMAGEN Y FRENTES OBTENIDOS CON LOS DISTINTOS MÉTODOS DE SEGMENTACIÓN DE OBJETOS SOBRE DOS DE LAS SECUENCIAS DEL BANCO DE PRUEBAS: BAILE 2(MULTIMODAL) Y HAMBRE (UNIMODAL).	107
FIGURA 5-25: DIAGRAMA DE BARRAS DE LA PRECISIÓN EN CERO EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO SG CON Y SIN MEJORA.....	111
FIGURA 5-26: DIAGRAMA DE BARRAS DE LA PRECISIÓN EN UNO EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO SG CON Y SIN MEJORA.....	111
FIGURA 5-27: DIAGRAMA DE BARRAS DEL <i>RECALL</i> EN CEROS EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO SG CON Y SIN MEJORA.....	112
FIGURA 5-28: DIAGRAMA DE BARRAS DEL <i>RECALL</i> EN UNOS EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO SG CON Y SIN MEJORA.....	112
FIGURA 5-29: IMAGEN Y MÁSCARAS DE FRENTE DEL ALGORITMO SG SIN Y CON MEJORA CAPTURADAS EN EL INSTANTE 61.....	113

FIGURA 5-30: MÁSCARAS DE FRENTE DEL ALGORITMO SG SIN Y CON MEJORA DIVIDIDAS EN PÍXELES CORRECTOS Y DETECTADOS DE VALOR 1.....	114
FIGURA 5-31: MÁSCARAS DE FRENTE DEL ALGORITMO SG SIN Y CON MEJORA DIVIDIDAS EN PÍXELES CORRECTOS Y DETECTADOS DE VALOR 0.....	114
FIGURA 5-32: MÁSCARAS DE FRENTE DEL ALGORITMO SG SIN Y CON MEJORA DIVIDIDAS EN PÍXELES CORRECTOS Y VERDADEROS (<i>GROUND TRUTH</i>) DE VALOR 1.....	115
FIGURA 5-33: MÁSCARAS DE FRENTE DEL ALGORITMO SG SIN Y CON MEJORA DIVIDIDAS EN PÍXELES CORRECTOS Y VERDADEROS (<i>GROUND TRUTH</i>) DE VALOR 0.....	115
FIGURA 5-34: DIAGRAMA DE BARRAS DE LA PRECISIÓN EN CERO EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO KDE CON Y SIN MEJORAS.....	118
FIGURA 5-35: DIAGRAMA DE BARRAS DE LA PRECISIÓN EN UNO EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO KDE CON Y SIN MEJORAS.....	118
FIGURA 5-36: DIAGRAMA DE BARRAS DEL <i>RECALL</i> EN CERO EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO KDE CON Y SIN MEJORAS.....	119
FIGURA 5-37: DIAGRAMA DE BARRAS DEL <i>RECALL</i> EN UNO EXPRESADA EN % POR CADA SECUENCIA PROCESADA POR CADA MÉTODO KDE CON Y SIN MEJORAS.....	119
FIGURA 5-38: IMAGEN Y MÁSCARAS DE FRENTE DEL ALGORITMO KDE SIN Y CON MEJORAS CAPTURADAS EN EL INSTANTE 188.....	121

INDICE DE TABLAS

TABLA 2-1. MEDIDAS DE EVALUACIÓN.....	30
TABLA 5-1: FACTORES CRÍTICOS DEL FRENTE Y FONDO [86].....	72
TABLA 5-2: CARACTERÍSTICAS DE LAS SECUENCIAS DE VIDEO DE PRUEBA ('DATASET')	76
TABLA 5-3: VECTORES DE PRUEBA PARA CADA PARÁMETRO DE CADA ALGORITMO.	78
TABLA 5-4: PARÁMETROS ESTADÍSTICOS DE LA EVALUACIÓN.....	78
TABLA 5-5. PARÁMETROS ÓPTIMOS PARA LOS MÉTODOS: DIFERENCIA DE FRENTE, MEDIANA E HISTOGRAMAS.	80
TABLA 5-6. PARÁMETROS ÓPTIMOS PARA EL MÉTODO GAUSSIANA SIMPLE.	81
TABLA 5-7. PARÁMETROS ÓPTIMOS PARA EL MÉTODO MEZCLA DE GAUSSIANAS (MOG).	82
TABLA 5-8. PARÁMETROS ÓPTIMOS PARA EL MÉTODO GAMMA.....	83
TABLA 5-9. PARÁMETROS ÓPTIMOS PARA EL MÉTODO ESTIMACIÓN DE NÚCLEO GAUSSIANO (KDE).	84
TABLA 5-10. PARÁMETROS ÓPTIMOS PARA EL MÉTODO COMBINADO DE HISTOGRAMAS Y GAUSSIANA SIMPLE.....	84
TABLA 5-11. TIEMPOS DE PROCESAMIENTO EN SEGUNDOS POR IMAGEN DE CADA MÉTODO DE SEGMENTACIÓN BASADO EN MODELADO DE FONDO.	86
TABLA 5-12 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S1.....	89
TABLA 5-13 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S2.....	92
TABLA 5-14 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S3.....	95
TABLA 5-15 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S4.....	97
TABLA 5-16 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S5.....	100
TABLA 5-17 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S6.....	102
TABLA 5-18 . RESULTADOS COMPARATIVOS DE TODOS LOS MÉTODOS EJECUTADOS SOBRE LA SECUENCIA S7.....	105
TABLA 5-19. PARÁMETROS ÓPTIMOS PARA EL MÉTODO GAUSSIANA SIMPLE CON Y SIN FILTRO ..	110

TABLA 5-20 . PARÁMETROS ÓPTIMOS PARA EL MÉTODO KDE SIN MEJORAS, CON LA MEJORA DEL FILTRO Y CON LA MEJORA DE LAS DOS COLAS.....	117
TABLA 5-21 . RESULTADOS COMPARATIVOS DEL MÉTODO KDE CON Y SIN MEJORAS.....	120
TABLA 5-22 . RESULTADOS COMPARATIVOS DEL MÉTODO KDE CON LAS MEJORAS: MEJORA 1 (FILTRO ESPACIO-TEMPORAL) Y MEJORA 2 (DOS COLAS DE PROCESAMIENTO).....	120

1 Introducción

1.1 Motivación

Al observar una imagen o un video es inmediato para un humano intentar hacer una descomposición o *segmentación* del mismo en distintos objetos. Dichos objetos tendrán una interpretación semántica dentro de la escena en base a distintos criterios: estructura, textura, color, movimiento u otra característica relevante. Esta segmentación que resulta sencilla de realizar en los humanos es uno de los mayores desafíos del procesamiento de imágenes y video.

La segmentación de regiones en movimiento en secuencias de video [1] constituye la primera etapa de análisis en multitud de áreas de aplicación como compresión de video [2], video-vigilancia [3], indexación de contenidos [4] e incluso efectos especiales en cine o TV [5] .

El objetivo de la segmentación de objetos en movimiento es distinguir aquellas regiones distintas al fondo más o menos estático de la secuencia de video capturada por una cámara generalmente fija. Así pues, esta etapa de análisis persigue la segmentación o clasificación de la secuencia de video en dos tipos de regiones: **frente**¹ (o *foreground*) y **fondo** estático de la escena (o *background*).

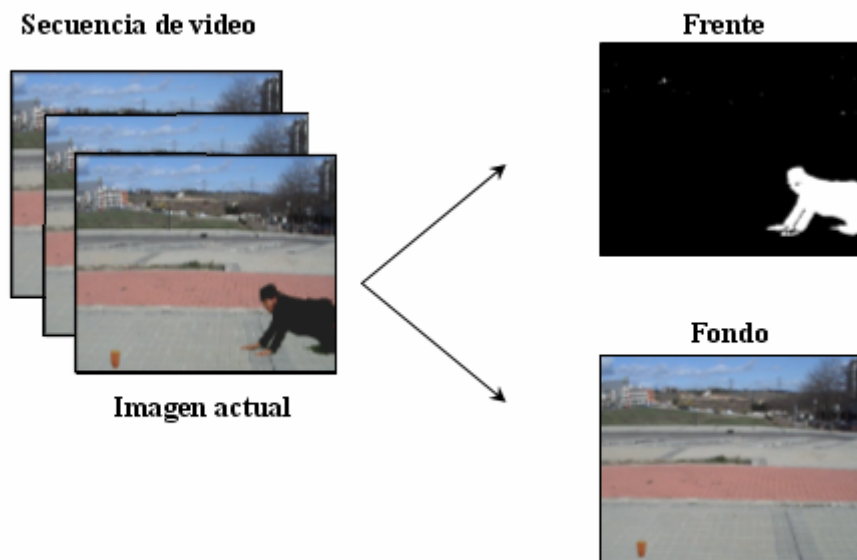


Figura 1-1: Segmentación en una secuencia de video en frente y fondo.

¹ A largo del documento se utilizarán indistintamente las palabras frente, primer plano o región en movimiento para referirse a los objetos en movimiento detectados en la secuencia de video

En este área de análisis, debido a las características específicas que presentan algunas secuencias de video capturadas por cámaras fijas (que son el tipo de secuencias con las que se va a trabajar en este proyecto), como puede ser la existencia o no de movimientos erráticos de la cámara (caso de una cámara de mano) o la presunción de que el primer plano suele ser distinto al fondo de la escena, los investigadores han desarrollado una amplia gama de técnicas basadas en el modelado del fondo de la escena conocidas como técnicas de substracción de fondo o algoritmos BGS (del inglés *BackGround Subtraction techniques*), que son utilizados junto con una serie de técnicas de pre/post-procesamiento que permiten mejorar el rendimiento final de la detección de regiones de primer plano.

La utilización de unas técnicas u otras depende de la aplicación deseada, es decir, de la información que se desee extraer, y está condicionada por las características que presenta la secuencia de video a analizar.

Existen varias clasificaciones de los métodos de segmentación de objetos en movimiento en la literatura [6][7][8], cada una agrupando los métodos con diferentes criterios; y las categorías en que se realiza la clasificación varían dependiendo de los trabajos. No obstante, en general, en entornos controlados, se diseñan algoritmos sencillos y de rápida ejecución (*'Frame difference'* [9], *'Runnin Average'* [10]) que a la vez son poco robustos frente al ruido y entornos complejos multimodales (donde los píxeles que conforman el modelo de fondo pueden presentar varios modos o patrones de variación). En cambio, en entornos no controlados, como pueden ser secuencias capturadas con iluminación exterior o con movimiento en el fondo (hojas, agua, fuego, etc.), para discernir entre el *ruido* existente en las imágenes de la secuencia de video y lo que realmente se está moviendo (frente) se requieren métodos más robustos (*'Simple Gaussian'* [11], *'Mixture of Gaussians'* [12], *'Kernel Density Estimation'* [13]) La discriminación de fondo y frente en estos casos se puede realizar mediante la estimación de la densidad de probabilidad (pdf) de las características de los objetos o mediante técnicas más generales, como pueden ser almacenamiento de los últimos valores del píxel, cálculo de rangos de valores del píxel, ajustes de funciones de predicción lineal, etc.

La motivación de este proyecto se basa en la importancia de la etapa de segmentación de objetos en movimiento en la cadena de procesamiento de un sistema: una arquitectura robusta para un sistema de análisis de alto nivel (seguimiento, detección de comportamientos, conteo de personas, etc.) requiere una etapa eficiente de extracción de objetos en movimiento.

1.2 Objetivos

El principal objetivo de este PFC es el desarrollo y evaluación de las técnicas de segmentación más destacadas del estado del arte en el ámbito de la substracción de fondo basadas en modelado de fondo. En este proyecto se realizará un análisis exhaustivo de cada técnica y se compararán las técnicas implementadas para saber en función de los distintos sistemas de análisis, el método de segmentación más adecuado a la información requerida y la secuencia de video a procesar. Para lograr los retos que se plantean en este proyecto, las tareas que se llevarán a cabo son las siguientes:

- i. **Estudio del estado del arte actual:** Inicialmente, se realizará un estudio de las diferentes técnicas de segmentación basadas en modelado de fondo, centrandó nuestra atención en las más representativas del estado del arte.
- ii. **Selección e implementación de los algoritmos más relevantes de substracción de fondo:** Una vez concluida la fase de recopilación de información y estudio, se pasará a la implementación de las técnicas estudiadas más relevantes.
- iii. **Desarrollo e implementación de un sistema de evaluación:** Se desarrollará un sistema de evaluación de segmentación de objetos en movimiento genérico con los siguientes objetivos:
 - o Determinar los parámetros más significativos de cada algoritmo y su influencia en las diferentes secuencias de video que forman parte del banco de pruebas del proyecto.
 - o Comparar los distintos métodos que se han implementado a través de diferentes medidas de evaluación; precisión, sensibilidad, especificidad, curvas ROC, *Fscore* y tiempos de ejecución.
 - o Comparar los resultados obtenidos de los algoritmos implementados con los que han obtenido otros investigadores diseñando el mismo tipo de algoritmos.
- iv. **Diseño de un sistema de detección de movimiento y evaluación:** se implementará un sistema en el que introduciremos nuestros algoritmos de substracción de fondo basados en modelado de fondo para evaluar el funcionamiento de cada algoritmo.
- v. **Análisis de resultados y conclusiones:** se extraerán una serie de conclusiones de cada algoritmo en función de la influencia de los diversos factores de las secuencias en los resultados y se ofrecerá una comparativa de los algoritmos.

En aras de realizar un análisis exhaustivo de los algoritmos, este Proyecto Fin de Carrera (PFC) utiliza un conjunto de datos de pruebas generado por el VPULab, que podemos encontrar en la siguiente dirección de Internet <http://www-vpu.ii.uam.es/CVSG/>.

1.3 Organización de la memoria

La memoria consta de los siguientes capítulos:

- **Capítulo 1:** Introducción, objetivos y motivación del proyecto.
- **Capítulo 2:** Estudio de los algoritmos de segmentación para la detección de objetos en movimiento: algoritmos de pre-procesado, métodos de segmentación basados en modelado de fondo y técnicas de post-procesado.
- **Capítulo 3:** Diseño e implementación de algoritmos de segmentación basados en modelado de fondo: métodos básicos, métodos paramétricos, métodos no paramétricos y combinación de algoritmos para un procesado a largo y corto plazo.
- **Capítulo 4:** Integración de los métodos en una arquitectura básica para su evaluación.
- **Capítulo 5:** Pruebas, resultados y análisis comparativo: descripción del conjunto de datos de prueba utilizado, pruebas, métricas, costes computacionales y resultados.
- **Capítulo 6:** Conclusiones y trabajo futuro.

2 Antecedentes y estado del arte

En este capítulo se describen conceptos básicos relacionados con la segmentación de objetos en movimiento para secuencias de video, así como una revisión de las técnicas más representativas.

Este estudio se encuentra dividido en las siguientes secciones: segmentación de objetos en movimiento (sección 2.1), técnicas más relevantes en el pre-procesado para la segmentación (sección 2.2), métodos de segmentación basados en modelado de fondo (sección 2.3) y algoritmos de post-procesado de video en segmentación (sección 2.4).

2.1 Introducción a la segmentación de objetos en movimiento

Actualmente, el uso de sistemas inteligentes de análisis de secuencias de video está cada vez más presente en una gran variedad de sistemas: video-vigilancia [15][16], indexación de contenidos [17], cine y TV [18], compresión de video [19]. Debido a la creciente cantidad de información visual generada por las cámaras y sensores de estos sistemas, es necesario desarrollar herramientas de análisis automático, y que operen en tiempo real en ciertos dominios de aplicación (por ejemplo, video-vigilancia), que permitan extraer las regiones de interés de la secuencia de video analizada. En este contexto, el primer problema es la localización de la región donde sucede algo relevante. Esta operación suele conocerse como segmentación.

La segmentación de objetos habitualmente tiene como objetivo discriminar los objetos en movimiento del primer plano (frente o *foreground*) de una imagen, del resto de los objetos o fondo (*background*). En el caso de una escena grabada por una cámara fija, que es el caso analizado por este proyecto, las técnicas de segmentación más eficaces son las basadas en el modelado y posterior substracción del fondo.

Existen muchas técnicas distintas de segmentación de secuencias de vídeo dependiendo del tipo de secuencias y de los objetivos deseados. La segmentación que emplearemos en este proyecto está basada en el modelado de fondo; consiste en una serie de algoritmos que mantienen un modelo matemático de apariencia del fondo a través del cual, para cada imagen, se pueden extraer (por comparación) los objetos en movimiento (primer plano) y el fondo.

Debido a que las secuencias de video no son ideales (es decir, objetos en movimiento sobre un fondo completamente estático), la segmentación automática de objetos presenta múltiples complicaciones, siendo una de las tareas más complicadas dentro del procesado de video. Como se ha señalado por Toyama en [20], Elgammal en [21] y Harville en [22], un buen algoritmo capaz de detectar objetos en movimiento debe lograr de manera computacionalmente eficiente, y preferentemente en tiempo real, solventar los siguientes problemas:

- **Cambios de iluminación de la escena a analizar (rápidos/lentos):** Son variaciones de iluminación que se pueden producir a lo largo del día si la escena ha sido capturada al aire libre [23], o tratándose de una secuencia de interior, por las distintas fuentes de iluminación que pueden estar apagadas o encendidas. Estos cambios pueden ser malinterpretados como parte de los objetos en movimiento o hacer que el fondo quede obsoleto si éste no se actualiza correctamente.
- **Sombras y reflejos presentes en la escena:** La interacción entre las fuentes luminosas y los diferentes objetos de la escena pueden producir efectos que suelen ser categorizados como objetos en movimiento (sombras y reflejos). Las sombras que acompañan a los objetos en movimiento no son parte del primer plano pero tampoco forman parte del fondo: deben ser eliminadas. Sin embargo, las sombras inherentes al fondo deben analizarse como parte de él. Existen algoritmos que eliminan las sombras aprovechando la información de brillo y cromaticidad en las imágenes. Por ejemplo, en [24] Horprasert presenta un modelo en el que determina que un píxel es sombra si la cromaticidad permanece constante y el brillo varía. En [13] Elgammal, Harwood y Davis elaboran un modelo similar pero de mayor sencillez.
- **Obtener y actualizar el fondo de la escena:** A menudo, es difícil extraer una imagen de fondo adecuada durante la fase de inicialización de dicho fondo. Este hecho puede ser debido a que, en ese momento, existan objetos en movimiento en la escena que no puedan ser discriminados como fondo. Algunos autores [25] analizan dicho problema de inicialización. Por otro lado, es necesario detectar los cambios susceptibles de producirse en el fondo de la escena, ya que, según la complejidad del escenario, dicho fondo podría verse alterado por multitud de objetos; por ejemplo, en un aparcamiento, por los coches que van estacionando o saliendo; o, en una estación de autobuses, por objetos robados o abandonados.
- **Determinar los parámetros de funcionamiento de los algoritmos:** Una de las tareas más importantes de este tipo de sistemas automáticos de análisis es el ajuste de los parámetros

de los algoritmos que determinan el fondo y los objetos en las secuencias. En cada algoritmo, podemos encontrar diferentes parámetros (umbrales, pesos,..) con distintas sensibilidades. Los parámetros más significativos serán los que alteren en mayor medida los resultados del sistema y por ello, tendremos que estudiarlos a fondo en función de las características de las secuencias.

- **Ruido introducido en la secuencia de imágenes:** El ruido que proviene de la captación de las imágenes por las cámaras de video puede producir errores en la segmentación de objetos de video y es necesario eliminarlo.
- **Fondos multimodales:** Son los que podemos encontrar en secuencias con objetos en movimiento lento y/o periódico (movimiento de las hojas de los árboles, movimiento ondulatorio del agua,...) y que, desde una perspectiva semántica, habitualmente se considera que pertenecen al fondo de la escena.
- **Camuflaje:** Este efecto aparece cuando los objetos del primer plano poseen el mismo color y textura que el fondo; por este motivo, el frente se confunde o *camufla* como fondo.

A fin de solucionar estos inconvenientes y extraer las regiones de interés con máxima fiabilidad, existen diversas aproximaciones. La mayoría de ellas, siguen un diagrama de flujo definido por Cheung y Kamath [26], que identifica cuatro etapas:

- 1) **Pre-procesado**: En esta etapa se realizan simples tareas de procesamiento de imágenes que cambian la señal de entrada de video en un formato que puede ser procesado por las etapas siguientes.
- 2) **Modelado de fondo**: En esta etapa se realizan tanto la inicialización del fondo como su mantenimiento.
- 3) **Detección de frente**: En esta etapa se realiza la substracción de frente; que permite obtener los objetos de interés en movimiento en la secuencia de video.
- 4) **Validación de datos o post-procesado**: En esta etapa se realizan diversas operaciones para mejorar los resultados obtenidos de las etapas anteriores.

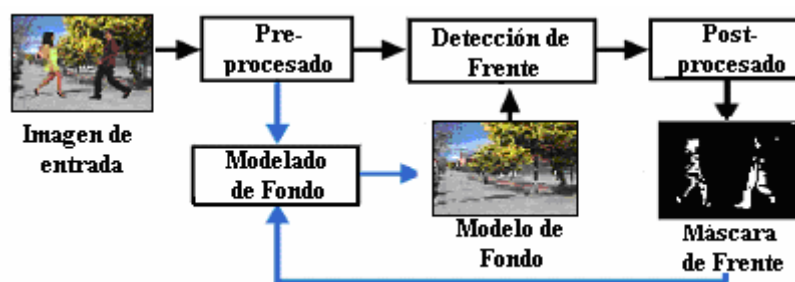


Figura 2-1: Diagrama de bloques de un algoritmo de BGS.

Aunque las etapas de modelado de fondo y detección de frente suelen mezclarse y confundirse, lo cierto es que se trata de etapas diferentes: el modelado de fondo comprende la elaboración y actualización de un fondo a partir de la secuencia de video generada por una cámara y, la detección de frente es el proceso mediante el cual cada imagen de un video se compara con el modelo de fondo a fin de determinar los píxeles pertenecientes al fondo y aquellos que son primer plano o frente.

En los capítulos siguientes abordaremos el problema de la detección de los objetos en movimiento mediante técnicas de extracción de fondo. Seguiremos el diagrama de flujo genérico presentado por Cheung y Katmath [26] analizando en primer lugar las técnicas de pre-procesado más utilizadas en diferentes situaciones, los algoritmos de modelado de fondo y extracción de frente así como las formas de actualizar e inicializar los modelos y finalmente, presentaremos diversas medidas para validar el rendimiento de los algoritmos.

2.2 Técnicas de pre-procesado en segmentación de video

El objetivo de un algoritmo de detección de objetos en movimiento es reconocer los cambios importantes que se producen en una secuencia de video y rechazar los de menor relevancia.

La función de la etapa de pre-procesamiento es filtrar los cambios menos significativos que se producen en la escena y/o mejorar las propiedades de las imágenes de la secuencia de video a analizar para simplificar la labor de detección de los objetos en movimiento más representativos que llevan a cabo las etapas posteriores.

Frecuentemente, en la etapa de pre-procesamiento se utilizan ajustes geométricos o ajustes de intensidad [27]. Existen otro tipo de técnicas que calculan la derivada de la imagen para extraer

información de los bordes de los objetos y así aprovechar esa información para facilitar la detección de los objetos en etapas posteriores.

En sistemas de tiempo real, se suele reducir el tamaño o la tasa de las imágenes para disminuir el tiempo de ejecución. También se utilizan filtros espaciales o temporales que eliminan ruido o variaciones del fondo que no se desean detectar como objetos en movimiento (por ejemplo, la lluvia o la nieve capturadas en aplicaciones al aire libre).

No obstante, las técnicas de pre-procesado no sólo permiten eliminar información poco relevante o capturar más información que facilite la detección; para obtener buenos resultados de detección en secuencias de fondo complejas con mucho movimiento en el fondo, un sistema con algoritmos de modelado de fondo y sustracción del frente sencillos algorítmicamente, por lo general, requiere técnicas de pre-procesado robustas. De ahí proviene la motivación de muchos investigadores a desarrollar técnicas más complejas de modelado de fondo que suelen conllevar un aumento de la carga computacional pero que también permiten ahorrar las técnicas de pre-procesado y post-procesado.

A continuación se describen algunas de las técnicas de pre-procesado más utilizadas:

- **Ajustes geométricos:** Las técnicas de ajustes geométricos permiten regenerar los valores de los píxeles afectados a causa de cambios de intensidad (que pueden ser debidos a cambios ambientales o cambios de la fuente de iluminación) o movimientos de cámara, por los valores de los píxeles adyacentes.

También estas técnicas permiten regenerar las máscaras de frente. Para ello, se agrandan o disminuyen regiones. De esta forma, se consigue incrementar la precisión en la detección del objeto de interés.

Entre las técnicas de ajustes geométricos se encuentran la erosión y dilatación, rotación, las transformaciones afines o proyectivas [27]-[31].

- **Ajustes radiométricos o de intensidad:** Existe una gran variedad de técnicas que sirven para compensar las variaciones de iluminación que se producen por cambios de intensidad o posición de las fuentes de iluminación.

Para hacer frente a los cambios de intensidad de la iluminación, algunos algoritmos [32][33][34] normalizan los valores de intensidad de los píxeles en una imagen para que tengan la misma media y varianza que los píxeles de la imagen de fondo y calculan umbrales de decisión independientes a los valores originales de intensidad.

Una forma sencilla de adaptarse a los cambios de iluminación puede realizarse con una actualización adecuada del fondo de la secuencia. Sin embargo, esta técnica tiene como inconveniente la estabilización del fondo en un tiempo elevado.

Aunque la mayoría de los algoritmos manejan la información de los píxeles con un único valor (o canal) por cada píxel [39]-[45], cada vez es más usual encontrar algoritmos de segmentación de objetos en movimiento que emplean varios valores (canales o colores) para procesar los píxeles de una imagen. Con varios canales por cada píxel se pueden diseñar etapas de pre-procesado que ajusten la iluminación en base a estos canales utilizando técnicas de contraste y eliminación de sombras[46]-[53].

- **Derivada de la imagen:** Existen técnicas de pre-procesado que calculan la derivada de la imagen (diferencia de frentes) para generar una representación de la escena de fondo invariante a los cambios de iluminación o para detectar objetos de alta velocidad. Por ejemplo, Pless en [54] combina diferencias espaciales y temporales con un modelado de fondo sencillo para analizar secuencias de vehículos en movimiento a gran velocidad.
- **Profundidad de la escena:** Los métodos de sustracción de fondo basados en técnicas de profundidad presentan algunos problemas [56][57] para llevar a cabo la segmentación de objetos en movimiento cuando los objetos y el fondo se encuentran muy próximos (por ejemplo, en secuencias con manos apoyadas en la pared o pies sobre el suelo). No obstante, estas técnicas se ven poco afectadas por los cambios de iluminación, por este motivo, el cálculo de profundidades en la escena puede combinarse con modelos de sustracción de fondo basados en modelado de fondo sencillos como realiza Gordon en [58] para llevar a cabo la segmentación de objetos en movimiento.

2.3 Modelado de fondo

El modelado de fondo, también denominado actualización de fondo, es la etapa fundamental en los algoritmos de segmentación de objetos en movimiento. Su función es la inicialización, actualización y representación de un modelo de fondo robusto de la secuencia de video analizada. El fondo se describe mediante un modelo matemático, para cada píxel de la imagen en cada instante de tiempo. El objetivo de esta etapa no es encontrar la información relativa a los objetos en movimiento, ya que esta tarea pertenece a una etapa posterior llamada extracción de frente.

Una primera aproximación para modelar el fondo de una escena sería suponer que los píxeles de fondo no cambian de valor a lo largo de la secuencia, es decir, sólo los objetos en movimiento se modificarán en los correspondientes píxeles. Esta suposición se suele realizar en entornos muy controlados (por ejemplo, interiores con una imagen de fondo conocida y estática y sin variaciones de iluminación) y requiere únicamente modelar factores externos como el ruido de la cámara. Este tipo de fondos se conoce como fondos unimodales. En cambio, en entornos menos controlados, los píxeles del fondo pueden cambiar su valor debido a cambios de iluminación (como sucede, por ejemplo, en escenas capturadas al aire libre y a distintas horas del día) o debido al movimiento de los objetos que pertenecen al fondo (como ocurre en escenas con árboles agitándose u olas de mar). Este tipo de fondos se conoce con el nombre de multimodales.

Los métodos de modelado de fondo unimodal, asumen que la distribución de cada uno de los píxeles de fondo tiene un único pico o modo y por tanto, el fondo de la secuencia analizada es esencialmente estático (como se puede observar en la primera gráfica de la **Figura 2-2**). Sin embargo, las secuencias con fondo no estático, en las que los píxeles de fondo experimentan variaciones notables, deben modelarse con distribuciones multimodales (que presentan varios picos o modos, como es el ejemplo de la segunda gráfica de la **Figura 2-2**). Por lo tanto, el conocimiento y modelado de los diferentes tipos o modos de variación es decisivo a la hora de analizar una secuencia y modelar su fondo.

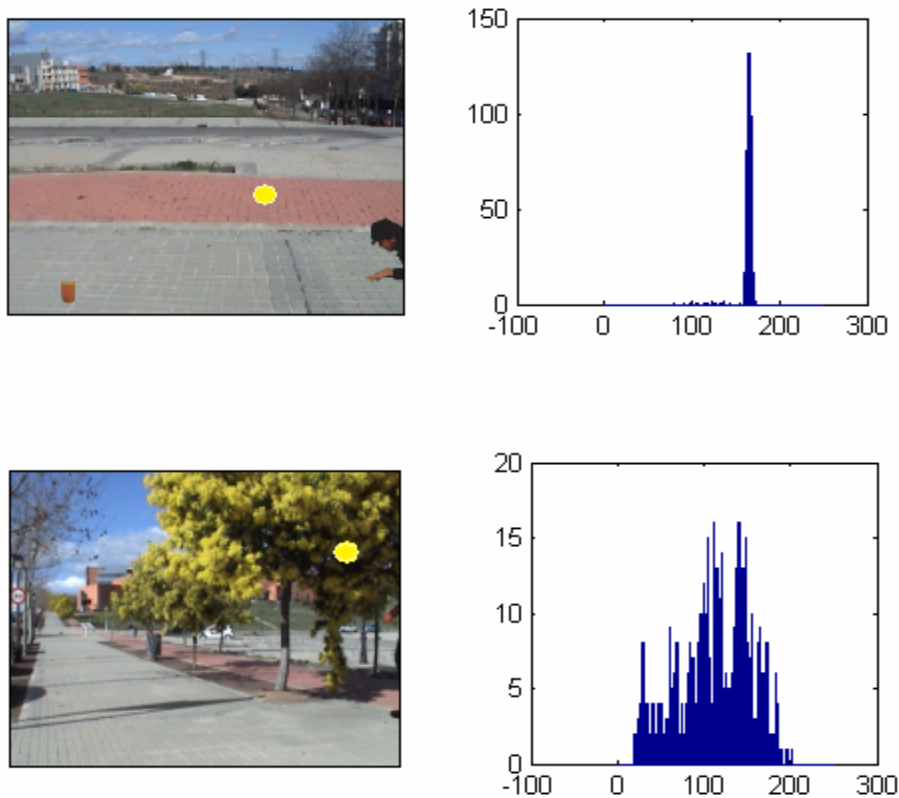


Figura 2-2: Secuencia unimodal (superior) y multimodal (inferior) y los histogramas del píxel situado en la posición (80, 313) marcado con un círculo amarillo en las imágenes a la derecha.

El problema de adaptación también es una cuestión importante que considerar a la hora de representar correctamente el fondo. Se debe tener en cuenta la velocidad de los cambios que sucedan en la escena (lo cual se suele modelar con un parámetro denominado factor de adaptación). Además, el fondo se debe adaptar tanto a las posibles variaciones de iluminación (por ejemplo, si la secuencia es capturada a distintas horas del día o existen cambios en la fuente que ilumina la escena), los posibles movimientos de objetos pertenecientes al fondo (por ejemplo, hojas o ramas agitándose o movimientos del agua o del mar) como a los objetos que se depositen o desaparezcan del fondo (debido al robo o abandono) ya que todas estas modificaciones en el fondo pueden convertirlo en otro completamente diferente al inicial.

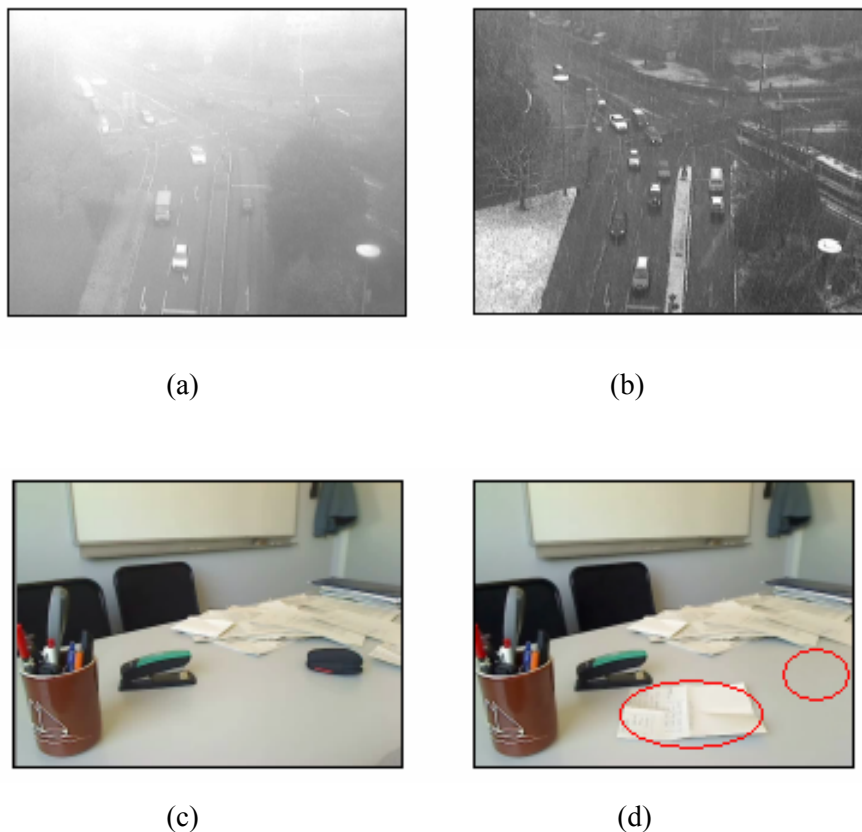


Figura 2-3: Secuencia con distinta iluminación debido a: (a) niebla y (b) lluvia. Secuencia (d) con un objeto aparecido (folio) y desaparecido (ratón) con respecto a la (c).

También es necesario estudiar la forma de inicializar la secuencia, ya que si se comienza la sustracción del fondo y extracción de objetos sin una imagen de fondo adecuada a la secuencia lo más posible es que no se detecten correctamente los objetos en movimiento y se actualice un fondo erróneo. La mayoría de los algoritmos de segmentación de fondo basados en modelado de fondo, inicializan la imagen de fondo a la primera imagen de la secuencia y, a partir de ella, actualizan y

modelan el fondo en base a las imágenes posteriores. Este mecanismo de inicialización, si se ejecuta sobre secuencias en las que la primera imagen tiene objetos moviéndose, acarrea multitud de problemas, puesto que la imagen de fondo que modelaría la secuencia tomaría objetos no deseables como parte del fondo impidiendo, por tanto, la detección correcta de los objetos en movimiento.

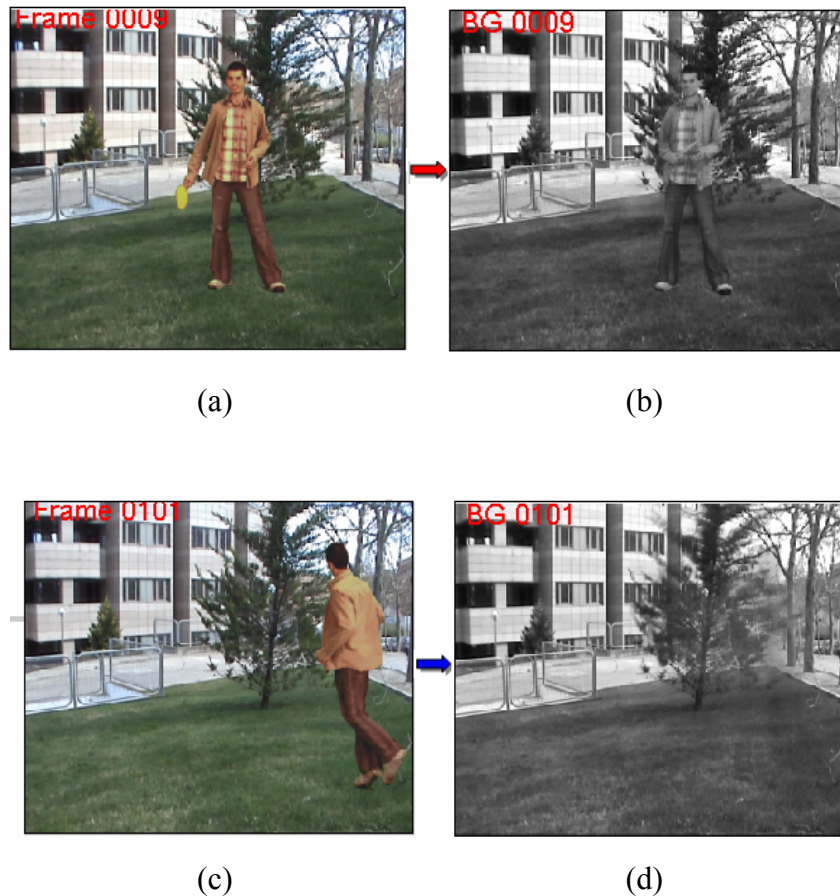


Figura 2-4: Secuencia (a) y fondo (b) inicializado erróneamente a la primera imagen capturando en él objetos en movimiento. Secuencia (c) y fondo (d) actualizado correctamente.

Según Cristani [59], las características principales para modelar un fondo son tres:

- 1. Modelo de representación:** describe el tipo de modelo matemático utilizado para representar el fondo.
- 2. Modelo de inicialización:** se refiere a la forma con la que se obtiene el primer fondo o estado del modelo de fondo del algoritmo.

3. Modelo de adaptación: mecanismo para adaptar el modelo a cambios en el fondo de la escena (por ejemplo, cambios de iluminación).

A continuación, se analizarán las diversas propuestas existentes en base a las tres características previamente descritas.

2.3.1 Modelo de representación del fondo

En secuencias para las que ya se dispone de un fondo conocido, como sucede en estudios de televisión en los que se diseña un fondo específico, resulta muy sencillo representar el fondo. En tales circunstancias, y si la cámara es fija, el diseño de los algoritmos de modelado de fondo se simplifica; los píxeles de fondo son valores constantes y, por tanto, el fondo se puede modelar con una imagen estática. Para representar esta imagen de fondo, en el caso de imágenes en escala de grises se puede utilizar un valor de intensidad para cada píxel. En el caso de una imagen a color, cada píxel sería representado con un canal R, G y B u otra escala (HSV,...).

No obstante, en la mayoría de las aplicaciones no se dispone del fondo o el fondo no es constante. Incluso, en el ejemplo anterior, factores externos como el ruido de los sistemas de captación o internos como las variaciones de iluminación debidas a la interacción entre objetos móviles y fuentes de luz, hacen que en la práctica el fondo no se modele como constante. Por este motivo, algunos autores elaboran modelos matemáticos que permiten representar el fondo y adaptarlo a los cambios existentes.

En este proyecto, los métodos de representación del modelo de fondo se clasifican en:

1.- Modelos básicos: modelos matemáticos sencillos para generar el fondo (diferencias, sumas, medias, medianas).

2.- Modelos paramétricos: el modelo que define el fondo se caracteriza por un conjunto de parámetros que describen la distribución de cada píxel. Aunque los parámetros de la distribución son variables, el tipo de distribución se asume como parte del modelo.

3.- Modelos no paramétricos: la representación del fondo se realiza por medio de funciones de densidad de probabilidad u otras funciones matemáticas, sin asumir distribuciones concretas controlables por parámetros.

A continuación se describen los distintos modelos:

2.3.1.1 Modelos básicos

Los métodos básicos de sustracción de fondo utilizan modelos matemáticos sencillos tales como, diferencias entre imágenes, promedios, máximos y mínimos,..etc., que permiten modelar simplificadaamente los píxeles de la imagen de fondo.

A continuación, se describen algunos de estos métodos:

- ***Diferencia de imágenes ('Frame differencing')***

El método de diferencia de imágenes [60], también llamado diferencia temporal, es posiblemente la forma más sencilla de sustracción de fondo.

Utiliza como modelo de fondo (B_t) para la imagen de la secuencia en el instante (I_t), la imagen anterior, es decir, la imagen del instante de tiempo $t - 1$.

$$B_t = I_{t-1} \quad (2.1)$$

Como ventajas, esta técnica posee una baja carga computacional, y el frente se adapta bien a los cambios rápidos.

Esta técnica es sensible al ruido y a las variaciones de iluminación. Además, debido a que sólo utiliza una única imagen anterior como modelo de fondo, la diferencia con el cuadro anterior no es capaz de identificar el interior de los objetos en movimiento, puesto que entre imágenes consecutivas, esos píxeles se mantienen invariantes.

- ***Filtro promedio temporal ('Average filter')***

En el filtro de promedio, el modelo de fondo (B_t) se calcula como una imagen estática hasta que se producen movimientos, en cuyo caso, el B_t corresponde al promedio de un conjunto de imágenes consecutivas en el tiempo.

$$\begin{aligned} |I_t - B_t| > \tau &\Rightarrow B_t = \frac{1}{N} \sum_{n=t}^{n=t-N} I_n \\ |I_t - B_t| \leq \tau &\Rightarrow B_t = I_t \end{aligned} \quad (2.2)$$

Este modelo no es robusto en secuencias con muchos objetos en movimiento sobre todo si se mueven lentamente. Además, no puede manejar fondos multimodales, adapta lentamente las variaciones del fondo y posee un único umbral para toda la imagen.

Koller en el artículo [61], utiliza un filtro basado en una ventana de promedio temporal que permite detectar los objetos en movimiento cuando se producen variaciones de iluminación.

Uno de los problemas que presenta esta técnica es que la información de movimiento se utiliza para actualizar el fondo. Si los objetos se mueven lentamente, no serán detectados. Por este motivo, una posible solución sería utilizar también la información de los píxeles no considerados objetos para actualizar el fondo.

- ***Filtro de mediana aproximado ('Approximated median filter')***

En este método, se almacena en un buffer los valores de los píxeles en un conjunto de imágenes. El fondo se define en cada píxel como la mediana de los valores almacenados en el buffer.

Es un método robusto cuyos resultados se asemejan a los de los métodos complejos. Se utiliza en los modelos de fondo elaborados en sistemas de vigilancia de tráfico urbano, donde la estimación de la mediana se incrementa por uno si el píxel de entrada es mayor que la estimación y disminuye en caso contrario.

$$\begin{aligned} I_t > B_t &\Rightarrow B_t = B_{t-1} + 1 \\ I_t < B_t &\Rightarrow B_t = B_{t-1} - 1 \\ I_t = B_t &\Rightarrow B_t = B_{t-1} \end{aligned} \tag{2.3}$$

Esta estimación converge a un valor tal que la mitad de los píxeles de entrada son mayores a él y la otra mitad son inferiores, es decir, es un estimador de la mediana.

El único inconveniente de este método, como observaron Cheung y Kamath [26], es que se adapta lentamente cuando se produce un cambio en el fondo. No obstante, existen técnicas para considerar la región del fondo descubierta por un objeto que se mueve después de haber permanecido estático durante un largo periodo de tiempo.

2.3.1.2 Modelos paramétricos

Los algoritmos basados en modelos paramétricos definen modelos de fondo más complejos, que permiten cierta tolerancia al ruido y a pequeñas fluctuaciones (hojas en movimiento, parpadeo de

luces, pequeñas vibraciones de las cámaras, cambios bruscos de luz,..). Describen la imagen de fondo en base a parámetros de una distribución de probabilidad estándar (usualmente Gaussiana). Algunos de estos ejemplos son los que se muestran a continuación:

- *Gaussiana simple ('Simple Gaussian')*

El método de la Gaussiana Simple (SG) modela los pequeños cambios que ocurren en la imagen de fondo B_t representando cada píxel con una distribución unimodal Gaussiana definida por dos parámetros: media μ_t y varianza σ_t^2 .

$$\begin{aligned}\mu_t(x, y) &= \sum_{i=1}^{i=t} \frac{B_i(x, y)}{t} \\ \sigma_t^2(x, y) &= \sum_{i=1}^{i=t} \frac{B_i^2(x, y)}{t} - \mu_t^2(x, y)\end{aligned}\tag{2.4}$$

En cada instante t , se determina si un píxel pertenece al fondo $B_t(x, y)$ si el valor de dicho píxel en la imagen recae dentro de la Gaussiana definida para ese píxel, es decir, si la diferencia entre el valor del píxel y el de la media considera para dicho píxel $\mu_t(x, y)$ es inferior a la desviación $\sigma_t^2(x, y)$.

Una gran cantidad de algoritmos utilizan como base de su segmentador de objetos en movimiento el principio de la Gaussiana simple. Es el caso de Gordon [65], que genera para cada canal (R, G y B) de un píxel una distribución Gaussiana.

El filtro de Kalman es una técnica recursiva de seguimiento lineal en la que los valores de los píxeles pueden modelarse por una distribución Gaussiana $N(\mu, \sigma^2)$. La media y la varianza del fondo se actualizan utilizando filtros adaptativos para acomodar los cambios de iluminación.

El método de la SG no es capaz de adaptarse a fondos multimodales, en los que cada píxel de fondo puede tomar valores muy diferentes, sin por ello, dejar de ser un píxel de fondo.

- *Mezcla de Gaussianas ('Mixture of Gaussian')*

En fondos multimodales, que contienen objetos no estáticos, tales como hojas de árboles en movimiento, olas, etc., hay píxeles cuyos valores de intensidad varían entorno a un conjunto finito de valores característicos. Por este motivo, un píxel no puede modelarse por medio de un valor (una media) y un conjunto en torno a éste (la varianza) utilizando una distribución Gaussiana. La Mezcla de Gaussianas (MoG) propone una solución a este problema [67] que consiste en modelar la intensidad de los píxeles con una mezcla de k distribuciones Gaussianas (donde k es un número pequeño, frecuentemente se utiliza de 3 a 5) definidas por los siguientes parámetros: media $\mu_{k,t}$, varianza $\sigma_{k,t}^2$, y peso $w_{k,t}$.

$$\begin{aligned}\mu_{k,t}(x, y) &= \sum_{i=1}^{i=t} \frac{B_i(x, y)}{t} \\ \sigma_{k,t}^2(x, y) &= \sum_{i=1}^{i=t} \frac{B_i^2(x, y)}{t} - \mu_i^2(x, y)\end{aligned}\tag{2.5}$$

Los píxeles de una imagen nueva se comparan con las k posibles subdistribuciones de fondo; si se encuentra parecido, es decir, el píxel no difiere de alguna de las medias en más de la varianza asociada, se caracteriza el píxel como fondo y se actualizan los parámetros del modelo para ese píxel: la media $\mu_{k,t}(x, y)$, la varianza $\sigma_{k,t}^2(x, y)$ y un peso $w_{k,t}(x, y)$ cuyo valor determina la cantidad de distribución que modela dicho píxel en cada instante. Si no se parece a ninguna de las distribuciones asociadas al píxel, la de menor peso se sustituye por una nueva Gaussiana de media el valor del píxel en la imagen actual y varianza un valor muy pequeño.

En el modelo de mezcla de Gaussianas, la combinación de k distribuciones Gaussianas cuya probabilidad de ocurrencia (suma de pesos $w_{k,t}(x, y)$) supere un determinado umbral, denominado umbral de frente, permitirá modelar cada píxel del fondo en cada instante.

Al igual que sucede en el modelo de la SG, la MoG contiene multitud de variantes; existen otras medidas para clasificar los píxeles de fondo y realizar el seguimiento de los objetos en movimiento. Por ejemplo, Javed [68] utiliza la MoG para llevar a cabo la substracción de fondo en cada canal de color, relacionando dichos canales a través de una matriz de covarianza. Harville [22], propone un método para modelar el fondo que combina la MoG con la información de profundidad y luminancia; este método se utiliza en detección de automóviles.

No obstante, la MoG también posee inconvenientes. En primer lugar, conlleva una alta carga computacional. Por otro lado, es muy poco robusta a cambios repentinos de iluminación. Además, los fondos multimodales requieren un número de distribuciones k elevado para modelar cada píxel, lo cual implica un incremento en la carga computacional. En este método de representación del fondo es muy importante la forma de actualizar las medias y las varianzas para adaptarse a los cambios del fondo.

A pesar de estos inconvenientes, la MoG es capaz de manejar una distribución multimodal de fondo ya que mantiene una función de densidad de probabilidad para cada píxel. Al ser un método paramétrico, puede adaptarse a los cambios de fondo sin necesidad de actualizar un gran buffer de almacenamiento de imágenes como, según se verá, requieren los métodos no paramétricos.

2.3.1.3 Modelos no paramétricos

Los modelos no paramétricos también son métodos complejos en los que no se asumen distribuciones estándar de probabilidad para modelar a los píxeles de fondo, sino técnicas más generales, como pueden ser almacenamiento de los últimos valores del píxel, cálculo de rangos de valores del píxel, ajustes de funciones de predicción, etc. Algunos modelos no paramétricos son los siguientes:

- ***Densidad de Núcleo ('Kernel Density Estimate')***

El método de representación del fondo Densidad de Núcleo ('KDE') descrito por Elgammal [21] estima la función de densidad de probabilidad de cada píxel de la imagen de fondo en cada instante de tiempo. Esta operación se realiza gracias a la información de la historia reciente de dicho píxel que se haya almacenada en un buffer. El objetivo es obtener mayor sensibilidad de detección que utilizando un método de representación de fondo con una distribución de probabilidad fija.

$$B_t(x, y) = \text{buffer}\{I_{t-L}(x, y), I_{t-L+1}(x, y), I_t(x, y)\} \quad (2.6)$$

La pertenencia al fondo B_t se estima mediante el promedio ($\text{Pr}(x_t)$) de funciones de núcleo K (por ejemplo, de tipo Gaussiano) centradas en valores anteriores x_i y evaluados en el valor del píxel actual x_t ; es decir, se calcula la probabilidad de parecido entre el píxel actual y los valores que dicho píxel ha tomado en las L imágenes anteriores y si esa probabilidad es suficientemente elevada (supera un determinado umbral U) se considera que el píxel es fondo; en caso contrario, será frente.

$$x_t \in B_t / \Pr(x_t) = \sum_{i=t-1}^{i=t-1-L} K(x_i - x_t) > U \quad (2.7)$$

Por ejemplo, eligiendo un estimador de función de núcleo K de tipo Gaussiano, cada muestra de las L almacenadas se considera que posee una distribución $N(\mu_{k,t}(x, y), \sigma_{k,t}^2(x, y))$.

$$\Pr(x_t) = \frac{1}{L} \sum_{i=t-1}^{i=t-1-L} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_t-x_t)^2}{2\sigma^2}} \quad (2.8)$$

El modelo KDE soporta parpadeo de fondo, ruido en la imagen y es capaz de adaptarse a los cambios rápidos y progresivos del fondo. No obstante, posee un requisito muy importante que es la alta carga computacional.

- **Modelos ocultos de Markov ('HMM')**

Los modelos estudiados en los apartados anteriores pueden adaptarse a cambios graduales en la iluminación, no obstante, un cambio súbito puede representar un gran problema.

Otro enfoque para modelar una amplia gama de variaciones en la intensidad de los píxeles sería utilizando los modelos ocultos de Markov (HMM) que representan estas variaciones como un conjunto de estados discretos correspondientes a los distintos modos de iluminación que pueden presentarse en la escena: luces de encendido, apagado, nublado, soleado... Por ejemplo, en seguimiento de tráfico [69] la intensidad de un píxel se puede modelar con tres estados de distribución Gaussiana: fondo, frente y sombra.

El modelo de fondo B_t quedaría definido por un conjunto de estados b_j caracterizados por una función de densidad de probabilidad para cada píxel localizado en la posición (x, y) :

$$B_t(x, y) = b_j(x, y) \quad (2.9)$$

Los modelos de representación del fondo basado en HMMs requieren un tiempo de cálculo elevado ya que suponen la valuación de un conjunto de estados, y la topología del sistema puede ser muy compleja según las características de la escena.

- **Métodos basados en códigos ('Codebooks')**

Cada píxel de fondo $B_t(x, y)$ puede codificarse por un conjunto de valores 'codewords' (c_m) que constituirían un descriptor o 'Codebook' (ζ).

Al procesar una nueva imagen de la secuencia, los valores de intensidad de los píxeles se comparan con los que forman su 'Codebook' mediante diferencias de color y luminancia, de tal forma que si existe parecido, se estiman y actualizan los códigos y, si no hay coincidencia, se inserta el nuevo código.

$$B_t = \{c / c_m \in \zeta \wedge \lambda_m \leq T_m\} \quad (2.10)$$

El hecho de mantener un conjunto de códigos para cada píxel que actualizar en cada instante hace que el método del 'Codebook' sea computacionalmente costoso. No obstante, se han desarrollado técnicas que resuelvan este problema; por ejemplo, Butler [71] clasifica los píxeles dotándoles de un peso que indica la cantidad de píxeles del mismo valor; la suma de los pesos es una estimación de la probabilidad de píxeles de fondo y puede utilizarse como umbral de decisión para determinar los píxeles pertenecientes al fondo y al frente.

2.3.2 Modelo de inicialización del fondo

Actualmente existe un gran interés por abordar las cuestiones de representación y modelado de fondo [25],[72]-[74]. Sin embargo, un tercer problema que ha recibido poca atención es la inicialización de fondo.

La mayoría de los modelos de fondo se basan en la generación inicial de un fondo mediante un conjunto de parámetros obtenidos a través de un número pequeño de imágenes de la secuencia en las que no están presentes los objetos en movimiento [75]. En algunas situaciones, esta suposición es difícil o prácticamente imposible de controlar, por ejemplo, en zonas de alta población caracterizadas por una presencia continua de objetos en movimiento u otros efectos perturbadores. En esos casos, parece necesario capacitar al modelo de un fondo previo robusto, sin objetos.

Existen varios métodos para generar un modelo inicial de fondo; por ejemplo, en sistemas de monitorización de tráfico [76] la imagen de fondo se inicializa por medio de la mediana de un conjunto de imágenes, basándose en el supuesto de que en el fondo cada píxel será visible en más del cincuenta por ciento del tiempo de la secuencia. También se puede utilizar la mediana de un

conjunto de imágenes para inicializar la imagen de fondo. La ventaja de usar la mediana frente a la media es que aquella evita la mezcla de valores de un píxel ya que la media del valor de intensidad de un píxel puede no corresponder al valor de cualquiera de los píxeles durante ese tiempo, en cuyo caso se produciría un error.

2.3.3 Modelo de adaptación del fondo

La forma de actualizar el fondo permite adaptar el modelo de fondo a los cambios que puedan producirse en los escenarios de las secuencias de video analizadas.

Esta operación de adaptación no es necesaria en aquellas secuencias en las que se conozca el modelo de fondo y se sepa que no tendrá variaciones en el tiempo (por ejemplo, en estudios de televisión donde se dispone de un fondo estático previamente diseñado).

Sin embargo, no siempre el modelo de fondo de la secuencia es conocido y se tiene la seguridad de que va a permanecer sin variaciones. Existen multitud de circunstancias por las que el fondo de una secuencia de video puede modificarse, por ejemplo, por los cambios de iluminación de la fuente o la captura de secuencias al aire libre tomadas a distintas horas del día, o por el hecho de que desaparezcan objetos pertenecientes del fondo (por ejemplo, objetos robados) u objetos que se tengan que incluir en el modelo fondo (por ejemplo, objetos abandonados). Por este motivo, la operación de adaptación es, en la mayoría de las secuencias de video, un requisito imprescindible para mantener un modelo de fondo adecuado.

Existen múltiples técnicas para adaptar el modelo de fondo: por ejemplo, una manera sencilla sería actualizar cada instante el modelo de fondo a través de la media o la mediana de los valores de las imágenes anteriores a dicho instante. Otras técnicas de mayor coste computacional son las que mantienen un buffer para cada píxel con los valores que dicho píxel ha tomado en instantes anteriores y actualizan el modelo de fondo en función de los valores del buffer (por ejemplo, construyendo funciones de probabilidad con los valores del buffer). Un método de actualización muy conocido en el estado del arte es la media móvil (*'Running average'*). Mediante esta operación, el modelo de fondo se adapta a las variaciones progresivas que puedan producirse en la escena, como por ejemplo, variaciones en la iluminación del día, ya que la actualización del fondo se realiza con parte de la imagen actual:

$$B_t(x, y) = \alpha I_t(x, y) + (1 - \alpha) B_{t-1}(x, y) \quad (2.11)$$

La velocidad de adaptación viene modelada por el parámetro α .

Este tipo de adaptación del modelo de fondo presenta como desventajas la detección como objetos en movimiento de sombras y reflejos y, que los objetos que desaparezcan del fondo se marcan como objetos móviles (fantasmas).

Todas las técnicas de actualización del modelo de fondo a su vez pueden realizarse mediante dos mecanismos diferentes: mecanismo de adaptación selectiva o a ciegas.

- **Mecanismo de adaptación selectiva:** Este mecanismo selecciona y actualiza (a través de cualquiera de las técnicas de actualización comentadas anteriormente) sólo los píxeles clasificados como fondo permitiendo incorporar los objetos que aparezcan en la escena y que pasado un determinado tiempo corresponden al modelo de fondo.



Figura 2-5: Ejemplo de la actualización selectiva sobre una secuencia de video en el instante t.

- **Mecanismo de adaptación a ciegas:** Este mecanismo actualiza (mediante uno de los métodos de actualización comentados anteriormente) todos los píxeles del modelo de fondo sin tener en cuenta ningún tipo de decisión previa. Por este motivo, estos mecanismos requieren técnicas de actualización de menor complejidad que los mecanismos selectivos, pero presentan como desventaja la actualización del fondo con los objetos en movimiento indeseados.



Figura 2-6: Ejemplo de la actualización a ciegas sobre una secuencia de video en el instante t.

2.4 Detección de frente

La detección de primer plano también denominada detección de frente, compara la entrada con el modelo de fondo e identifica los píxeles candidatos a ser objeto en movimiento (frente o primer plano).

Para obtener esta clasificación, la comparación entre la imagen y el fondo se puede llevar a cabo por diferencia y la clasificación fijando un umbral que depende de la escena, el ruido de la cámara, y las condiciones de iluminación.

En las siguientes secciones se va a examinar primero cómo generar la diferencia dado una modelo de fondo y la imagen actual y, a continuación, se describen algunas técnicas para hallar el umbral con el que clasificar los píxeles de frente.

2.4.1 Técnicas de detección de frente

Las técnicas de detección de frente son algoritmos sencillos cuando se dispone de un modelo de fondo adecuado a la escena de análisis ya que por comparación entre la imagen actual y fondo se pueden detectar fácilmente los objetos en movimiento. Algunas de estas técnicas son las siguientes:

2.4.1.1 Basadas en diferencia

El método más básico para realizar la detección de frente F_t es calculando la diferencia entre dos imágenes: los píxeles cuyos valores de diferencia son altos normalmente corresponderán a objetos en movimiento.

$$\begin{aligned}(I_l(x, y) - B_l(x, y)) \leq \tau &\Rightarrow F_l(x, y) = 0 \\(I_l(x, y) - B_l(x, y)) > \tau &\Rightarrow F_l(x, y) = 1\end{aligned}\tag{2.12}$$

La diferencia entre dos imágenes se puede calcular mediante diversos métodos: el valor absoluto de la diferencia o la diferencia normalizada, a los que posteriormente aplicar un umbral con el que discernir los píxeles que forman parte del fondo (ausencia de movimiento) y aquéllos que pertenecen al frente (objeto en movimiento).

En la literatura podemos encontrar diferentes formas de diferencia:

- **Diferencia absoluta:** La diferencia absoluta entre la imagen actual y la imagen de fondo proporciona bordes gruesos que son útiles en detección de frente ya que permiten aligerar el post-procesado para rellenar contornos [52].
- **Diferencia relativa:** Consiste en aplicar un diferencia en función de la distribución espacial de localización. Por ejemplo, definiendo para las diferencias de regiones de bajo contraste un umbral menor [77].
- **Diferencia normalizada:** Es otro método de detección de frente que se realiza aplicando un umbral sobre las estadísticas normalizadas.
- **Diferencia basada en dos umbrales:** Algunos autores, aplican varios umbrales a la diferencia entre la imagen y el fondo para detectar los objetos en movimiento y de esta forma, incrementar el grado de precisión en la detección. Boulton en [62], mantiene un registro de valores mínimo y máximo de N imágenes y a la diferencia entre cada imagen y su modelo de fondo se le aplica dos umbrales uno alto TL y otro bajo TH. Los píxeles cuya diferencia sea inferior al umbral menor TL se considerarán fondo debido al gran parecido (poca diferencia) entre la imagen y el fondo y, los superiores al máximo TH se clasifican como frente. Por otro lado Jabri [78], realiza una diferencia entre los canales de color por separado de la imagen y su correspondiente modelo de fondo (caracterizado por una media y una desviación estándar) y, a continuación, combina los resultados.

2.4.1.2 Basadas en estadísticos

Las técnicas de detección de frente basadas en estadísticos pueden utilizarse si se conoce la función densidad de ruido de la cámara. Estos métodos, en lugar de umbralizar la diferencia de la imagen,

comparan el comportamiento estadístico de un conjunto de valores de un píxel pertenecientes a una serie de imágenes de la secuencia con el píxel de la imagen actual situado en la misma posición que éstos.

Los métodos estadísticos dependen de la distribución con la que se modele el fondo, de tal forma que cada modelo de fondo utiliza una estrategia diferente para detectar los píxeles de frente. Por ejemplo:

- **Basados en una Gaussiana:** Si se utiliza como modelo de fondo la SG, la imagen actual se compara con el modelo de fondo de distribución Gaussiana (μ_t, σ_t^2) mediante la medición de la probabilidad en luminancia o color y si la probabilidad es pequeña, el píxel se clasifica como frente ($F_t = 1$), de lo contrario, será clasificado como fondo ($F_t = 0$).

$$\begin{aligned} |I_t(x,y) - \mu_t(x,y)| \geq c\sigma_t(x,y) &\Rightarrow F_t(x,y) = 1 \\ |I_t(x,y) - \mu_t(x,y)| < c\sigma_t(x,y) &\Rightarrow F_t(x,y) = 0 \end{aligned} \quad (2.13)$$

- **Basados en Mezcla de Gaussianas:** En una nueva imagen procesada, el valor de cada píxel se compara con las k distribuciones que modelan el píxel en una MoG (μ_t, σ_t^2) a fin de decidir si el elemento pertenece al fondo ($F_t = 0$), recae dentro de la Gaussiana k-ésima) o al frente ($F_t = 1$), si la diferencia supera la varianza permitida.

$$\begin{aligned} \sum_{i=1}^{i=K} w_i \leq U, |I_t(x,y) - \mu_t(x,y,k)| > c\sigma_t(x,y,k) &\Rightarrow F_t(x,y) = 1 \\ \sum_{i=1}^{i=K} w_i > U, |I_t(x,y) - \mu_t(x,y,k)| \leq c\sigma_t(x,y,k) &\Rightarrow F_t(x,y) = 0 \end{aligned} \quad (2.14)$$

- **Basados en la estimación de densidad de Núcleo e Histogramas:** Se generan estimaciones de probabilidad para cada píxel en función de los píxeles almacenados de un número de imágenes anteriores a una procesada, de modo que, un píxel se considera primer plano ($F_t = 1$) si la probabilidad de pertenecer a esta estimación es inferior a un umbral.

$$\begin{aligned} f_{x,y}(imagen[x,y]) > \tau &\Rightarrow F_t(x,y) = 0 \\ f_{x,y}(imagen[x,y]) \leq \tau &\Rightarrow F_t(x,y) = 1 \end{aligned} \quad (2.15)$$

- **MRF:** Desarrollado por Migdal y Grimson [79], la detección de los conocimientos se logra mediante tres diferencias: M1, codifica las relaciones espaciales; M2, es una extensión de M1 y M3, explota las relaciones temporales.

2.4.2 Selección del umbral

La elección de un umbral adecuado es una operación decisiva para encontrar los objetos de interés en la secuencia. Como resultado del cálculo de la imagen diferencia, los niveles de gris pertenecientes a los píxeles de frente tendrán valores superiores a los píxeles en ausencia de movimiento, por lo tanto, el problema de localización de objeto en movimiento se resolverá buscando del umbral óptimo.

La salida de una operación de umbralización será una imagen binaria con valor 0 (negro) si pertenece al fondo y valor 1 (blanco) si se considera objeto en movimiento.



Figura 2-7: Máscara del frente

La eficiencia sobre el cálculo del frente depende de la selección del umbral, tal y como se observó en la sección 3). El umbral se puede establecer empíricamente [40], [80], [81], o a través de un cálculo adaptable [36], [39], [82]. En el primer caso, se fija un umbral para todos los píxeles que, generalmente se obtiene experimentalmente sobre una base de datos. En el segundo caso, el umbral se adapta siguiendo diversos criterios, como pueden ser: el número de falsas alarmas, características locales, características del modelo de fondo empleado o información espacial.

2.5 Técnicas de post-procesado

Los sistemas automáticos de análisis de video requieren una detección precisa de los objetos móviles existentes en la secuencia para lograr unos objetivos específicos, como pueden ser: conteo de personas, detección de objetos robados u abandonados, seguimiento (*'tracking'*) de objetos en movimiento o análisis de comportamientos, entre otros.

En la práctica, las diversas estrategias de sustracción de objetos en movimiento, en general, no proporcionan resultados todo lo eficientes que sería deseable y requieren la aplicación técnicas de post-procesado que reduzcan los posibles problemas que pueden afectar a la detección y permitan simplificar la labor de los algoritmos posteriores. Entre los problemas más importantes se encuentran: el ruido presente en los resultados obtenidos, la falta de iluminación uniforme, las sombras y reflejos de los objetos o la ocultación parcial de objetos. Todos estos inconvenientes se han estudiado por multitud de investigadores que han elaborado diversos algoritmos que permiten si no solucionarlos por completo, reducirlos en gran medida. Cucchiara en su trabajo [51] realiza un resumen de las diversas técnicas de sustracción de fondo y propone algunos métodos de post-procesado, por ejemplo: a fin de eliminar sombras, expone algoritmos basados en el análisis de color (RGB o espacio HSV); en el estudio de contornos, técnicas de crecimiento de regiones o cálculo de imagen diferencia; explica diferentes algoritmos de seguimiento y, por último, describe algunos métodos para reducir las oclusiones.

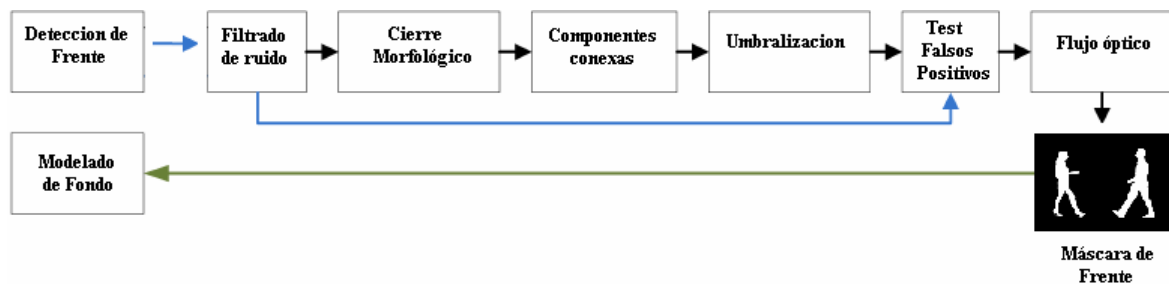


Figura 2-8: Diagrama de un sistema de segmentación de objetos con técnicas de post-procesado.

Por lo tanto, es importante tener en cuenta en este proyecto, que sólo abordaremos uno de los módulos de la cadena de procesamiento de un sistema real, la sustracción de fondo, y el objetivo no es proporcionar los mejores resultados posibles, labor que sería sencilla combinando estas técnicas con las de pre y post-procesado; si no la de estudiar en profundidad los modelos de sustracción de fondo más relevantes del estado del arte a fin de proporcionar un análisis comparativo que contribuya al estado del arte.

2.6 Medidas de evaluación de algoritmos

Generalmente, los resultados de los algoritmos de segmentación de objetos en movimiento son máscaras binarias donde se indica la presencia de objeto en movimiento o la presencia de fondo de la escena para cada píxel. Este hecho, a nivel de píxel, acarrea una serie de problemas como los vistos en apartados anteriores: píxeles relativos a ruido, píxeles detectados aislados, agujeros en el interior de los objetos detectados, etc.

Cheung y Kamath [26] definen como proceso de evaluación de datos un conjunto de medidas que se ejecutan comparando la salida de los algoritmos de sustracción de fondo con las máscaras perfectas generadas a mano de las secuencias analizadas. En esta comparación se pueden analizar diferentes medidas: falsos positivos, verdaderos positivos, falsos negativos y verdaderos negativos. Y a su vez, éstas pueden combinarse en otras medidas muy utilizadas en el estado del arte: precisión, ‘*recall*’ o sensibilidad, especificidad, ‘*fscore*’ o las curvas ROC.

Para evaluar el nivel de fallos se utilizan los indicadores:

- **Falsos positivos, falsas alarmas o ‘*False positive*’ (FP):** Se producen cuando los objetos detectados (píxeles marcados con valor 1) han sido etiquetados incorrectamente, es decir, son píxeles en ausencia de movimiento (deberían ser píxeles con valor 0).
- **Falsos negativos, ‘*False negative*’ (FN):** Ocurren si los píxeles pertenecientes al objeto en movimiento (deberían tener valor 1) no han sido detectados, por tanto, se han etiquetado erróneamente como fondo (son iguales al valor 0).

Este tipo de errores podría solucionarse de forma sencilla mediante técnicas de post-procesado; por medio de filtros que eliminen pequeños grupos de píxeles o, a través de operaciones morfológicas que suavicen los bordes de los objetos.

Por otro lado, se puede evaluar el acierto en los píxeles de frente y fondo con las siguientes medidas:

- **Verdaderos positivos, detecciones de objeto correctas o ‘*True positive*’ (TP):** Corresponden a los objetos (con valor 1) etiquetados correctamente.
- **Verdaderos negativos, detecciones de fondo correctas, ‘*True negative*’ (TN):** Son los píxeles pertenecientes al fondo (con valor 0) detectados correctamente.

En la siguiente tabla se describen estas medidas:

<i>Medida</i>	<i>Máscara de la segmentación (‘Foreground mask’)</i>	<i>Máscara de video base (‘ground truth mask’)</i>
FP	V	F
FN	F	V
TP	V	V
TN	F	F

Tabla 2-1. Medidas de evaluación.

Una forma de llevar el recuento de los píxeles de frente y fondo es calcular la equivalencia entre las medidas anteriores y el número total de píxeles positivos y negativos de la máscara correcta (*‘ground truth’*). De tal forma que, el número de positivos totales en la secuencia será la suma de los **TP** (detectados como parte del objeto y acertados) más **FN** (detectados como fondo y erróneos); mientras que los negativos totales de la máscara serán los **FP** (detectados fondo y correctos) más **TN** (detectados objeto cuando realmente pertenecen al fondo).

En la literatura, podemos encontrar otras medidas que combinan las anteriores con el fin de comparar algoritmos en base de distintos criterios. Son las expuestas a continuación

- **Precisión, valor de predicción positivo:** Esta medida está definida por el cociente entre el número de ejemplos que se clasifican correctamente para una clase C_z y el número total de ejemplos detectados por el sistema de esa clase. Si se denomina N_{zz} el número de elementos clasificados correctamente dentro de la clase C_z , y N_{jz} todos los elementos detectados por el sistema dentro de la clase C_z , la precisión del conjunto C_z viene dada por la ecuación (2.16):

$$\mathbf{Precision} (\text{Modelo}) = \frac{(N_{11} + N_{22} + \dots + N_{zz})}{\text{Total}_{\text{ejemplos}}} \quad (2.16)$$

$$\mathbf{Precision} (C_1) = \frac{N_{11}}{N_{11} + N_{21} + \dots + N_{z1}}$$

$$\mathbf{Precision} (C_2) = \frac{N_{22}}{N_{12} + N_{22} + \dots + N_{z2}}$$

$$\mathbf{Precision} (C_z) = \frac{N_{zz}}{N_{1z} + N_{2z} + \dots + N_{zz}}$$

- **Especificidad o ‘True Negative Rate’:** Es la proporción de ejemplos que han sido clasificados dentro de las otras clases diferente a la clase C_z .

$$\text{TNRate} = 1 - \text{FP} = 1 - \frac{\text{FP}}{\text{FP} + \text{TN}} = \frac{\text{TN}}{\text{TN} + \text{FP}} \quad (2.17)$$

$$\text{TNRate}(C_z) = \frac{N_{22} + N_{33} + \dots + N_{zz}}{(N_{21} + \dots + N_{2z}) + (N_{31} + \dots + N_{3z}) + \dots + (N_{z1} + \dots + N_{zz})} \quad (2.18)$$

- **Sensibilidad o TPR (‘True Positive Rate’) o ‘Recall’:** Corresponde al número de elementos N_{zz} detectados que realmente tienen la clase C_z de entre todos los elementos N_{zj} que son de la clase C_z .

$$\mathbf{Recall} (C_1) = \frac{N_{11}}{N_{11} + N_{12} + \dots + N_{1z}} \quad (2.19)$$

$$\mathbf{Recall}(C_2) = \frac{N_{22}}{N_{21} + N_{22} + \dots + N_{2z}}$$

$$\mathbf{Recall}(C_z) = \frac{N_{zz}}{N_{z1} + N_{z2} + \dots + N_{zz}}$$

- **F-Measure:** Es una medida que combina la Precisión con el ‘Recall’ o TPR con la misma probabilidad para la clase C_z .

$$Fscore(C_z) = \frac{2 \cdot \text{Precision}(C_z) \cdot \text{Recall}(C_z)}{\text{Precision}(C_z) + \text{Recall}(C_z)} \quad (2.20)$$

- **Curvas ROC.** El espacio ROC (*'Receiver Operating Characteristic'*), define un sistema de coordenadas que representan el compromiso entre sensibilidad (efectividad o *recall*) y especificidad del sistema (precisión); un aumento en la sensibilidad (efectividad o *recall*) está acompañado por un decremento en la especificidad (precisión). Es decir, las curvas ROC muestran la relación entre las muestras clasificadas adecuadamente y las muestras que no pertenecen a la clase pero se clasificaron como si lo fueran.

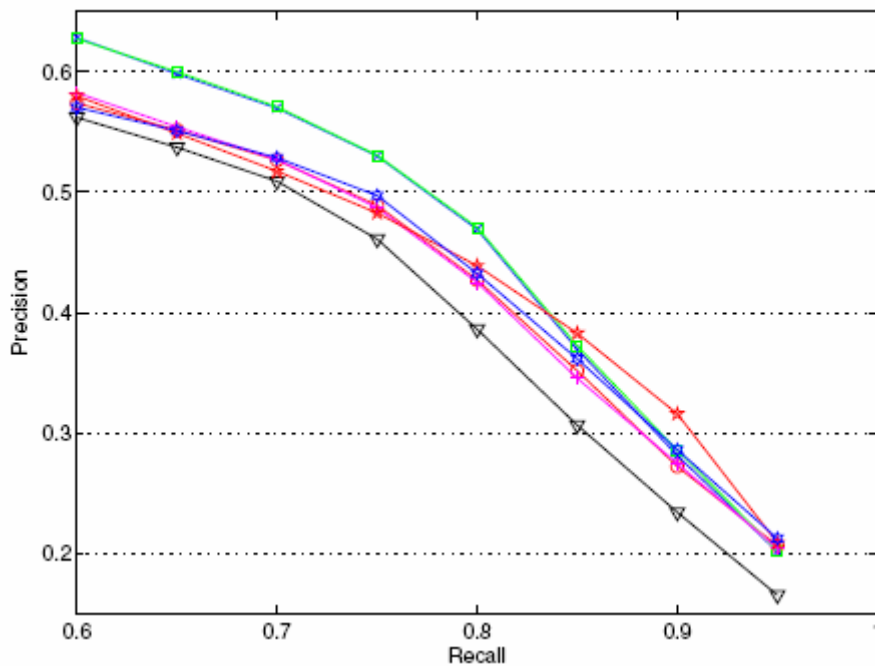


Figura 2-9: Ejemplo de las curvas ROC obtenidas con distintos algoritmos de segmentación de objetos sobre diferentes secuencias de un banco de pruebas.

3 Diseño

A lo largo de este capítulo se describen los algoritmos de segmentación basados en modelado de fondo implementados en este proyecto clasificándolos en tres categorías: modelos básicos, métodos paramétricos y métodos no paramétricos.

En las siguientes subsecciones se abordarán tres cuestiones importantes en el diseño de cada algoritmo: encontrar las regiones de fondo, mantener el modelo de fondo en el tiempo y extraer las regiones donde se encuentran los objetos en movimiento. Además se analizarán los bloques funcionales creados para cada algoritmo y las razones por las cuales se han tomado determinadas decisiones.

En cuanto a la notación, en adelante se utilizará $B_t(x, y)$, $I_t(x, y)$ y $F_t(x, y)$ para designar el valor de intensidad del fondo, la imagen de la secuencia y la máscara del frente, respectivamente, en el píxel localizado en la posición (x, y) y el instante t .

Los algoritmos de segmentación de objetos basados en modelado de fondo implementados siguen un diagrama de flujo que mostramos en la **Figura 3-1**. Cada módulo realiza diferentes operaciones matemáticas en función del método, que serán explicadas detalladamente en los apartados siguientes. Las flechas marcadas en cursiva sólo se ejecutan una vez en el procesamiento de la secuencia y el módulo marcado con un recuadro naranja sólo se ejecuta en los métodos paramétricos. El diagrama de flujo es el siguiente:

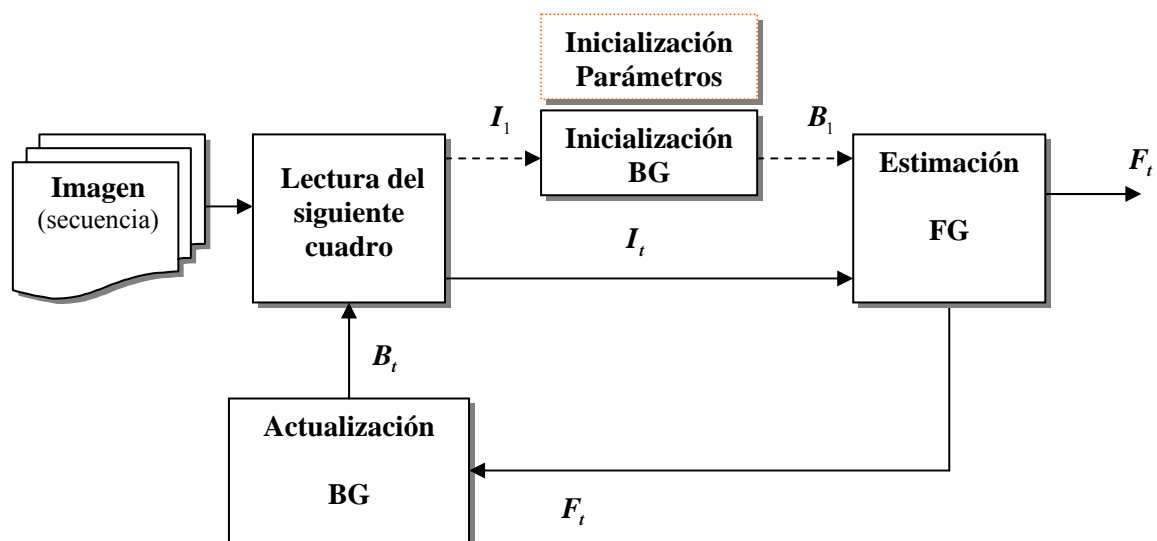


Figura 3-1: Diagrama de flujo de los algoritmos de segmentación.

3.1 Modelos básicos

El término “modelos básicos de sustracción de fondo” sirve para designar al conjunto de métodos matemáticos de operaciones sencillas con las que modelar los píxeles de fondo.

El modelo de fondo en estos algoritmos corresponde, en cada instante, a una medida de la variación temporal de un número de imágenes anteriores. Según la operación matemática aplicada a esta variación para obtener el modelo de fondo de la secuencia, podemos dividir estos algoritmos en:

- **Diferencia con el cuadro anterior:** El modelo de fondo corresponde a la imagen anterior.
- **Mediana de los ‘N’ cuadros previos:** La imagen de fondo se calcula con la mediana de un número de imágenes anteriores.
- **Generación del fondo basada en histogramas:** El fondo corresponde, para cada píxel, a la media del histograma de sus valores en un número de imágenes anteriores.

En general, se puede destacar de estas técnicas la rápida adaptación a los cambios de fondo, ya que no dependen de las imágenes anteriores más allá de las consideradas, y este número suele ser bajo. Por otro lado, en esta situación, el requisito de almacenamiento no es significativo. Sin embargo, no son algoritmos robustos frente al ruido, cambios de iluminación o cambios del fondo. Estas son, entre otras, las ventajas e inconvenientes de estos modelos que veremos en más detalle para cada método en los apartados posteriores.

3.1.1 Diferencia con el cuadro anterior

La diferencia con el cuadro anterior, método también denominado diferencia de imágenes (*Framme difference*), es posiblemente la técnica de sustracción de fondo más sencilla de las encontradas en el estado del arte.

El modelo de fondo corresponde en cada instante de la secuencia a la imagen procesada en el instante anterior. Para detectar el objeto en movimiento, se mide la diferencia de los píxeles de imágenes consecutivas como ilustra la **Figura 3-2**:

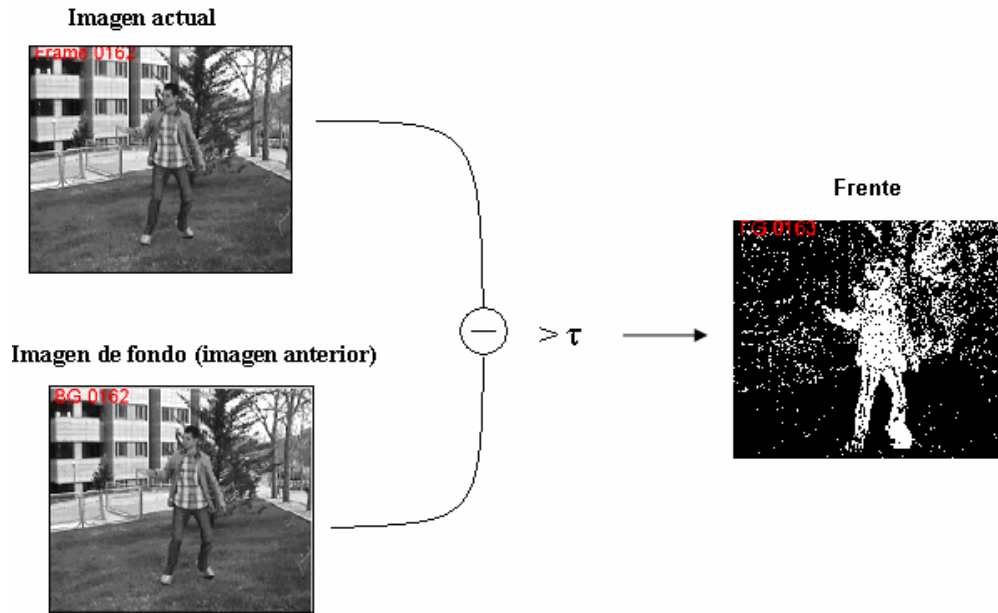


Figura 3-2: Diferencia entre la imagen actual y la de fondo y cálculo de la imagen binaria del frente.

En el diseño que se ha implementado para este método, el modelo de fondo denominado B_t se inicializa al comienzo de la ejecución (módulo del diagrama Inicialización BG) para cada píxel situado en la posición (x, y) con el valor de la primera imagen I_1 de la secuencia en la misma posición (x, y) .

$$B_1(x, y) = I_1(x, y) \quad (3.1)$$

La estimación del frente F_t se realiza comparando la imagen de fondo B_t con la imagen actual I_t calculando su diferencia. A esta diferencia se le aplica un umbral τ : si la diferencia supera el umbral, se considera que la variación entre imágenes ha sido elevada a causa de un objeto en movimiento; los píxeles en los que la diferencia sea inferior a dicho umbral son píxeles con ausencia de movimiento, es decir, pertenecientes al fondo.

$$\begin{aligned} |I_t(x, y) - B_t(x, y)| > \tau &\Rightarrow F_t(x, y) = 1 \\ |I_t(x, y) - B_t(x, y)| \leq \tau &\Rightarrow F_t(x, y) = 0 \end{aligned} \quad (3.2)$$

, donde τ es un umbral predefinido que determina el parecido entre imágenes consecutivas.

La actualización del modelo se obtiene asignando al fondo la imagen anterior de la secuencia.

$$B_1(x, y) = I_1(x, y) \quad (3.3)$$

Este modelo, es muy sensible a fondo en movimiento, movimiento de la cámara, cambios de brillo y ruido; por este motivo, para estimar el movimiento y compensar pequeños cambios del fondo se proponen nuevos métodos, como por ejemplo, el cálculo de la mediana de N imágenes anteriores o el uso de histogramas, ya que son menos sensibles las situaciones mencionadas anteriormente.

3.1.2 Mediana de los 'N' cuadros previos

El filtrado de mediana, resuelve alguno de los inconvenientes que presentaba el modelo anterior al utilizar como modelo de fondo, en cada instante, la mediana de un número N de imágenes anteriores. A continuación, como en el método anterior, el fondo se resta a la imagen actual y esa diferencia se compara con un umbral para determinar los píxeles de frente.

En el diseño que se ha desarrollado, el fondo B_t se inicializa con la primera imagen I_1 .

$$B_1(x, y) = I_1(x, y) \quad (3.4)$$

La detección del objeto en movimiento F_t se lleva a cabo umbralizando de manera similar al método de la diferencia.

$$\begin{aligned} |I_t(x, y) - B_t(x, y)| > \tau &\Rightarrow F_t(x, y) = 1 \\ |I_t(x, y) - B_t(x, y)| \leq \tau &\Rightarrow F_t(x, y) = 0 \end{aligned} \quad (3.5)$$

La actualización del fondo funciona de la siguiente forma: si un píxel en la imagen actual tiene un valor mayor que el correspondiente píxel de fondo, el fondo se incrementa en uno. Del mismo

modo, si el píxel de la imagen actual es menor que el píxel de fondo localizado en la misma posición (x, y) , el fondo B_t se decrementa en una unidad.

$$\begin{aligned}
 I_t(x, y) > B_t(x, y) &\Rightarrow B_t(x, y) = B_{t-1}(x, y) + 1 \\
 I_t(x, y) < B_t(x, y) &\Rightarrow B_t(x, y) = B_{t-1}(x, y) - 1 \\
 I_t(x, y) = B_t(x, y) &\Rightarrow B_t(x, y) = B_{t-1}(x, y)
 \end{aligned}
 \tag{3.6}$$

De esta forma, el fondo con el tiempo converge a una estimación en la que la mitad de los píxeles de entrada son mayores que el fondo, y la mitad tienen menor valor que el fondo, lo cual es una aproximación de la mediana (la convergencia puede variar dependiendo de la velocidad de la secuencia y la cantidad de movimiento de la escena).

La aproximación por mediana es un método robusto ya que produce unos resultados comparables a otros métodos de mayor complejidad. No obstante, presenta inconvenientes a la hora de almacenar y procesar un número elevado de imágenes que requieran una gran cantidad de memoria.

3.1.3 Generación del fondo basada en histogramas.

La generación del fondo basada en histogramas consiste en obtener para cada píxel una función de distribución de probabilidad mediante el cálculo del histograma su valor en un conjunto de imágenes consecutivas.

La siguiente figura (**Figura 3-3**) muestra la imagen y frente obtenidos con el método de Generación de Histogramas, capturados en el instante 46 de una de las secuencias de prueba. Con un círculo amarillo se marca el píxel seleccionado para observar su histograma.



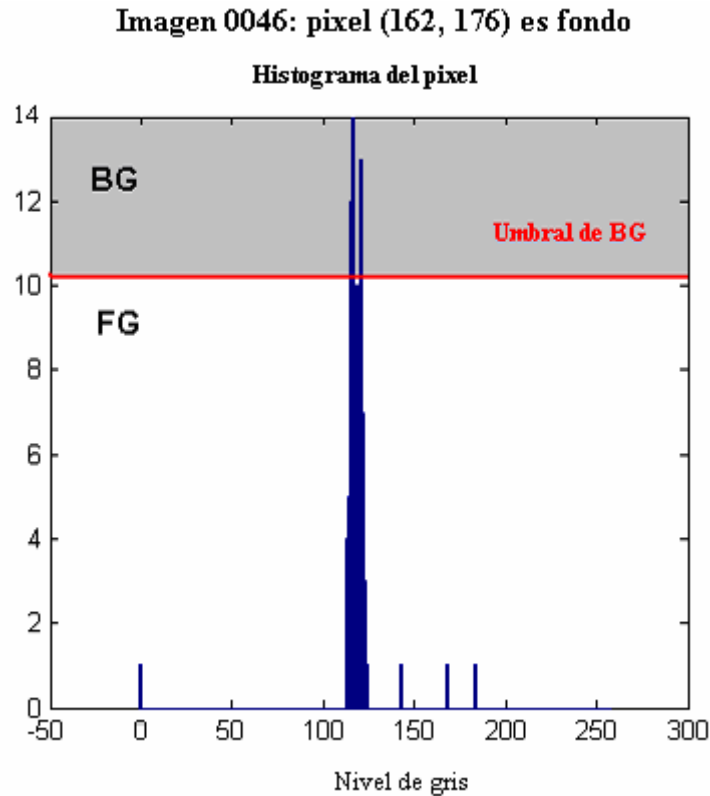


Figura 3-3: Histograma del píxel localizado en la posición (162,176) de la imagen 46 de la secuencia Tennis1 de los videos de prueba.

Al histograma de cada píxel se le asigna un umbral; así pues los futuros valores del píxel que sean inferiores al valor del umbral serán considerados objeto en movimiento (FG), mientras que, los que tengan un valor superior serán considerados píxeles de fondo (BG).

En primer lugar, se deben almacenar en un buffer para cada píxel los valores que permitirán configurar su histograma, es decir, las N imágenes anteriores a cada instante considerado. Si el número de valores almacenados (N) es menor al de imágenes procesadas i , el histograma se construye con los valores de las imágenes procesadas; si ya se han procesado N imágenes, el histograma se calcula con los valores de los píxeles de las N imágenes anteriores.

$$\begin{aligned}
 i < N &\Rightarrow B_1(x, y) = \text{Hist} [I_1(x, y), \dots, I_i(x, y)] \\
 i \geq N &\Rightarrow B_1(x, y) = \text{Hist} [I_{i-N}(x, y), \dots, I_i(x, y)]
 \end{aligned}
 \tag{3.7}$$

El umbral para cada píxel en el método de Generación de Histogramas se denomina $U(x, y)$ para distinguirlo del resto de umbrales (τ) de otros métodos que no se calculan para cada píxel sino que

se fijan para toda la imagen a un valor inicial. Este umbral $U(x,y)$ se determina como el valor máximo de los valores almacenados en el buffer dividido por el número de valores almacenados:

$$\begin{aligned} i < N &\Rightarrow U(x, y) = \frac{\max(B_1(x, y))}{i} \\ i \geq N &\Rightarrow U(x, y) = \frac{\max(B_1(x, y))}{N} \end{aligned} \quad (3.8)$$

Ese umbral se utiliza para las siguientes imágenes de la secuencia hasta que se actualiza el buffer para determinar los píxeles que pertenecen al fondo si su valor es inferior al del umbral, y los que son frente si superan dicho umbral.

$$\begin{aligned} I_t(x, y) > U(x, y) &\Rightarrow F_t(x, y) = 0 \\ I_t(x, y) \leq U(x, y) &\Rightarrow F_t(x, y) = 1 \end{aligned} \quad (3.9)$$

La actualización del buffer con el que se construyen los histogramas se realiza cada L imágenes, donde $L \ll N$.

Existe una gran variedad de formas de implementar este algoritmo: por ejemplo, la elección del umbral variará en función de las características de la secuencia; por otro lado, se podría calcular además de la media, una varianza y así tener un rango mayor de píxeles para discernir entre fondo y frente. También se podrían obtener varios umbrales locales para un único histograma y aplicarlos para extraer el objeto en movimiento [62]. Estas y otras variantes se han implementado aunque solamente reflejaremos los resultados del método sencillo representado por las fórmulas anteriores ya que la utilización de varianzas supone una mayor complejidad del método y estos cálculos se abordarán en los métodos paramétricos.

El inconveniente más importante de este método es el requerimiento de memoria; se necesitan tantos buffers de tamaño N como píxeles tenga la imagen. Por otro lado, no ofrece buenos resultados en fondos con pequeñas variaciones ya que para decidir el frente sólo utiliza un único valor de píxel con el que decidir los píxeles del frente y del fondo durante L imágenes en las que se mantiene el histograma, de modo que si los píxeles de la escena sufren cambios en torno a una varianza no será capaz de detectarlos hasta las L imágenes siguientes cuando el histograma del píxel se actualice. Este es el motivo por el que aparecen los métodos paramétricos, en los que cada píxel de la imagen se modela con una función de distribución de probabilidad.

3.2 Métodos paramétricos

Los modelos paramétricos estiman el valor de cada píxel del fondo de las secuencias por medio de funciones de distribución de probabilidad que se actualizan con cada nueva imagen de entrada. En lugar de mantener un buffer con el que estimar el fondo, los píxeles de fondo se representan por los parámetros de la distribución, lo cual supone un ahorro de memoria muy importante.

La función de distribución más utilizada para modelar el fondo es la Gaussiana $N(\mu, \sigma^2)$. Existen dos variantes fundamentales de este método:

- **Gaussiana simple:** Cada píxel del fondo se representa con dos parámetros (μ, σ^2) .
- **Mezcla de Gaussianas:** Los píxeles del fondo se estiman mediante una combinación de varias distribuciones de tipo Gaussiano, con un número menor a un máximo K definido para el modelo, cada una de ellas determinada por dos parámetros (μ, σ^2) .

A continuación, se explicará de manera detallada la formulación empleada para implementar estos métodos.

3.2.1 Gaussiana Simple (SG)

Tal y como describe Wren en los artículos [11][30], el modelado de fondo por medio de una Gaussiana Simple (SG) se basa en el ajuste de la densidad de probabilidad de una función Gaussiana (pdf) en cada píxel de la imagen de la secuencia de video a procesar.

Para cada píxel del modelo de fondo $B_t(x, y)$, se mantiene un valor de luminancia medio $\mu_t(x, y)$ y una desviación estándar $\sigma_t(x, y)$ (para imágenes en color existirá una media y una desviación para cada canal de color).

La **Figura 3-4** presenta la imagen, frente y fondo capturados en el instante 184 de la ejecución del método Gaussiana Simple sobre una de las secuencias de prueba. En la imagen aparece un círculo amarillo que representa el píxel seleccionado para analizar su modelo de fondo Gaussiano.

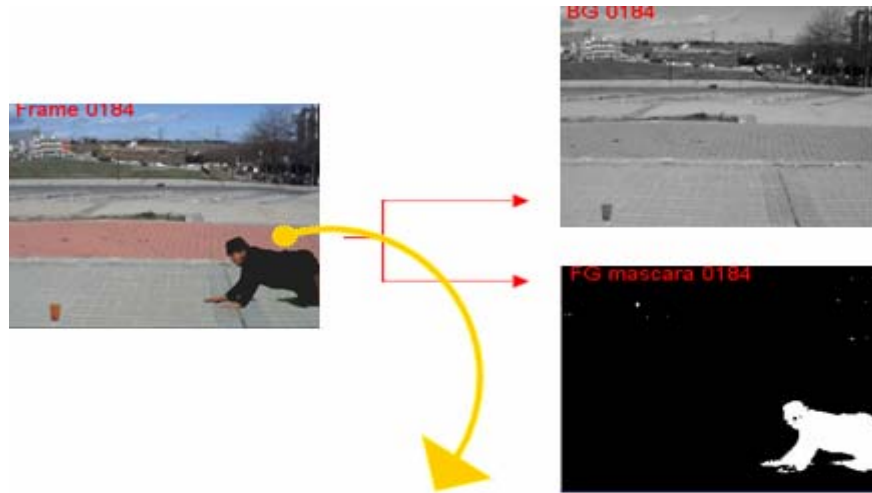


Imagen 0184: pixel (188, 321) es fondo

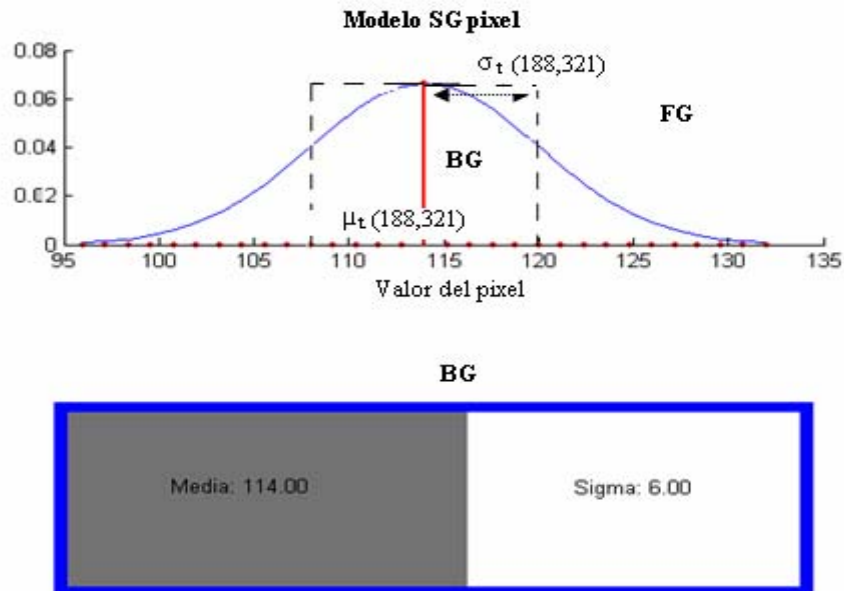


Figura 3-4: Distribución Gaussiana del píxel (188, 321) de la secuencia Hombre hambriento en el instante 184.

La gráfica superior representa la distribución Gaussiana de media $\mu_{184}(188,321)=114$ y desviación $\sigma_{184}(188,321)=6$ del píxel localizado en la posición (188, 321) en la imagen 184 de la secuencia “Hombre hambriento”.

Los píxeles de cada nueva imagen de la secuencia se comparan con los valores de media que modelan el fondo, de forma que si la diferencia supera c veces la desviación establecida para cada píxel (típicamente se suele usar un valor de 2-3 desviaciones), es decir, no encaja en el *interior* de la Gaussiana que modela el píxel, se decide que el píxel es parte del objeto en movimiento (FG). Por el contrario, si el píxel posee un valor de entre los que se encuentran en el interior de la

Gaussiana (su diferencia con la media no excederá 2-3 veces la desviación) se considerará píxel del fondo (BG).

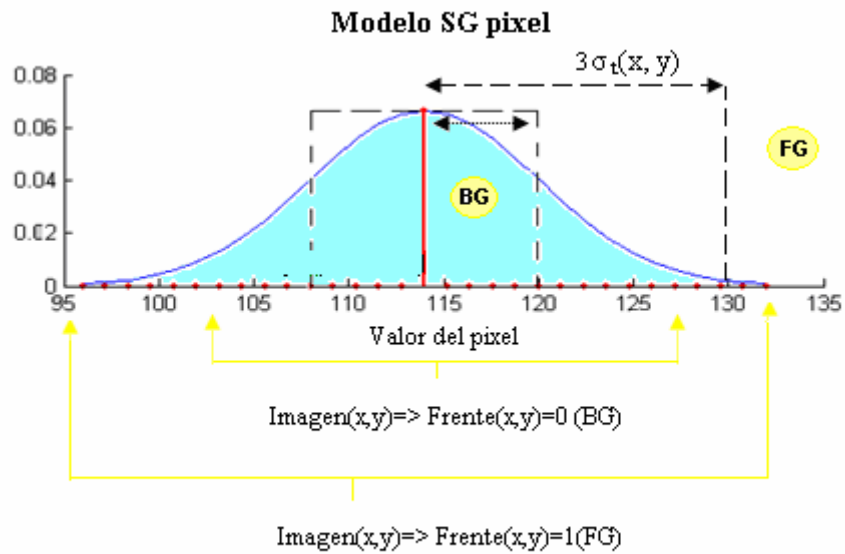


Figura 3-5: Distribución Gaussiana de un píxel.

A continuación, se explicarán las operaciones realizadas para implementar este algoritmo.

❑ Inicialización BG y parámetros

El método inicializa la imagen de fondo a la primera imagen de la secuencia y como valor de desviación se marcan todos los píxeles a un valor pequeño determinado (P). Este valor de desviación varía en función de las características de la secuencia: para secuencias cuyo fondo contiene mucho movimiento será elevado (por ejemplo, 25 o 35) para modelar a los píxeles con una Gaussiana más ancha y permitir más rango de variación en los valores que los píxeles pueden tomar en las imágenes de la secuencia. Y será un valor pequeño P (por ejemplo, 5 o 10) si la secuencia contiene un fondo de escaso movimiento, para ajustar el fondo a la secuencia y discernir con mayor precisión los píxeles que forman parte de objeto en movimiento.

$$B_1(x, y) = \left\{ \mu_1(x, y), \sigma_1(x, y) \right\} / \begin{cases} \mu_1(x, y) = I_1(x, y) \\ \sigma_1(x, y) = P \end{cases} \quad (3.10)$$

Estimación del FG

La extracción del objeto en movimiento se realiza calculando la probabilidad de que el píxel de la nueva imagen procesada pertenezca a la Gaussiana que modela el fondo. Una función Gaussiana modela una densidad de probabilidad que deja de tener valor a aproximadamente de 2 a 3 veces el valor de la desviación de la función. Por este motivo, si la diferencia entre la imagen y el modelo de fondo (la media) del píxel difiere en $c = 2-3$ veces el valor de la desviación, se considera que no hay parecido entre el nuevo píxel y su modelo, y se determina que el píxel es objeto en movimiento (F_t). Mientras que, si sucede lo contrario, será marcado como píxel de fondo.

$$\begin{aligned} |I_t(x, y) - \mu_t(x, y)| \geq c \sigma_t(x, y) &\Rightarrow F_t(x, y) = 1 \\ |I_t(x, y) - \mu_t(x, y)| < c \sigma_t(x, y) &\Rightarrow F_t(x, y) = 0 \end{aligned} \quad (3.11)$$

Actualización BG

Los parámetros que modelan el fondo de la secuencia se actualizan por medio de una media móvil caracterizada por la velocidad de cambio de la secuencia ($\alpha \ll 1$). Este parámetro, que depende de la rapidez de los movimientos de la escena, es muy decisivo en los resultados y ha sido estudiado en profundidad a lo largo de este proyecto.

$$B_{t+1}(x, y) / \begin{cases} \mu_{t+1}(x, y) = \alpha I_t(x, y) + (1 - \alpha)\mu_t(x, y) \\ \sigma_{t+1}^2(x, y) = \alpha (I_t(x, y) - \mu_t(x, y))^2 + (1 - \alpha)\sigma_t^2(x, y) \end{cases} \quad (3.12)$$

Entre las ventajas que presenta este modelo se encuentran: la sencillez, rapidez de ejecución, adaptación automática del ruido y de los cambios progresivos del fondo. Por el contrario, no soporta fondos multimodales en los que los píxeles oscilan entre varios valores muy distintos (parpadeo de hojas, luces intermitentes,..). Además, cada secuencia posee un valor diferente de varianza (σ^2) y velocidad de cambio (α) que es necesario ajustar para optimizar los resultados.

3.2.1.1 Filtro de análisis espacial.

Con el objetivo de soportar fondos multimodales en los que se producen movimientos de los píxeles a píxeles cercanos (por ejemplo, movimientos de hojas), estimamos la probabilidad de

parecido del píxel actual no sólo con el localizado en el fondo en la misma posición, si no con un bloque $V \times V$ de píxeles en torno a éste de modo que si el píxel (x, y) se parece a alguno de su entorno (i, j) entonces el píxel se considera que se ha movido pero es parte del fondo. Esta técnica de post-procesado permite eliminar gran parte de píxeles detectados erróneamente como frente.

$$\begin{aligned} \exists(i, j) \in [V, V] / |I_t(x, y) - \mu_t(i, j)| \leq c\sigma_t(i, j) &\Rightarrow F_t(x, y) = 0 \\ \forall(i, j) \in [V, V] / |I_t(x, y) - \mu_t(i, j)| > c\sigma_t(i, j) &\Rightarrow F_t(x, y) = 1 \end{aligned} \quad (3.13)$$

A continuación, se muestra una figura (**Figura 3-6**) con la imagen, fondo y frente obtenidos en el instante 130 al ejecutar el método de la SG con el filtro de análisis espacial sobre una de las secuencias de prueba. En las tres imágenes se marca un píxel con un círculo amarillo y un rectángulo rojo a su alrededor que indica el vecindario con el que se realiza el filtrado del píxel.

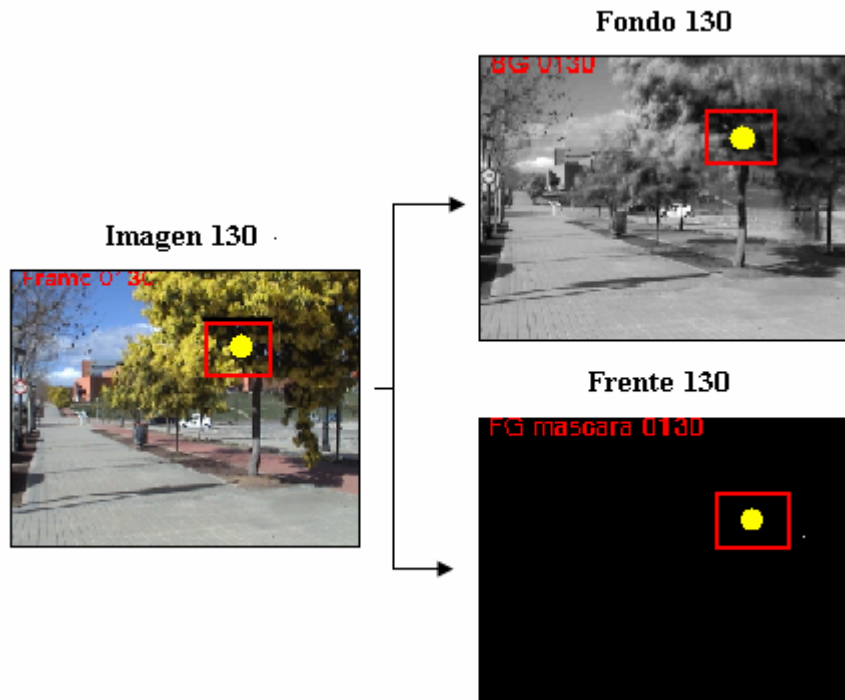


Figura 3-6: Imagen, modelo SG de fondo y frente capturados en el instante 130 de la secuencia. El círculo amarillo indica el píxel seleccionado y el rectángulo rojo a su alrededor es el vecindario escogido para realizar el filtrado espacial

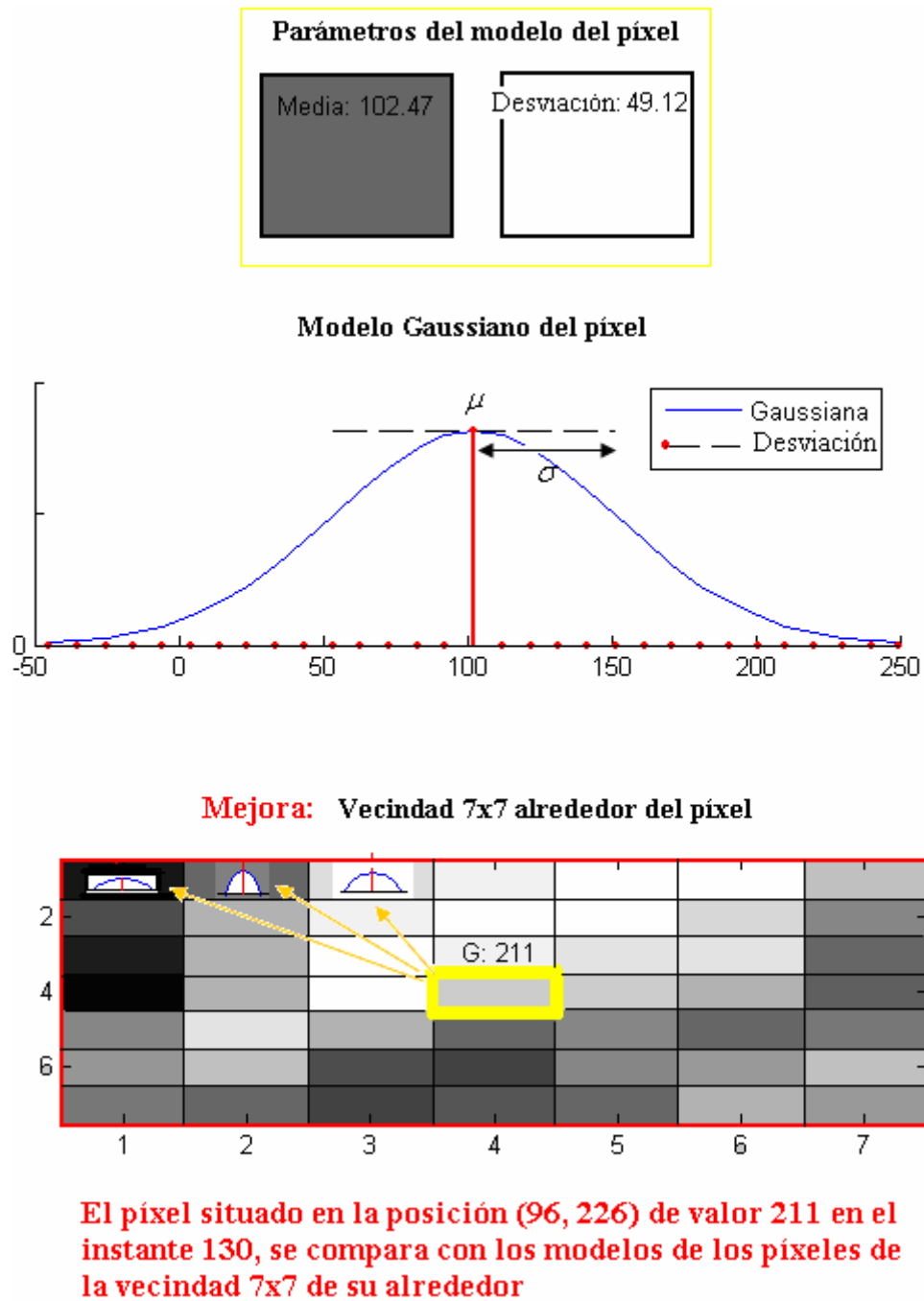


Figura 3-7: Parámetros del modelo de Fondo Gaussiano del píxel situado en la posición (96, 226) en el instante 130 de la secuencia.

El primer elemento de la **Figura 3-7** muestra los valores de la media y varianza del modelo Gaussiano del píxel localizado en la posición (96, 226) en el instante 130. La gráfica situada más abajo representa la función Gaussiana que modela al píxel en el instante 130 de captura. Y la última representación, es un ejemplo que muestra la imagen del modelo de fondo (medias de las Gaussianas) de los 7 x 7 píxeles alrededor del píxel (96, 226). El modelo de cada uno de los píxeles de la vecindad se compara con el valor del píxel en el instante 130 para calcular si existe parecido

(su valor difiere menos de tres veces la varianza del modelo de algún píxel de alrededor) y, por este motivo, determinar que el píxel (considerado en una primera detección como frente) es fondo.

3.2.2 Mezcla de Gaussianas (MoG)

La Mezcla de Gaussianas (MoG) es uno de los métodos complejos más utilizado en la literatura [67][68]. Entre sus ventajas destaca la robustez, ya que permite modelar fondos multimodales, es decir, manejar múltiples modos de distribución (o tipos de movimiento). Por ejemplo, una hoja agitándose bajo un cielo azul presenta dos modos o tipos de movimiento: el de las hojas y el cielo.

En este algoritmo, el fondo no se modela por los valores de una imagen sino a través de un modelo paramétrico cuyo objetivo es aproximar una función de distribución a los últimos valores de cada píxel. Cada píxel p_t localizado en la posición (x, y) de la imagen de fondo en el instante t se representa por un número K de distribuciones Gaussianas cuya combinación constituye una función de densidad de probabilidad $F(p_t)$.

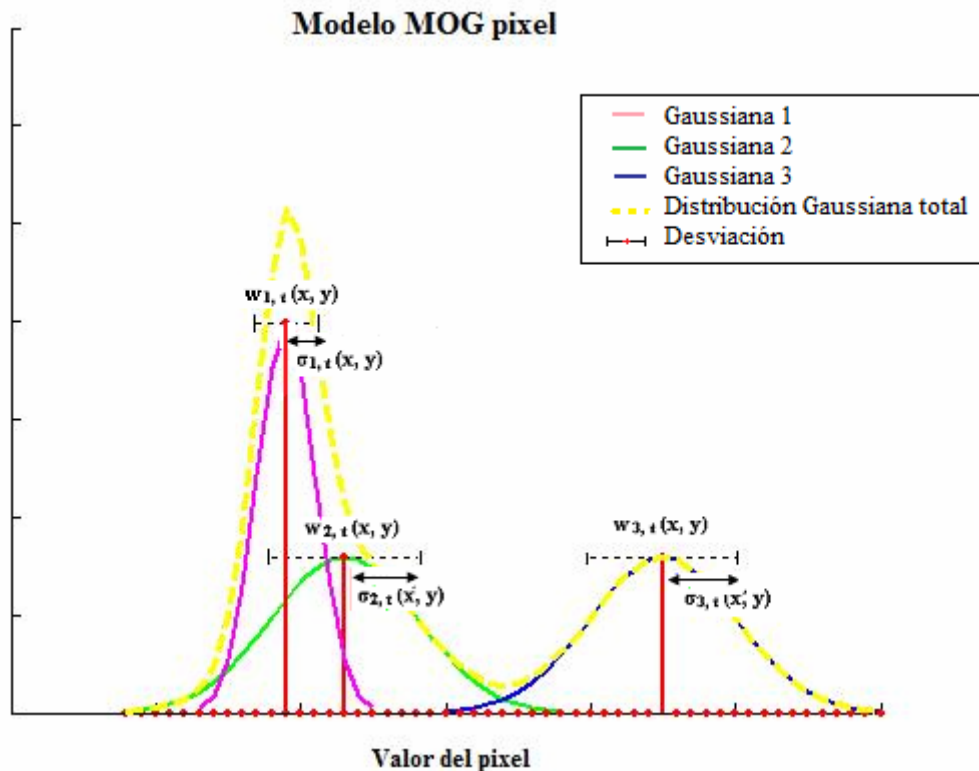


Figura 3-8: Distribución de la Mezcla de Gaussianas de un píxel.

$$F(p_t) = \sum_{i=1} w_{i,t} \cdot \eta(p_t; \mu_{i,t}, \sigma_{i,t}) \quad (3.14)$$

$$\sum_{i=1}^{i=K} w_{i,t} = 1 \quad (3.15)$$

Donde $w_{i,t}$ es el parámetro que indica el peso de la componente Gaussiana i -ésima en el instante t y $\eta(p_t; \mu_{i,t}, \sigma_{i,t})$ la componente o Gaussiana i -ésima del píxel p_t , que posee una media $\mu_{i,t}$ y una desviación $\sigma_{i,t}$, representada por:

$$\eta(p_t; \mu_{i,t}, \sigma_{i,t}) = \frac{1}{(2\pi)^{\frac{1}{2}} \sigma_{i,t}} e^{-\frac{1}{2} \left(\frac{p_t - \mu_{i,t}}{\sigma_{i,t}} \right)^2} \quad (3.16)$$

Las medias de la Gaussianas (o componentes) son los valores medios que toman los píxeles a lo largo de la secuencia ponderados por un peso. El número de Gaussianas K (componentes por píxel) son típicamente 3-5 y se escoge generalmente en función de las limitaciones de memoria.

A continuación se explicarán paso a paso las operaciones realizadas para modelar el fondo con MoG y extraer el frente.

❑ Inicialización de Parámetros

En primer lugar, se inicializan los parámetros que representan el fondo: media, desviación y peso. Para ello, es necesario tomar una serie de decisiones iniciales; en nuestro caso, hemos inicializado la media de una de las Gaussianas (por ejemplo, la Gaussiana $i = 1$) a la primera imagen y el resto de medias a valores aleatorios. Con ello consideramos que la primera Gaussiana modelará el píxel en la primera imagen. Por este motivo, el peso w o porcentaje de distribución para $k = 1$ debe ser un valor alto (próximo a 1) y para el resto será un valor muy pequeño (próximo a 0), ya que la suma de los pesos de las Gaussianas debe ser 1. En cuanto a la desviación inicial dependerá del tipo de secuencia, generalmente, se da un valor elevado a las tres componentes.

$$B_1(x, y) = \{ \mu_i(x, y), \sigma_i(x, y), i = 1 \dots K \} \quad /$$

$$\begin{cases} i = 1 \Rightarrow \mu_{i,1}(x, y) = I_1(x, y) & \sigma_{i,1}(x, y) = P & w_{i,1} = 1 \\ 1 < i \leq K \Rightarrow \mu_{i,1}(x, y) = \text{aleatorio} & \sigma_{i,1}(x, y) = P & w_{i,1} = 0 \end{cases}, \quad \sum_{i=1}^K w_{i,t} = 1 \quad (3.17)$$

Estimación del BG y actualización de parámetros

Para obtener los píxeles pertenecientes al fondo, se calcula la imagen diferencia entre los K modelos posibles; si se encuentra parecido con alguna distribución, es decir, si la diferencia difiere en c (2-3) veces el valor estimado de la desviación correspondiente a la distribución, entonces el píxel se marca como fondo y se actualizan los parámetros de dicha distribución a través de una media móvil:

$$\exists i \in [1, K] / |I_t(x, y) - \mu_{i,t}(x, y)| \leq c \sigma_{i,t}(x, y) \quad (3.18)$$



$$\begin{cases} \mu_{i,t+1}(x, y) = \rho I_t(x, y) + (1 - \rho) \mu_{i,t}(x, y) \\ \sigma_{i,t+1}^2(x, y) = \rho (I_t(x, y) - \mu_{i,t}(x, y))^2 + (1 - \rho) \sigma_{i,t}^2(x, y) \\ w_{i,t+1}(x, y) = w_{i,t}(x, y) \end{cases} \quad (3.19)$$

Donde el factor $\rho = \alpha \Pr(I_t(x, y) / \mu_{i,t-1}(x, y), \sigma_{i,t-1}(x, y))$, α es un parámetro predefinido al inicio de la ejecución, $\mu_{i,t}(x, y)$ es el valor de la media del píxel en el instante t e $I_t(x, y)$ es el valor del píxel en el instante t .

Las componentes que no han sido marcadas como píxeles de fondo se actualizan de la siguiente forma; la media y varianza queda con el mismo valor y el peso desciende lo cual significa que la probabilidad de modelar el píxel con esta Gaussiana disminuye:

$$\exists i \in [1, K] / |I_t(x, y) - \mu_{i,t}(x, y)| > c \sigma_{i,t}(x, y) \quad (3.20)$$



$$\begin{cases} \mu_{i,t+1}(x, y) = \mu_{i,t}(x, y) \\ \sigma_{i,t+1}^2(x, y) = \sigma_{i,t}^2(x, y) \\ w_{i,t+1}(x, y) = (1 - \alpha) w_{i,t}(x, y) \end{cases} \quad (3.21)$$

Para que la suma de pesos siga valiendo uno, se normalizan los pesos:

$$w_{i,t+1}(x, y) = \frac{w_{i,t+1}(x, y)}{\sum_{j=1}^{j=K} w_{j,t+1}(x, y)} \quad (3.22)$$

Si no se ha encontrado parecido con ninguna de las Gaussianas que modelaban el píxel, se reemplaza la de menor peso por una nueva Gaussiana de peso muy pequeño y de media el valor del píxel; con ello, se consigue introducir un modo en la distribución del píxel.

Una vez actualizados los parámetros, para obtener una imagen de fondo se calcula para cada píxel la media ponderada de las K distribuciones:

$$B_t(x, y) = \sum_{i=1}^K w_{i,t}(x, y) \cdot \mu_{i,t}(x, y) \quad (3.23)$$

❑ Obtención del FG

En el proceso de detección de píxeles de frente y fondo, en primer lugar, se ordenan los pesos de mayor a menor y se escogen como modelo de fondo de cada píxel las B Gaussianas de mayor peso, es decir, aquellas cuyo peso acumulado supere un umbral (por ejemplo, un 0.85). Si para la Gaussiana b que supere dicho umbral la diferencia entre la media y el píxel de la imagen actual es inferior a c veces su desviación, se considerará píxel de fondo. En caso contrario, será objeto en movimiento F_t (FG).

$$\begin{aligned} B &= \arg \min_b \sum_{i=1}^b w_{i,t}(x, y) > \tau \\ \exists b \in [1, B] / |I_t(x, y) - \mu_{b,t}(x, y)| &\leq c\sigma_{b,t}(x, y) \Rightarrow F_t(x, y) = 0 \\ \forall b \in [1, B], |I_t(x, y) - \mu_{b,t}(x, y)| &> c\sigma_{b,t}(x, y) \Rightarrow F_t(x, y) = 1 \end{aligned} \quad (3.24)$$

La mezcla de Gaussianas es un algoritmo complejo algorítmicamente pero produce buenos resultados. No obstante, presenta un problema de ajuste de parámetros (el porcentaje de la componente de fondo τ , la desviación estándar σ inicial, el número de veces que se toma la desviación de las gaussianas c para decidir entre fondo y frente, el factor de aprendizaje ρ , el número total de componentes de Gaussianas K y el número máximo de componentes B en el modelo de fondo) todos ellos con un impacto significativo sobre el desempeño del algoritmo.

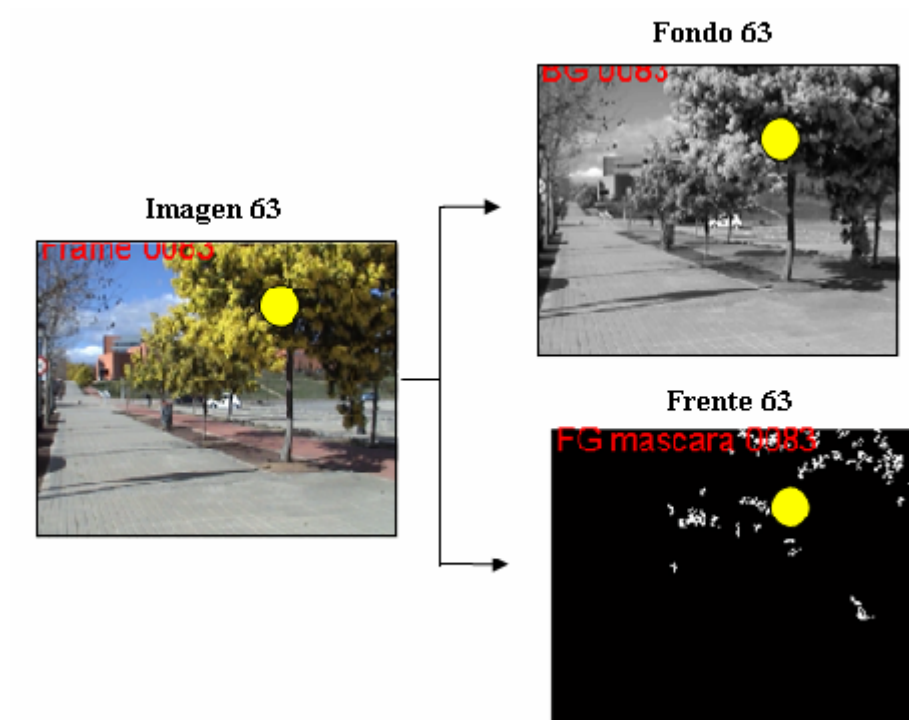


Figura 3-9: Imagen, modelo MoG del Fondo y Frente en la evaluación del píxel (88, 233) de la secuencia ‘Baile 2’ en la imagen 83.

En la figura superior (**Figura 3-9**) se muestra la imagen, el fondo y frente de la secuencia ‘Baile 2’ en el instante 63. Se selecciona el píxel de la imagen de fondo situado en la posición (88, 233) marcado con el círculo amarillo en las tres imágenes para observar el valor de sus parámetros y su evolución a lo largo de la secuencia que se muestra en la **Figura 3-10**.

Imagen 0083: pixel (88, 233) es fondo

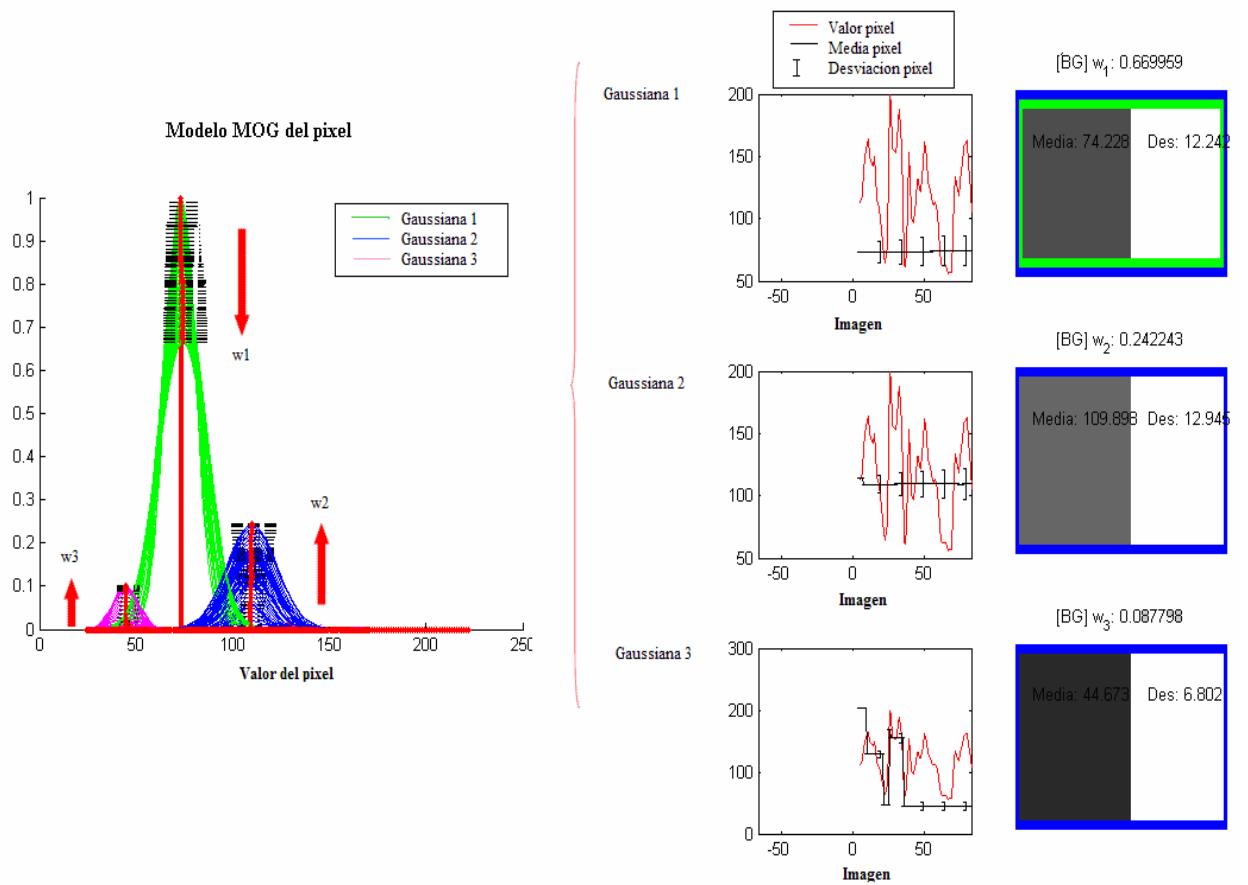


Figura 3-10 : Ejemplo del modelo MOG del píxel (88, 233) de la secuencia Baile Multimodal en la imagen 83.

La gráfica de la izquierda representa la evolución de las tres Gaussianas que modelan el fondo e el píxel localizo en la posición (88,233); la de mayor peso (de color verde) disminuye su peso w en el tiempo mientras que las otras dos (de color rosa y azul) lo incrementan. La figura de la derecha representa para cada Gaussiana el valor del píxel en cada imagen de la secuencia (color rojo) y el valor estimado, es decir la media (color negro) y la desviación (barras verticales de color negro). Finalmente, los tres cuadros son el color en tono de gris del valor medio estimado para cada Gaussiana. El píxel es fondo porque su desviación es inferior a la de la primera Gaussiana (rodeada de verde en el cuadrado de media y desviación) y se modela con las tres Gaussianas porque la suma de los pesos es superior al umbral definido en la etapa de extracción de frente.

3.2.3 Método Gamma

En el método Gamma el modelo de fondo se representa, para cada píxel, por una media y una varianza generadas dinámicamente basándose en información temporal de la secuencia de video analizada.

La detección de los píxeles de fondo y objeto se realiza mediante una umbralización de la diferencia al cuadrado entre la imagen actual I_t y de la imagen que representa el fondo de la secuencia B_t .

$$\begin{aligned} \left(|I_t(x, y) - B_t(x, y)| \right)^2 > \tau &\Rightarrow F_t(x, y) = 1 \\ \left(|I_t(x, y) - B_t(x, y)| \right)^2 \leq \tau &\Rightarrow F_t(x, y) = 0 \end{aligned} \quad (3.25)$$

Para disminuir el efecto del ruido impulsivo de las secuencias de video, la anterior diferencia se promedia en ventanas de tamaño $W \times W$ centradas en cada píxel (x, y) de estudio.

$$\begin{aligned} g(x, y) &= \sum_{i=-W}^{i=W} \sum_{j=-W}^{j=W} \left(|I_t(x+i, y+j) - B_t(x+i, y+j)| \right)^2 > \beta \Rightarrow F_t(x, y) = 1 \\ g(x, y) &= \sum_{i=-W}^{i=W} \sum_{j=-W}^{j=W} \left(|I_t(x+i, y+j) - B_t(x+i, y+j)| \right)^2 \leq \beta \Rightarrow F_t(x, y) = 0 \end{aligned} \quad (3.26)$$

La función de densidad de probabilidad f_{dp} de una diferencia al cuadrado modelada por distribuciones Gaussianas (como es el caso de la función $g(x, y)$ de la fórmula anterior 3.27), se puede demostrar que sigue una distribución Chi-cuadrado (χ_q^2) con q grados de libertad (donde q es el tamaño de la ventana alrededor del píxel que estamos analizando, es decir, $2W + 1$):

$$f_{dp} = \sum_{i=1}^{i=q} D^2 \cong \chi_q^2 \quad (3.27)$$

Inicialmente se supone que no existe ningún objeto moviéndose en la imagen diferencia (hipótesis H) y que el cambio de valor de un píxel es debido al ruido de la cámara (considerado aditivo y que sigue una distribución Gaussiana con varianza σ). La función de densidad condicional está definida por:

$$f(g(x, y) / H) = \frac{1}{2^{\frac{q}{2}} \sigma^2 \Gamma(\frac{q}{2})} g(x, y)^{\frac{(q-2)}{2}} e^{-\frac{g(x, y)}{2\sigma^2}} \quad (3.28)$$

Donde $g(x, y)$ es la suma de diferencias al cuadrado entre la imagen y el fondo, $\Gamma(x)$ es la función gamma, σ es la varianza del ruido y q el tamaño de la ventana sobre el píxel de estudio (x, y) .

A partir de esta función de densidad de probabilidad se puede derivar en la siguiente probabilidad:

$$P(g(x, y) \leq \tau(x, y) / H) = \frac{\Gamma(\frac{q}{2}, \frac{g(x, y)^2}{2\sigma^2})}{\Gamma(\frac{q}{2})} \Rightarrow F_t = 0$$

$$P(g(x, y) > \tau(x, y) / H) = \frac{\Gamma(\frac{q}{2}, \frac{g(x, y)^2}{2\sigma^2})}{\Gamma(\frac{q}{2})} \Rightarrow F_t = 1$$
(3.29)

En el que se establece para cada píxel un determinado umbral $\tau(x, y)$ (obtenido empíricamente con valores desde 10^{-2} hasta 10^{-6}). Si el valor de la probabilidad es menor que ese umbral, el píxel se asignará a un objeto en movimiento. En caso contrario el píxel se asignará al fondo.

Este método expuesto es sencillo de aplicar pero tiene el problema de que es bastante dependiente de la secuencia de imágenes y del ruido que introduce la cámara que queremos analizar.

Modelado y Actualización de fondo (BG)

El modelo de fondo se construye con una media y una varianza (como el modelo de fondo del método de la Gaussiana Simple).

Además, el método Gamma desarrollado incorpora los cambios en el fondo debidos a objetos estacionarios (nuevos objetos en el fondo denominados también fantasmas) mediante un contador. Estos cambios se guardan en la variable que se ha denominado fondo antiguo B_t antiguo. La incorporación al fondo antiguo (B_t antiguo) se realiza utilizando un mecanismo de actualización selectiva de los píxeles de fondo por medio del contador (*contador*), que marca los píxeles detectados como frente en imágenes consecutivas. Cuando el contador supera un determinado límite M , se actualiza el fondo antiguo (B_t antiguo) con el valor de la imagen actual I_t en los píxeles detectados como objetos estacionarios F_t y con los valores del fondo en los píxeles detectados como fondo $not(F_t)$.

$$F_t = |I_t - B_t| > K \cdot \sigma$$

$$contador > M \Rightarrow B_{t,antiguo} = I_t(F_t) + B_t(not(F_t)) \quad (3.30)$$

Para actualizar la imagen de fondo a los cambios recientes se utiliza una media móvil:

$$B_{t,nuevo} = \alpha I_t + (1 - \alpha) B_t \quad (3.31)$$

El resultado de estas dos actualizaciones da lugar al modelo de fondo:

$$B_t = B_{t,nuevo} + B_{t,antiguo} \quad (3.32)$$

La varianza se actualiza con una media móvil igual a la utilizada en la Gaussiana simple:

$$\sigma_{t+1}^2(x, y) = \alpha (I_t(x, y) - \mu_t(x, y))^2 + (1 - \alpha) \sigma_t^2(x, y) \quad (3.33)$$

Estimación de FG

Por cada píxel se toma un bloque de píxeles de tamaño $W \times W$ entorno a él y se calcula la media de la diferencia al cuadrado entre la imagen actual y el modelo de fondo estimado en dicho $g_t(x, y)$.

$$g_t(x, y) = \sum_{i=x-\frac{W}{2}}^{i=x+\frac{W}{2}} \sum_{j=y-\frac{W}{2}}^{j=y+\frac{W}{2}} (I_t(i, j) - B_t(i, j))^2 \quad (3.34)$$

Para cada píxel del bloque diferencia se clasifican como píxeles de frente aquellos cuya probabilidad supere un umbral τ :

$$P(g_t(x, y) > \tau / H) \Rightarrow F_t(x, y) = 0$$

$$P(g_t(x, y) \leq \tau / H) \Rightarrow F_t(x, y) = 1 \quad (3.35)$$

Este método presenta un alto porcentaje de aciertos en secuencias con ruido impulsivo gracias al promediando del valor de cada píxel con los valores de los píxeles de la ventana de alrededor. El único inconveniente es el ajuste de los parámetros media, varianza y umbral óptimos para la secuencia.

3.3 Métodos no paramétricos

En esta sección se describe el marco teórico del método no paramétrico de modelado de fondo y proceso de sustracción de frente desarrollado en este proyecto. Este método se denomina Estimación de Densidad del Núcleo (KDE). El algoritmo utiliza el valor exacto de los píxeles como característica básica de modelado de fondo; manteniendo una muestra formada por un conjunto de valores de intensidad para cada píxel de la imagen y utilizando esta muestra para estimar la función de distribución de probabilidad de la intensidad de los píxeles. Por lo tanto, el modelo es capaz de estimar la probabilidad de cualquier valor de la intensidad observada recientemente. Este método puede manejar situaciones en que el fondo de la escena no es completamente estático, por ejemplo, debido a pequeños movimientos de las ramas de árboles y arbustos. Por otro lado, se actualiza continuamente y, por tanto, se adapta a los cambios de la escena del fondo.

3.3.1 Estimación de Densidad del Núcleo (KDE)

El modelo de fondo se genera a través de una muestra de L imágenes de la secuencia de los valores de cada píxel $(x_{t-L}, x_{t-L+1}, \dots, x_{t-1}, x_t)$. Tras almacenar la muestra, se obtiene una estimación de densidad del píxel $\Pr(x_t)$ aplicando una función de estimación de densidad de núcleo K . La probabilidad que posee el píxel x_t de pertenecer al fondo en cada instante t se calcula como:

$$\Pr(x_t) = \sum_{i=t-L}^{i=t-1} K(x_i - x_t) \quad (3.36)$$

Donde K es la función de densidad de núcleo. Si utilizamos una función de núcleo Gaussiana, la probabilidad se puede calcular de la siguiente forma:

$$\Pr(x_t) = \frac{1}{L} \sum_{i=t-1}^{i=t-1-L} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-x_t)^2}{2\sigma^2}} \quad (3.37)$$

Utilizando esta función de probabilidad, se considera que un píxel pertenece al frente si $\Pr(x_t)$ es menor que un umbral (este umbral depende de la secuencia y se puede ajustar para obtener un determinado porcentaje de falsos positivos). En la práctica, la estimación de la probabilidad se puede calcular de manera rápida empleando tablas pre-calculadas para los valores del núcleo, dada una diferencia $(x_t - x_i)$ y el ancho de banda del núcleo σ .

Además, realizar una suma de la evaluación parcial de cada píxel con sus L valores anteriores suele ser suficiente para superar el umbral que se impone en la mayoría de los píxeles de la imagen, ya que la mayor parte de la imagen es fondo. Lo cual permite realizar aplicaciones muy rápidas.

Dado que la densidad del núcleo es una estimación general, ésta puede converger a cualquier valor de la función de densidad de los píxeles. Por este motivo, la estimación se basa en las muestras más recientes de cada píxel, de modo que la adaptación del modelo puede lograrse únicamente mediante la adición de nuevas muestras y haciendo caso omiso a las anteriores.

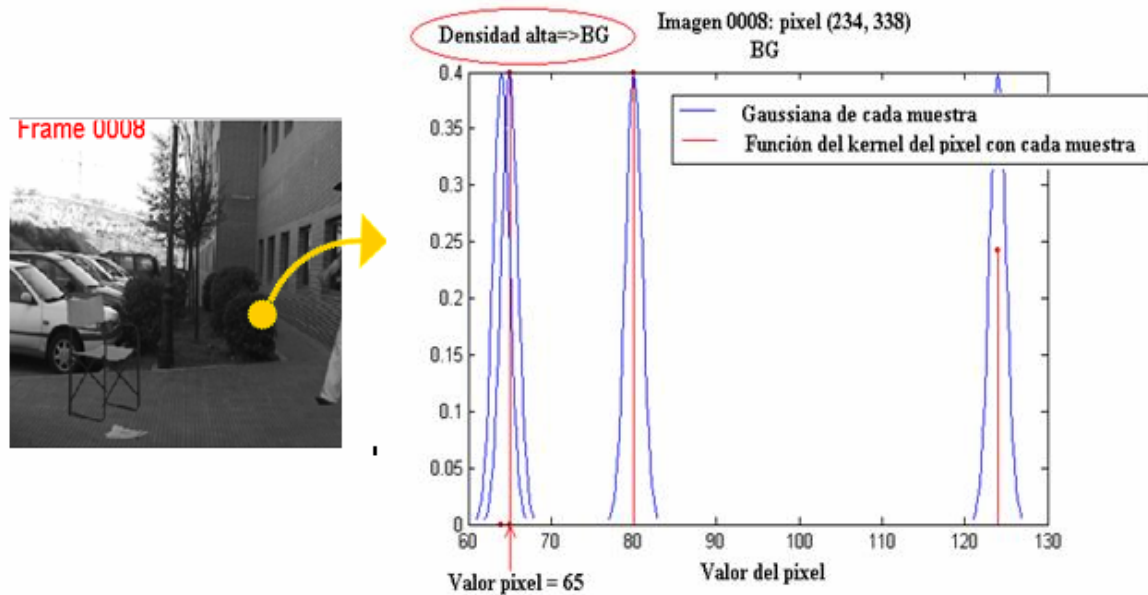


Figura 3-11: Representación de las Funciones de Núcleo Gaussianas centradas en valores anteriores y evaluadas en el valor actual del píxel localizado en la posición (234, 338).

Para seleccionar el ancho de banda de la función o desviación σ , se asume que la distribución de los píxeles en el tiempo es Gaussiana $N(\mu, \sigma)$, entonces la distribución de la desviación de

diferencias entre el píxel actual y los píxeles almacenados $(x_i - x_t)$ es también Gaussiana $N(0, 2\sigma)$. Como se puede demostrar, la distribución $N(0, 2\sigma)$ es simétrica, y por ello, la probabilidad de que dicha distribución sea mayor a la mediana (med) de las diferencias entre el píxel actual y los almacenados en el buffer es igual a un cuarto. Es decir:

$$\Pr\left(N\left(0, 2\sigma^2\right) > med\right) = \frac{1}{4} \quad (3.38)$$

Donde la mediana de diferencias entre el píxel actual y los almacenados en el buffer es:

$$med(x_t) = \text{mediana}_{i=1..L-1} \left(\left| I_{t-i}(x_t) - I_t(x_t) \right| \right) \quad (3.39)$$

Y, por lo tanto, la desviación estándar de la primera distribución puede estimarse como:

$$\sigma = \frac{med}{0.68\sqrt{2}} \quad (3.40)$$

Finalmente, los píxeles que pertenecen al fondo serán aquellos cuya densidad sea mayor que un umbral global τ . En caso contrario serán parte del frente F_t :

$$\begin{aligned} \Pr(x_t) > \tau &\Rightarrow F_t(x_t) = 0 \\ \Pr(x_t) \leq \tau &\Rightarrow F_t(x_t) = 1 \end{aligned} \quad (3.41)$$

A continuación, se explicarán paso a paso las operaciones realizadas para implementar este método:

Inicialización

El algoritmo comienza su ejecución desde la imagen tercera de la secuencia ($t = 3$), debido a que el cálculo de la varianza requiere las dos imágenes previas a la procesada. Desde dicho instante, se leen y almacenan en un buffer los píxeles y se calcula la distribución de cada píxel $\Pr(x_t)$ con los valores de los píxeles almacenados en el buffer (si aún no se han procesado L imágenes, la $\Pr(x_t)$ se calcula con los píxeles almacenados que serán menos de L).

Estimación del FG

La estimación de los píxeles que pertenecen al objeto en movimiento se realiza calculando la función de distribución de cada píxel $Pr(x_t)$, aplicando las fórmulas descritas anteriormente. Se estiman los píxeles del frente si la probabilidad no supera un cierto umbral y, en caso contrario, se consideran píxeles del fondo.

Actualización del BG

El modelo de fondo en este algoritmo es una cola FIFO.

La actualización se lleva a cabo almacenando en el buffer los L píxeles de las imágenes anteriores al píxel que se acaba de procesar. Para ello, en cada instante se elimina de la primera posición del buffer la muestra relativa a la imagen $L+1$ anterior y se introduce en la última posición la primera muestra que se acaba de procesar, de modo que el buffer contiene siempre L muestras de cada píxel con las que calcular su distribución, excepto en las L primeras imágenes procesadas en las que el buffer por primera vez se encuentra vacío.

Imagen de la secuencia en $t=138$



Representación del Fondo de la Imagen en $t=138$

B_{11}	B_{12}	B_{13}	...	B_{1M}
B_{21}	B_{22}	B_{23}	...	B_{2M}
B_{31}	B_{32}	B_{33}	...	B_{3M}
...
B_{K1}	B_{K2}	B_{K3}	...	B_{KM}

Figura 3-12: Representación de la imagen procesada y el modelo de fondo en el instante 138.

La imagen de fondo contiene por cada píxel (x, y) un buffer B donde se almacenan los L valores anteriores al instante actual t .

Por ejemplo, en el instante 138 en el buffer del píxel $(2,3)$ se introduce el valor actual y se elimina el del instante actual menos L .

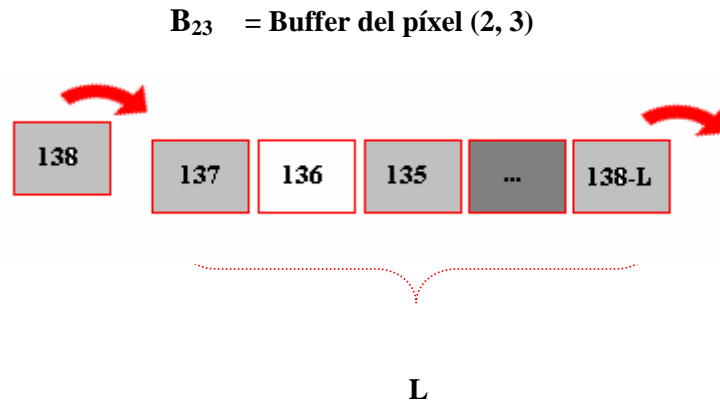


Figura 3-13: Buffer del píxel (2, 3) que contiene los L valores anteriores al instante de procesamiento 138.

3.3.1.1 Mejora del KDE: Filtro de análisis espacial y temporal.

En entornos de exteriores, existen numerosas fluctuaciones en el fondo que producen falsas detecciones de primer plano. En primer lugar, existen falsas detecciones debido al ruido aleatorio que se espera que sea homogéneo en toda la imagen y durante toda la secuencia. En segundo lugar, las falsas detecciones pueden producirse debido a pequeños movimientos en el escenario de fondo que no están representados por el modelo de fondo. Esto puede ocurrir localmente, por ejemplo, si una rama del árbol se mueve más allá de lo aceptado por el modelo de fondo, o bien, si se producen movimientos de cámara. Este tipo de falsas detecciones generalmente son agrupadas espacialmente en la imagen, y no son fáciles de eliminar con las técnicas morfológicas o los filtros de ruido, ya que estas operaciones pueden afectar también a la detección de pequeños objetivos y/o oclusiones.

Si una parte del fondo (por ejemplo, una rama del árbol) se mueve y ocupa un nuevo píxel que no forma parte del modelo de fondo, entonces será detectado como objeto en movimiento. Sin embargo, este objeto tendrá una alta probabilidad de pertenecer al fondo por tener un gran parecido a la distribución del píxel original.

Suponiendo que un pequeño desplazamiento puede ocurrir entre píxeles consecutivos, se decidirá si el píxel detectado es un objeto en movimiento o parte del fondo teniendo en cuenta la distribución del fondo de un conjunto de píxeles a su alrededor.

Siendo x_t el valor observado del píxel x en el instante t , se define la probabilidad de desplazamiento del píxel $P_N(x_t)$ como la máxima (max) probabilidad de que el valor observado pertenezca a la distribución de fondo en el vecindario N donde se encuentran los píxeles de fondo:

$$P_N(x_t) = \max_{y \in N(x_t)} (\Pr(x_t / B_y)) \quad (3.42)$$

Donde y es cada píxel del vecindario $N(x_t)$ y B_y la distribución del modelo de fondo del píxel y .

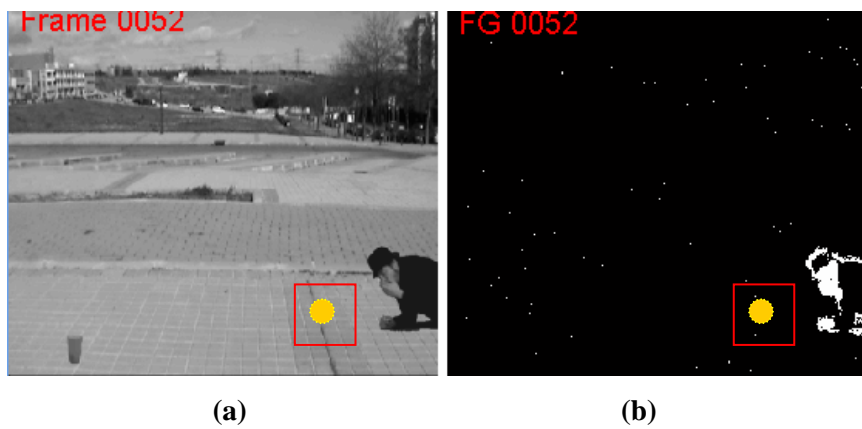


Imagen 0052: píxel (233, 326) es fondo

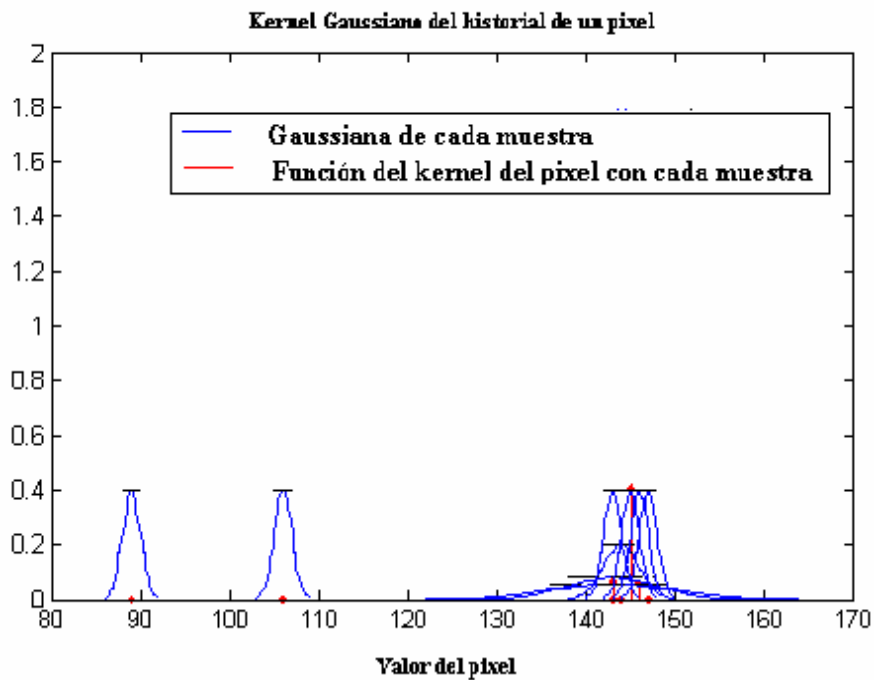


Figura 3-14: Modelo de KDE con filtrado espacio-temporal en la secuencia Hombre hambriento.

Las imágenes (a) y (b) de la **Figura 3-14** son una representación de la imagen y la máscara de frente obtenida tras el filtrado espacial en la secuencia “Hombre hambriento” en el instante 52. Abajo, se encuentra una gráfica del modelo de fondo KDE formado por las Gaussianas cuyas medias son los valores del píxel (233, 326) de las 10 imágenes anteriores a la actual.

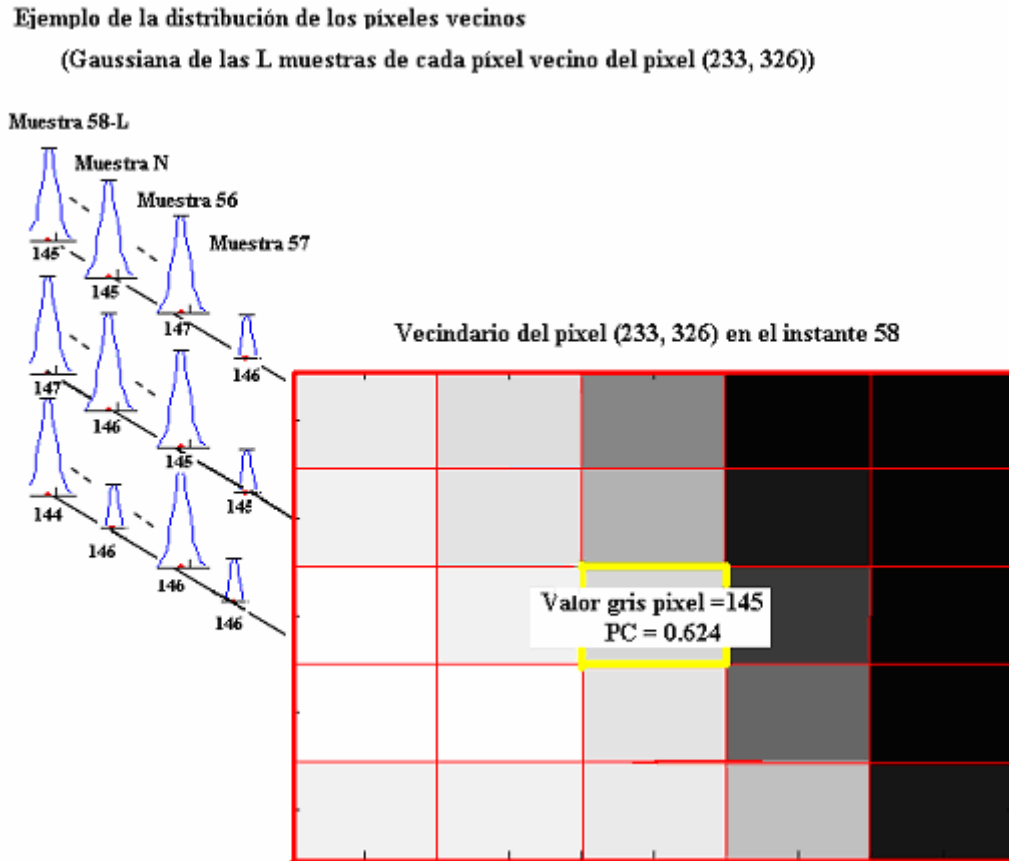


Figura 3-15: Vecindario 5 x 5 del píxel (233, 326) en el instante 58.

A la derecha de esta **Figura 3-15** se encuentra una representación de los valores de la imagen en el vecindario 5 x 5 en torno al píxel (233, 326) en el instante 58, así como la probabilidad máxima de parecido (P_C) que se explicará más adelante, con la distribución media de cada uno de sus vecinos. A la izquierda, un ejemplo del tipo de distribución que siguen algunos de sus vecinos compuesta por las Gaussianas centradas en las L muestras anteriores al instante actual.

La probabilidad de detección $\Pr(x_t/B_y)$ se calcula utilizando el núcleo K como función de estimación. Con el umbral P_N se consigue eliminar falsas detecciones procedentes de pequeños movimientos del fondo de la escena. Para evitar la pérdida de detecciones correctas que podrían ser similares al fondo (por ejemplo, objetos camuflados) se añade una limitación a los objetos

detectados en movimiento que consiste en realizar el desplazamiento en un vecindario cercano al píxel detectado. La probabilidad de desplazamiento a un lugar cercano (P_C) se define como:

$$P_C(x_t) = \prod_{x_i \in C} (P_N(x_i)) \quad (3.43)$$

Un píxel se considera parte del fondo solo si cumple:

$$\begin{aligned} P_C(x_t) &> \text{Umbral } 1 \\ P_N(x_t) &> \text{Umbral } 2 \end{aligned} \quad (3.44)$$

3.3.1.2 Mejora del método KDE: Combinación de modelos a corto y largo plazo.

Para lograr un modelo que se adapte rápidamente a los cambios en la escena y apoyar la sensibilidad en la detección así como disminuir los falsos positivos, Elgammal propone combinar los resultados de dos modelos de fondo: uno a largo plazo y otro a corto.

Modelo a corto plazo: Se trata de un modelo de la historia reciente de la escena. Se adapta a los cambios rápidamente para incrementar la sensibilidad de la detección. Este modelo consiste en mantener los últimos L valores de las muestras de la imagen. Estas muestras se actualizan de forma selectiva utilizando una máscara de valor 1 si el píxel debe ser actualizado en el instante t y 0 en caso contrario. A través del modelo a corto plazo se prevé que habrá dos tipos de falsos positivos: los debidos a eventos que no están representados en el modelo y la persistencia de falsos positivos debido a una detección incorrecta o la persistencia de los falsos positivos debidos a cambios en la escena del fondo.

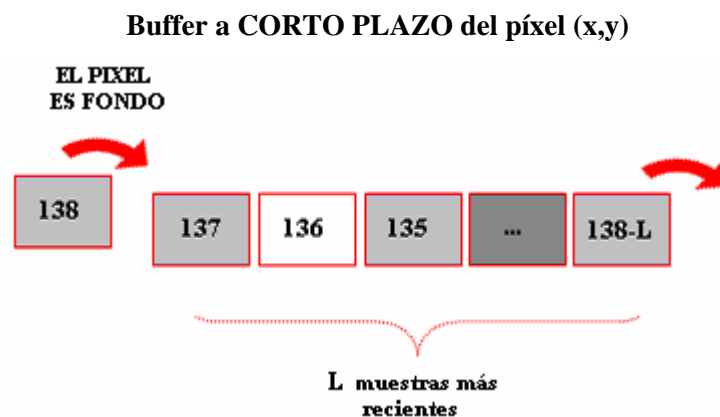


Figura 3-16: Buffer del píxel (x, y) que contiene los L valores anteriores al instante de procesamiento.

Método a largo plazo: Este modelo capta una representación más estable de la escena de fondo y se adapta a los cambios lentamente. Consiste en una cola de L muestras tomadas en una ventana de tiempo más grande $W \gg L$, es decir, de forma no consecutiva en el tiempo. La muestra se actualiza siguiendo un mecanismo ciego de actualización, de modo que cada nueva muestra, se añade al modelo independientemente de la clasificación. Este modelo se espera que tenga más falsos positivos, ya que no es el modelo más reciente de los antecedentes, y más falsos negativos debido a los píxeles de objeto que podrían incluirse en la muestra.

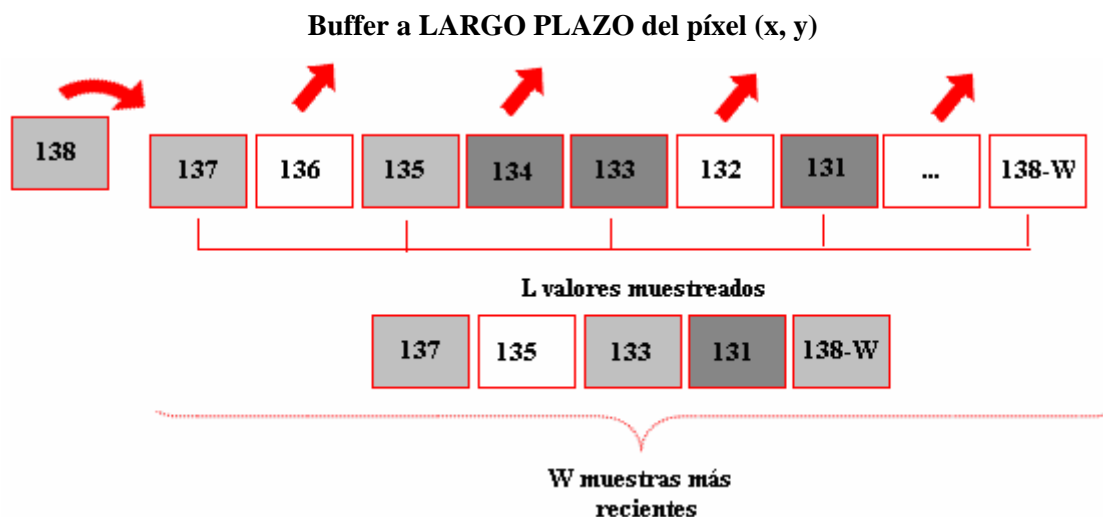


Figura 3-17: Buffer del píxel (x, y) que contiene los L valores anteriores al instante de procesamiento en una ventana W ($W \gg L$).

3.4 Método combinado: Histogramas y Gaussiana

Actualmente, existe cierta tendencia a agrupar y combinar distintos algoritmos de segmentación de objetos en movimiento para cubrir las carencias que tienen unos gracias a las ventajas que proporcionan otros. Utilizando esta aproximación, se han realizado diversas combinaciones de los algoritmos descritos anteriormente con el fin de evaluar las ventajas y desventajas introducidas por esta aproximación.

En este apartado, explicaremos un algoritmo de segmentación de objetos en movimiento que se ha realizado siguiendo la idea que propone E. Monari y C. Pasqual en [92] en el que se combinan dos de los métodos ya estudiados:

- **Corto plazo:** Gaussiana Simple.
- **Largo plazo:** Generación basada en histogramas.

En nuestro algoritmo, hemos elegido como modelo a corto plazo una Gaussiana simple ya que es muy rápido y en escenarios con pequeños movimientos ofrece muy buenos resultados. No obstante, en entornos más complejos, con muchos tipos de movimientos en el fondo, sería necesario emplear un modelo de fondo de Mezcla de Gaussianas para modelar adecuadamente el fondo, por el contrario, este algoritmo requiere una gran potencia de cálculo. Por este motivo, hemos optado finalmente por implementar la Gaussiana Simple.

Como modelo a largo plazo, generamos un histograma para cada píxel de la imagen de fondo a partir de las N imágenes anteriores con el objetivo de describir la distribución de los píxeles en un periodo de tiempo elevado de la secuencia de video analizada.

Estimación del FG/BG Largo plazo: Generación basada en histogramas

El comportamiento de cada píxel es modelado con un histograma que describe la distribución de sus valores en las últimas N imágenes; es decir, el tamaño del buffer de modelo de fondo que determina el tamaño del histograma es N . Los píxeles que pertenezcan a fondos estáticos presentarán un histograma con muy poca variación; los que sean de fondos muy variables, tendrán un histograma con mucha variación.

El píxel que sea máximo local en el histograma será el valor más frecuente en la secuencia. Se clasificarán los píxeles de fondo como aquellos que superen un umbral fijado en un porcentaje del valor del máximo local, y serán parte del frente, aquellos que tengan un valor inferior.

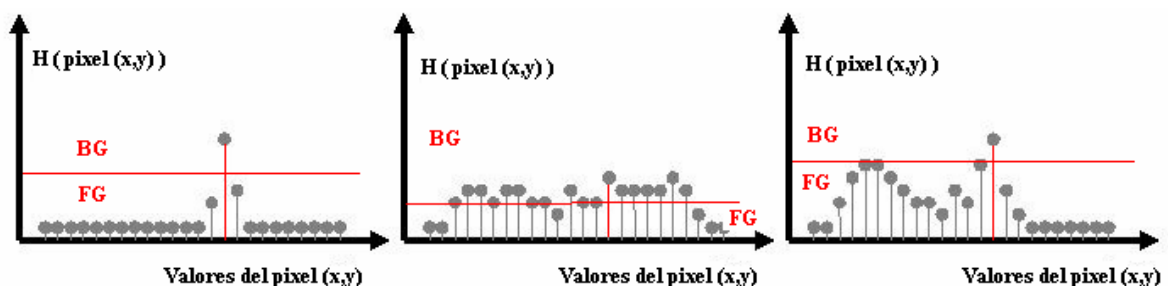


Figura 3-18: Ejemplos de los histogramas de un píxel: a) fondo estático, b) fondo pequeñas variaciones de intensidad. c) fondo dinámico.

El modelado por medio de histogramas presenta dos propiedades: por una parte, permite modelar distintas distribuciones del fondo a través de la elección del máximo local (por ejemplo, cambios de

intensidad). Por otro lado, es capaz de adaptar los cambios de iluminación generando la función de distribución de los píxeles de las N imágenes actuales.

Estimación del FG/BG Corto plazo: Gaussiana simple.

En el método a corto plazo, cada píxel del fondo se representa con los parámetros de una Gaussiana Simple (media y varianza), los cuales se actualizan en el tiempo a través de una media móvil. Los resultados que se obtienen mediante este método se caracterizan por el bajo corte computacional y la rapidez, propiedades muy importantes en métodos a corto plazo ejecutados en tiempo real. Además, es adecuado para fondos de escenarios estáticos y con pequeñas variaciones. Por otro lado, se adapta a los cambios de iluminación en un periodo largo de ejecución incrementando la varianza de la Gaussiana que modela el fondo.

Fusión de modelos.

El método empleado para fusionar las máscaras binarias de frente calculadas por los dos algoritmos es utilizar una operación lógica que debe decidir entre cuatro posibles estados:

- Ambos métodos clasifiquen el píxel como frente.
- Sólo el método a corto plazo estime el píxel como frente.
- Sólo el método a largo plazo estime el píxel como frente.
- Ambas métodos clasifiquen el píxel como fondo.

En la primera y última situación el resultado es claro: si ambos métodos coinciden en sus resultados el píxel se clasifica como hayan decidido los dos métodos. En los otros dos casos, los métodos difieren en sus estimaciones y nuestro algoritmo decide que el píxel sea clasificado como la decisión tomada por la Gaussiana por ser un método más robusto que la Generación de Histogramas.

4 Integración en una arquitectura para su evaluación

En este proyecto se ha desarrollado un sistema modular que permite integrar y evaluar los distintos esquemas de modelado de fondo descritos en el capítulo 3 de una manera sencilla. Este sistema se utiliza para los siguientes propósitos: ejecución de secuencias de video (almacenadas como secuencias de imágenes), evaluación de parámetros óptimos de funcionamiento de los algoritmos y comparación entre los distintos algoritmos.

La arquitectura del sistema está compuesta por dos grandes módulos que realizan distintas tareas: el primero es el modelado del fondo y extracción del frente; y el segundo, la validación de los resultados.

1. Módulo de modelado y extracción del frente: El objetivo de este módulo es la generación de las imágenes de fondo y frente. Para ello, realiza las siguientes tareas:

- 1.1. Lectura de la imagen.
- 1.2. Almacenamiento de la imagen en un buffer de datos.
- 1.3. Fijación de los parámetros del algoritmo.
- 1.4. Inicialización del fondo.
- 1.5. Modelado y actualización del fondo.
- 1.6. Calculo de tiempo de procesamiento del modelado de fondo.
- 1.7. Extracción de la máscara de frente.
- 1.8. Almacenamiento de la máscara de frente.
- 1.9. Representación de la imagen, el fondo y la máscara del frente de la secuencia procesada en cada instante.
- 1.10. Representación gráfica de los modelos de fondo.

2. Modulo de validación de los resultados: Este módulo posee dos objetivos: durante la primera etapa del proyecto, se utiliza para evaluar los parámetros más adecuados de cada método; y, durante una segunda etapa el proyecto, sirve para comparar los métodos implementados. Realiza las siguientes tareas:

- 2.1. Lectura de la máscara binaria del frente obtenido en la etapa anterior y de la máscara del frente anotada ('ground truth').
- 2.2. Cálculo de medidas de evaluación.
- 2.3. Cálculo de porcentajes de detección correcta/incorrecta a nivel de píxel.

- 2.4. Cálculo las medidas de evaluación más elaboradas: precisión, especificidad, 'f-measure', sensibilidad ('recall') y curvas ROC.
- 2.5. Elección de los parámetros óptimos.
- 2.6. Almacenamiento de todas las medidas en tablas Matlab.
- 2.7. Apertura de un fichero secuencia_metodo.txt, almacenamiento de las medidas en tablas en el fichero y cierre del fichero.
- 2.8. Representación gráfica de todas las medidas.

En el siguiente diagrama se muestra la arquitectura descrita anteriormente:

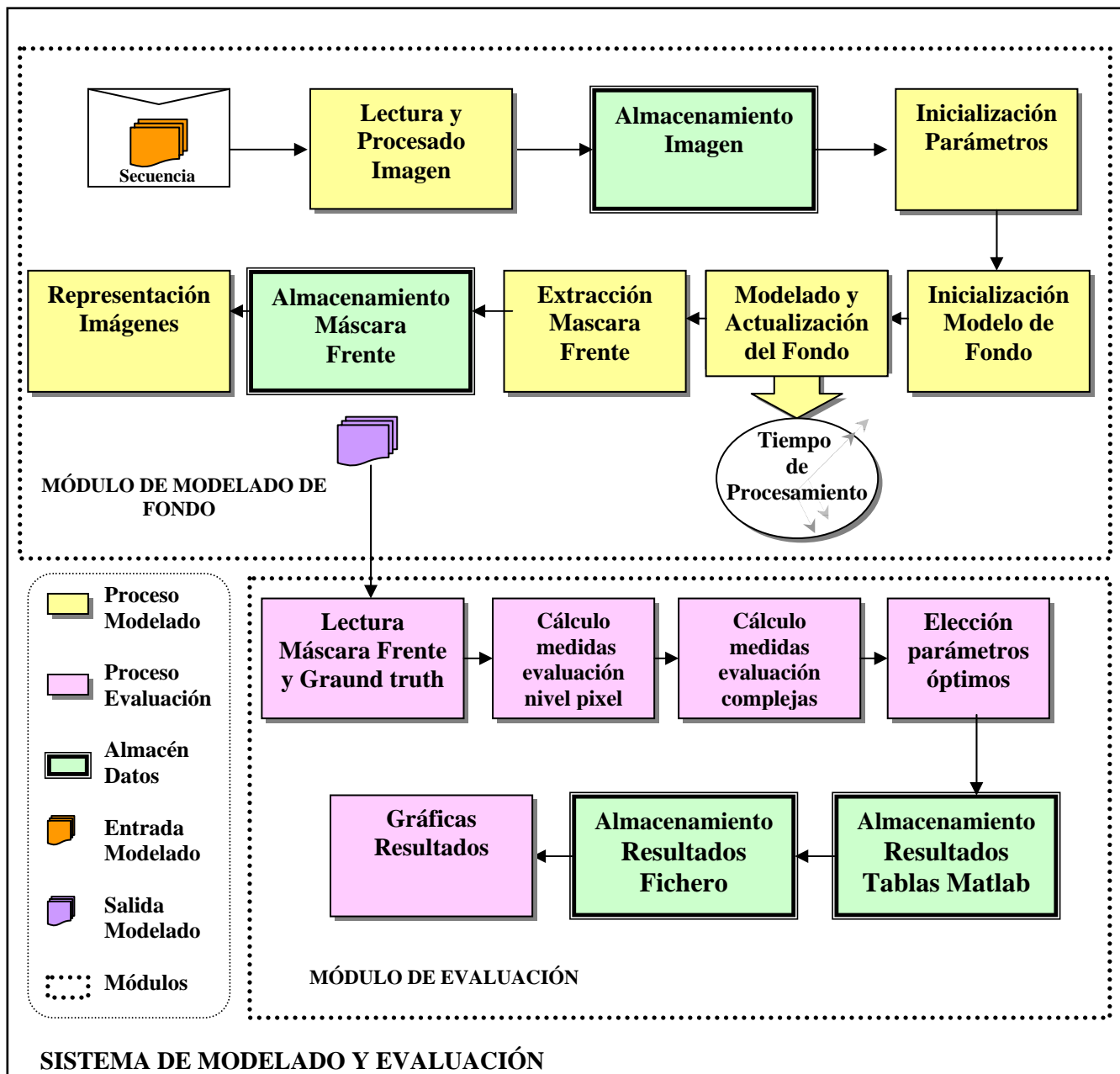


Figura 4-1: Arquitectura del Sistema de modelado de fondo y evaluación.

5 Resultados: pruebas y análisis comparativo

La segmentación de objetos en secuencias de video ha generado un gran interés en la comunidad científica en los últimos años. Debido a ello, se han desarrollado diversas estrategias para evaluar la calidad de los sistemas de segmentación propuestos. Las técnicas disponibles con este fin se pueden clasificar en: subjetivas [83] y objetivas [84][85]. Las técnicas subjetivas dependen del criterio humano de evaluación, por lo tanto, para aplicarse es necesario que se realicen sobre un número elevado de personas. Las objetivas, suelen utilizar una segmentación de referencia, que generalmente se anota a mano, denominada ‘*ground-truth*’, para compararla con la segmentación obtenida del algoritmo de segmentación implementado. Esta comparación se puede utilizar para calcular el número de píxeles erróneos y correctos detectados por el segmentador, así como para obtener muchas otras medidas propuestas en la literatura [84][85]. Estas medidas permiten evaluar la calidad de la segmentación.

Las medidas utilizadas en este proyecto para evaluar los algoritmos de segmentación son técnicas objetivas que se basan en una segmentación de referencia calculada para un conjunto de secuencias de video de características diferentes que constituyen el grupo de videos de prueba (‘*dataset*’) utilizados en el proyecto, grupo que explicaremos en el siguiente apartado.

5.1 Descripción del dataset utilizado

El conjunto de secuencias de videos de prueba que se ha utilizado para evaluar los algoritmos de segmentación fueron creados por el *Video Processing & Understanding Lab* de la Escuela Politécnica Superior de la UAM [86] y están disponibles en la siguiente página de Internet <http://www-vpu.ii.uam.es/CVSG/>.

Las secuencias de video utilizadas están clasificadas por su complejidad, es decir, por el grado de dificultad que presenta para un algoritmo de segmentación analizar dichas secuencias. Este grado de dificultad depende de una serie de propiedades denominadas ‘factores críticos’ entre las que se encuentran: las características específicas de los objetos, el tipo de fondo o el movimiento de la cámara (aunque en nuestro caso, se ha trabajado con secuencias de video sin movimiento de cámara).

El siguiente cuadro resume los factores críticos en relación con el primer plano y el fondo:

FRENTE (OBJETOS)					FONDO		
OBJETOS SIMPLES				GRUPOS		Textura	Modalidad
Textura	Velocidad	Estructura	Conectividad	Tamaño	Velocidad		

Tabla 5-1: Factores críticos del frente y fondo [86].

A continuación, se describen los factores críticos y el grado de complejidad que suponen en una secuencia de video.

5.1.1 Factores críticos del objeto

Las características del fondo y algunas de las propiedades de los objetos, si se trabaja en vecindades, influyen significativamente en la precisión de detección. Las propiedades de los objetos (como la textura y velocidad) que afectan a objetos completos no a píxeles aislados, no son importantes en los algoritmos de substracción de fondo basados en modelado de fondo donde se decide entre fondo y frente píxel a píxel. Las propiedades de los objetos y fondo más destacadas son las siguientes:

- *Complejidad en base a la textura de los objetos:* Sólo afecta a los algoritmos de estimación de movimiento, no a los de background-subtraction. Los objetos de alta textura son, en general, más fácilmente identificables que los de baja textura debido a que si un objeto posee alta textura, el parecido con el fondo es menor (la diferencia entre imagen y fondo es mayor. Por este motivo, el algoritmo de segmentación tiene menos complejidad a la hora de procesar secuencias con objetos de alta textura y, más complejidad si las secuencias poseen objetos de baja textura.
- *Complejidad vs. velocidad aparente de los objetos:* Los pequeños cambios temporales de los objetos fáciles de confundir con la inestabilidad y/o el ruido de la cámara son difíciles de discriminar, es decir, representan secuencias de video de alta complejidad. Si los objetos

poseen una velocidad muy alta son difíciles de detectar por algoritmos de estimación de movimiento y exigen un análisis de movimiento de mayor complejidad.

- *Complejidad vs. estructura del objeto:* Cuando el objeto en movimiento puede degradarse debido a la inmovilidad de algunas de sus partes resulta complicado detectar el objeto completo en movimiento (por ejemplo, los pies de una persona al caminar permanecen estáticos mientras impulsa su cuerpo hacia adelante). Los objetos rígidos son menos complejos de detectar que lo que no lo son.
- *Complejidad vs. conectividad:* En los algoritmos de estimación de regiones, la falta de conectividad en regiones y puntos del objeto en movimiento hace que su detección sea más difícil. Este hecho ocasiona una disminución en la precisión así como un incremento en la complejidad en los algoritmos que utilicen la evolución temporal para detectar al objeto. Esta característica no afecta a los algoritmos de background-subtraction
- *Complejidad vs. tamaño del objeto:* Sólo debe ser considerado en algoritmos que procesen secuencias de video con movimiento de cámara (no es el caso de este PFC).

En relación con los conjuntos de objetos existen las siguientes características:



- *Complejidad vs. velocidad de las imágenes:* Esta característica no afecta a los algoritmos de sustracción de fondo. En un algoritmo de estimación de movimiento que procese secuencias cuyas imágenes tengan objetos con velocidades diferentes (lentas y rápidas) será más difícil establecer un umbral para la detección de objetos en movimiento.
- *Complejidad vs. interacciones entre objetos:* Esta es una de las características que puede afectar a los algoritmos de sustracción de fondo. Si los objetos de la secuencia de video se unen o cruzan resulta a veces difícil detectarlos. Por este motivo, este tipo de secuencias son consideradas de alta complejidad.

5.1.2 Factores críticos del fondo

Estos factores se refieren a las propiedades del fondo que pueden influir en la detección del objeto en movimiento. Estas propiedades son las siguientes:

- *Complejidad vs. textura del fondo:* La textura del fondo tampoco afecta los algoritmos de sustracción de fondo. Al igual que ocurriría con la textura del frente, si el fondo tiene una textura alta (fondo no uniforme), los algoritmos de segmentación que empleen el cambio temporal para detectar los objetos en movimiento serán más complejos que aquellos cuya textura sea baja (fondo uniforme).
- *Complejidad vs. modos del fondo:* El modo del fondo se refiere al tipo (o modo) de movimiento que existe en el fondo. Si el fondo carece de movimiento salvo el debido al ruido introducido por la cámara, se puede representar por un tipo de movimiento y se denomina fondo unimodal. Sin embargo, si en el fondo existen varios tipos de movimiento (por ejemplo, debido al movimiento de árboles) se dice que es multimodal. A medida que aumentan los tipos de movimiento en el fondo, se incrementa la complejidad para detectar los objetos en movimiento.

Los videos de prueba utilizados para evaluar los algoritmos de segmentación implementados en este proyecto, se han clasificado en base a las características descritas anteriormente y que influyen en los algoritmos de sustracción de fondo implementados. Además, se utilizará un identificador (S1, S2,..) para agrupar a cada conjunto de secuencias de características similares (con el mismo modelo de fondo). Los videos utilizados para evaluar los métodos de segmentación desarrollados son los siguientes:

<i>SECUENCIA</i> (Imagen 720x576)	<i>Identificador del conjunto de secuencias</i>	<i>Situación (Interior o exterior)</i>	<i>Modalidad Fondo</i>
<p>BAILE 1</p>  <p>Nº imágenes = 750</p>	S1	Interior	Unimodal
<p>BAILE 2</p>  <p>Nº imágenes = 750</p>	S2	Exterior	Multimodal(alto)

<p>HOMBRE CALUROSO</p>  <p>Nº imágenes = 749</p>	S2	Exterior	Multimodal(alto)
<p>NIÑA PELUCHE</p>  <p>Nº imágenes = 619</p>	S2	Exterior	Multimodal(alto)
<p>INGRAVIDEZ</p>  <p>Nº imágenes = 671</p>	S3	Exterior	Multimodal(alto)
<p>TENIS 1</p>  <p>Nº imágenes = 307</p>	S4	Exterior	Multimodal(medio)
<p>TENIS 2</p>  <p>Nº imágenes = 691</p>	S4	Exterior	Multimodal(medio)
<p>SILLA</p>  <p>Nº imágenes = 749</p>	S5	Exterior	Multimodal(medio)



HAMBRE APREMIA  N° imágenes = 749	S6	Exterior	Unimodal
BENDER  N° imágenes = 732	S7	Exterior	Unimodal

Tabla 5-2: Características de las secuencias de video de prueba ('dataset')

5.2 Diseño de pruebas experimentales

Con el objetivo de obtener los parámetros de funcionamiento más adecuados de cada algoritmo, se han estudiado y ejecutado las pruebas experimentales

Por otro lado, se han comparado los parámetros y los resultados obtenidos por los métodos de segmentación de objetos implementados en este PFC, con los mismos métodos realizados por otros autores sobre secuencias de video de características similares a las que se encuentran en el conjunto de secuencias de video de prueba de este proyecto.

Para desarrollar estas pruebas experimentales, se ha supuesto que los parámetros de cada algoritmo son independientes. Para cada parámetro se ha realizado una serie de pruebas con un conjunto de valores predefinidos (rango de prueba), fijando el resto de parámetros a valores coherentes según las características de las secuencias, pruebas preliminares y los informes de otros autores.

Los rangos de prueba se han escogido en función de los factores críticos (descritos en el apartado anterior 5.1) de las secuencias; por ejemplo, las secuencias de fondo multimodal exigen mayor variación en el fondo y, por tanto, varianzas (σ^2) más elevadas que aquellas secuencias cuyo fondo sólo tenga un tipo de movimiento. Por otro lado, las secuencias con objetos en movimiento lento requieren un factor de aprendizaje (α) más bajo que aquellas en las que los movimientos de los objetos sean rápidos. Adicionalmente, se han comparado los rangos de prueba seleccionados para los parámetros de cada método de modelado de fondo, con los tomados por otros autores en

los métodos que han implementado se sustracción de fondo [87][88][89]. Aplicando los rangos de prueba a cada parámetro, se han ido encontrando los parámetros más adecuados de cada método de modelado de fondo.

A continuación, se muestra una tabla que recoge para cada algoritmo los parámetros más representativos y los rangos de prueba con los que se han realizado las ejecuciones sobre las secuencias del apartado anterior.

Los rangos de prueba están representados por dos tipos de vectores expresados en notación Matlab:

- El primer tipo de vector utilizado para representar el rango de los parámetros está formado por tres números separados por dos puntos: el primer número del vector indica el valor inicial del rango, el segundo el paso, y el último, el valor final del rango (por ejemplo, el rango de pruebas del parámetro Umbral es [20: 20: 160], lo que quiere decir, que el rango con el que se prueba el parámetro Umbral comienza con el valor 20, prosigue en valores de 20 en 20 hasta finalizar con el valor 160).

- El segundo tipo de vector, está formado por un conjunto de valores separados por una coma que representan directamente los valores de prueba concretos (por ejemplo, el rango de pruebas del parámetro Alpha es [10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}], lo cual significa, que el parámetro Alpha se ejecuta con cada uno de los valores indicados en el vector).

Como se observa en la tabla, existen varios tipos de vectores de prueba para cada parámetro de cada algoritmo. La utilización de estos vectores depende de las formulas del propio algoritmo y de las propiedades (factores críticos) de cada secuencia de prueba.

Algoritmo de modelado de fondo	Parámetros más representativos	Rango de prueba
DIFERENCIA	Umbral (τ)	[20: 20: 160], [2: 2: 20], [20: 20: 360], [50: 50: 800]
MEDIANA	Umbral (τ)	[10: 10: 90]
HISTOGRAMAS	Umbral (U)	[0.75:0.05:0.95]
GAUSSIANA SIMPLE	Alpha (α)	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}], [10^{-8} , $5 \cdot 10^{-8}$, 10^{-7} , $5 \cdot 10^{-7}$, 10^{-6} , $5 \cdot 10^{-6}$, 10^{-5}]
	Varianza (σ^2)	[1: 3: 18], [1: 5: 30], [35: 5: 65], [20: 5: 50]
MEZCLA DE GAUSSIANAS	Alpha (α)	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}], [10^{-8} , $5 \cdot 10^{-8}$, 10^{-7} , $5 \cdot 10^{-7}$, 10^{-6} , $5 \cdot 10^{-6}$, 10^{-5}]
	Varianza (σ^2)	[2: 2: 18], [24:2:34], [20: 5: 50], [10:5:40],
	Umbral de BG (B)	[0.75:0.05:0.95]

GAMMA	Umbral (τ)	[10^{-7} , 10^{-5} , 10^{-3}]
KDE	Prob. de Kernel (Pr)	[0.25 : 0.05: 0.5]
	Nº de imágenes (L)	[20, 40, 80]
COMBINADO	Umbral (U)	[0.75:0.05:0.95]
	Alpha (α)	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}], [10^{-8} , $5 \cdot 10^{-8}$, 10^{-7} , $5 \cdot 10^{-7}$, 10^{-6} , $5 \cdot 10^{-6}$, 10^{-5}]
	Varianza (σ^2)	[1: 3: 18], [1: 5: 30], [35: 5: 65], [20: 5: 50]

Tabla 5-3: Vectores de prueba para cada parámetro de cada algoritmo.

5.3 Métricas utilizadas

Las métricas que se han utilizado para evaluar los métodos de segmentación desarrollados en este PFC son medidas objetivas que se basan en la comparación de las máscaras binarias del frente generadas por cada método y el ‘ground-truth’. Aunque se han encontrado distintas medidas perceptuales en la literatura [82] finalmente se ha optado por las medidas objetivas a nivel de píxel que recordamos a continuación:

- **Verdaderos positivos (TP):** Es el número de detecciones correctas de valor uno (frente).
- **Verdaderos negativos (TN):** Es el número de detecciones correctas de valor cero (fondo).
- **Falsos positivos (FP):** Es el número de detecciones incorrectas de valor uno.
- **Falsos negativos (FN):** Es el número de detecciones incorrectas de valor cero.

En la siguiente tabla se muestran estas medidas y sus relaciones con la máscara del ‘ground truth’ y la máscara obtenida a la salida del algoritmo de segmentación.

		Ground truth		
		Positivos(unos)	Negativos(ceros)	
Algoritmo de Segmentación	Positivos(unos)	TP	FP	TP+FP
	Negativos(ceros)	FN	TN	FN+TN
		TP+FN	FP+TN	

Tabla 5-4: Parámetros estadísticos de la evaluación.

Las medidas anteriores suelen combinarse y dar lugar a otras muy usuales en el estado del arte de segmentación de objetos. Estas medidas son las siguientes:

- **Precisión:** Se define como el número total de píxeles correctos de un tipo con respecto al total de los detectados por nuestro algoritmo de segmentación de ese tipo. Las máscaras que se obtienen de la segmentación son binarias presentando píxeles con valor cero y valor uno. Por otro lado, nuestro objetivo a la hora de evaluar los algoritmos no es únicamente la correcta detección del objeto sino también, que los píxeles en ausencia de movimiento sean correctamente detectados, esto supone el cálculo de precisión en los dos grupos:

$$\mathbf{Precisión} (\text{píxeles_de_valor_0}) = P0 = \frac{TN}{TN + FN} \quad (5.1)$$

$$\mathbf{Precisión} (\text{píxeles_de_valor_1}) = P1 = \frac{TP}{TP + FP} \quad (5.2)$$

- **Recall:** Es el número de píxeles correctos detectados de un tipo con respecto al total real (que marcará el 'ground truth') de ese tipo. Por el mismo motivo que en el caso anterior, existen dos grupos de valores en las máscaras obtenidas a la salida de los segmentadores implementados (las dos con igual importancia) y, por ello, se ha calculado un 'recall' para cada conjunto:

$$\mathbf{Recall} (\text{píxeles_de_valor_0}) = R0 = \frac{TN}{TN + FP} \quad (5.3)$$

$$\mathbf{Recall} (\text{píxeles_de_valor_1}) = R1 = \frac{TP}{TP + FN} \quad (5.4)$$

- **Fscore:** Una forma de combinar la correcta detección del algoritmo con respecto a las detecciones realizadas por el mismo ('precisión') con la correcta detección del algoritmo con la realidad ('recall') es a través del *f-score*, que concede a ambas medidas igual peso. El *f-score* también ha sido calculado para cada conjunto de valores de la máscara binaria resultado del segmentador:

$$\mathbf{Fscore} (\text{píxeles_de_valor_0}) = FS0 = \frac{2 \cdot P0 \cdot R0}{P0 + R0} \quad (5.5)$$

$$\mathbf{Fscore} (\text{píxeles_de_valor_1}) = FS1 = \frac{2 \cdot P1 \cdot R1}{P1 + R1} \quad (5.6)$$

Para lograr el objetivo de evaluar y encontrar los parámetros óptimos de los algoritmos se han ponderado de manera equivalente la detección de frente/fondo en la máscara binaria (unos y ceros)

fusionando las medidas anteriores a través de una suma (si bien se podía haber utilizado cualquier otra función). La selección del parámetro óptimo se realiza eligiendo la suma máxima:

$$SUMA = FS0 + FS1 \quad (5.7)$$

5.4 Selección de parámetros

Cada algoritmo de segmentación de objetos ha sido ejecutado con los rangos de prueba adecuados a sus características (definidos en la tabla Tabla 5-3).

Los parámetros óptimos se han obtenido seleccionando la suma máxima de la medida *f-score* en ceros (5.5) y unos (5.6) calculada para cada rango de cada secuencia de prueba. Las tablas completas que se han obtenido de estas ejecuciones se hallan en el Anexo A.

El resultado de los parámetros óptimos de cada método se recoge en las siguientes tablas:

ID SECUENCIA	DIFERENCIA		MEDIANA		HISTOGRAMAS	
	RANGO	τ óptimo	RANGO	τ óptimo	RANGO	U óptimo
S1	[20:20:160]	140	[10:10:90]	20	[0.75:0.05:0.95]	0.75
S2	[50:50:800]	450	[10:10:90]	60	[0.75:0.05:0.95]	0.75
S3	[20:20:320]	180	[10:10:90]	20	[0.75:0.05:0.95]	0.75
S4	[20:20:160]	40	[10:10:90]	10	[0.75:0.05:0.95]	0.85
S5	[20:20:160]	40	[10:10:90]	10	[0.75:0.05:0.95]	0.85
S6	[20:20:160],[2:2:20]	20	[10:10:90]	10*	[0.75:0.05:0.95]	0.85
S7	[20:20:160]	40	[10:10:90]	10	[0.75:0.05:0.95]	0.9

Tabla 5-5. Parámetros óptimos para los métodos: diferencia de frentes, mediana e histogramas.

Los valores de los rangos de prueba elegidos para el umbral del método de la Diferencia de Frentes dependen de las características de la secuencia. Por ejemplo, secuencias de fondo complejo requieren umbrales mayores (como el umbral de 450 en la secuencia S2) ya que la diferencia al

cuadrado entre la imagen y el modelo de fondo que se utiliza para discriminar entre fondo y frente en este método, en tales secuencias, alcanza valores muy elevados. Sin embargo, secuencias de fondo unimodal requieren rangos menores puesto que la diferencia al cuadrado entre imagen y fondo es inferior.

Los parámetros óptimos del umbral para cada secuencia del método de la mediana, se han obtenido probando un único tipo de rango de valores de umbral ([10:10:90]) para todas las secuencias. Se ha escogido un rango bastante amplio, que ha permitido seleccionar umbrales bajos en secuencias de baja complejidad (baja varianza), y umbrales elevados en secuencias de alta complejidad (alta varianza).

En el caso de la generación de histogramas, se ha escogido un rango [0.75:0.005:0.95] para todas las secuencias ya que en el método se especifica que el umbral con el que se discrimina entre frente y fondo debe ser mayor que 0.75. Para determinar el límite superior del rango se han realizado varias pruebas y se ha observado que en ninguna de las secuencias el umbral óptimo superaba el valor de 0.9, por este motivo, se ha optado por el valor de 0.95 para el número final del rango. Este método se ha ejecutado con histogramas de tamaño 100 (100 valores para cada píxel) que se han actualizado cada 75 imágenes.

<i>GAUSSIANA SIMPLE</i>				
<i>ID SECUENCIA</i>	<i>RANGO ALPHA</i>	<i>ALPHA ÓPTIMO</i>	<i>RANGO DESVIACIÓN</i>	<i>DESVIACIÓN ÓPTIMA</i>
S1	[10 ⁻⁸ , 5 10 ⁻⁸ , 10 ⁻⁷ , 5 10 ⁻⁷ , 10 ⁻⁶ , 5 10 ⁻⁶ , 10 ⁻⁵]	1*10 ⁻⁶	[1:3:18]	6
S2	[10 ⁻⁴ , 5 10 ⁻⁴ , 10 ⁻³ , 5 10 ⁻³ , 10 ⁻² , 5 10 ⁻² , 10 ⁻¹]	1*10 ⁻²	[30:5:55]	35
S3	[10 ⁻⁴ , 5 10 ⁻⁴ , 10 ⁻³ , 5 10 ⁻³ , 10 ⁻² , 5 10 ⁻² , 10 ⁻¹]	1*10 ⁻³	[1:3:18]	10
S4	[10 ⁻⁴ , 5 10 ⁻⁴ , 10 ⁻³ , 5 10 ⁻³ , 10 ⁻² , 5 10 ⁻² , 10 ⁻¹]	5*10 ⁻²	[30:5:55]	35
S5	[10 ⁻⁴ , 5 10 ⁻⁴ , 10 ⁻³ , 5 10 ⁻³ , 10 ⁻² , 5 10 ⁻² , 10 ⁻¹]	5*10 ⁻⁴	[1:3:18]	5
S6	[10 ⁻⁸ , 5 10 ⁻⁸ , 10 ⁻⁷ , 5 10 ⁻⁷ , 10 ⁻⁶ , 5 10 ⁻⁶ , 10 ⁻⁵]	1*10 ⁻⁶	[1:3:18]	6
S7	[10 ⁻⁸ , 5 10 ⁻⁸ , 10 ⁻⁷ , 5 10 ⁻⁷ , 10 ⁻⁶ , 5 10 ⁻⁶ , 10 ⁻⁵]	1*10 ⁻⁵	[1:3:18]	5

Tabla 5-6. Parámetros óptimos para el método Gaussiana simple.

En el método de la Gaussiana Simple (en siglas inglesas, SG), para elegir los parámetros más adecuados se han realizado algunas pruebas iniciales, con objeto de determinar cuáles eran más sensibles, es decir, los parámetros que al cambiar levemente su valor alteraban en gran medida los resultados finales. Una vez concluida esta fase, se han probado las secuencias de prueba con los rangos más adecuados de los parámetros (según las características de las secuencias), de mayor a menor sensibilidad (primero el parámetro alpha y después la desviación). Uno de los parámetros que no mostramos en la tabla superior es el número de veces que se toma la desviación para determinar si los píxeles se encuentran en el interior de la Gaussiana y por tanto, pertenecen al fondo, ya que en todas las pruebas que se han realizado, los resultados óptimos se han obtenido en tres veces la varianza, que matemáticamente corresponde a los valores en los que la función Gaussiana posee valor mayor que cero.

<i>MEZCLA DE GAUSSIANAS</i>						
<i>SECUENCIA PRUEBA</i>	<i>RANGO PRUEBA ALPHA</i>	<i>ALPHA ÓPTIMO</i>	<i>RANGO DESVIACIÓN</i>	<i>DESVIACIÓN ÓPTIMA</i>	<i>RANGO UMBRAL</i>	<i>UMBRAL ÓPTIMO</i>
S1	[10^{-7} , $5 \cdot 10^{-7}$, 10^{-6} , $5 \cdot 10^{-6}$, 10^{-5} , $5 \cdot 10^{-5}$, 10^{-4}]	$5 \cdot 10^{-5}$	[24: 2: 34]	28	[0.75:0.05:0.95]	0.9
S2	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-2}$	[20: 5: 50]	40	[0.75:0.05:0.95]	0.85
S3	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$5 \cdot 10^{-3}$	[10: 5: 40]	30	[0.75:0.05:0.95]	0.9
S4	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$5 \cdot 10^{-3}$	[5: 5: 35]	15	[0.75:0.05:0.95]	0.95
S5	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-3}$	[5: 5: 35]	20	[0.75:0.05:0.95]	0.85
S6	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-4}$	[24: 2: 34]	32	[0.75:0.05:0.95]	0.85
S7	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-3}$	[2: 2: 18]	4	[0.75:0.05:0.95]	0.95

Tabla 5-7. Parámetros óptimos para el método mezcla de Gaussianas (MoG).

Una vez obtenidos los parámetros más adecuados de la SG, se ha ejecutado la MoG con una idea aproximada de los rangos en los que posiblemente se encontrarían los parámetros más adecuados de este método ya que aunque los métodos son diferentes y, por tanto, los parámetros óptimos también lo son, las secuencias de prueba son las mismas (con las mismas características) y la distribución con la que se modelan los píxeles en ambos métodos es similar (una Gaussiana en la SG y un conjunto de Gaussianas en la MoG). Para comprobar la veracidad de los parámetros

óptimos que se han obtenido, estos parámetros se han comparado con los parámetros utilizados por algunos autores [87][88] que han probado el método de MoG en secuencias de video con características similares a los videos de prueba que se han utilizado en este PFC.

<i>GAMMA</i>		
<i>ID SECUENCIA</i>	<i>RANGO UMBRAL</i>	<i>UMBRAL ÓPTIMO</i>
S1	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻⁵
S2	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻⁵
S3	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻⁷
S4	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻³
S5	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻⁷
S6	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻⁷
S7	[10 ⁻³ 10 ⁻⁵ 10 ⁻⁷]	10 ⁻³

Tabla 5-8. Parámetros óptimos para el método Gamma.

El rango de prueba para encontrar el umbral óptimo del método Gamma se ha seleccionado en base a los criterios que marcaba el método 3.2.3. La ventana de cálculo se ha fijado a 3 x 3, y los parámetros alpha y desviación que se han tomado son los óptimos encontrados en el método de la SG.

<i>KDE</i>				
<i>ID SECUENCIA</i>	<i>RANGO PRUEBA PROB. KDE</i>	<i>PROB. KDE ÓPTIMO</i>	<i>RANGO TAMAÑO BUFFER (L)</i>	<i>TAMAÑO BUFFER(L) ÓPTIMO</i>
S1	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.009	[20:20:80]	20
S2	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.005	[20:20:80]	20
S3	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.005	[20:20:80]	20
S4	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.009	[20:20:80]	20
S5	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.009	[20:20:80]	60

S6	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.009	[20:20:80]	80
S7	[0.005, 0.007, 0.009, 0.011, 0.013, 0.015]	0.009	[20:20:80]	80

Tabla 5-9. Parámetros óptimos para el método Estimación de núcleo Gaussiano (KDE).

La elección de los rangos de prueba del KDE se ha realizado en función de los parámetros elegidos por otros autores [86][87] que han ejecutado este método sobre videos de prueba con características similares a las de los videos de prueba que se han utilizado en este PFC.

<i>COMBINADO</i>						
<i>ID SECUENCIA</i>	<i>RANGO ALPHA</i>	<i>ALPHA ÓPTIMO</i>	<i>RANGO DESVIACIÓN</i>	<i>DESVIACIÓN ÓPTIMA</i>	<i>RANGO UMBRAL</i>	<i>UMBRAL ÓPTIMO</i>
S1	[10^{-7} , $5 \cdot 10^{-7}$, 10^{-6} , $5 \cdot 10^{-6}$, 10^{-5} , $5 \cdot 10^{-5}$, 10^{-4}]	$5 \cdot 10^{-5}$	[24: 2: 38]	28	[0.75:0.05:0.95]	0.75
S2	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-2}$	[20: 5: 50]	40	[0.75:0.05:0.95]	0.75
S3	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$5 \cdot 10^{-3}$	[10: 5: 40]	30	[0.75:0.05:0.95]	0.75
S4	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$5 \cdot 10^{-3}$	[5: 5: 35]	15	[0.75:0.05:0.95]	0.85
S5	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-3}$	[5: 5: 35]	20	[0.75:0.05:0.95]	0.85
S6	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-4}$	[24: 2: 34]	32	[0.75:0.05:0.95]	0.85
S7	[10^{-4} , $5 \cdot 10^{-4}$, 10^{-3} , $5 \cdot 10^{-3}$, 10^{-2} , $5 \cdot 10^{-2}$, 10^{-1}]	$1 \cdot 10^{-3}$	[2: 2: 18]	4	[0.75:0.05:0.95]	0.9

Tabla 5-10. Parámetros óptimos para el método Combinado de Histogramas y Gaussiana Simple.

Los parámetros elegidos para el método que fusiona Histogramas y SG han sido los óptimos encontrados de los métodos que componen este método (Generación de histogramas y SG) por separado.

5.5 Análisis comparativo

A lo largo de esta sección, se exponen los resultados obtenidos en las ejecuciones de los algoritmos implementados sobre los videos de prueba. En el primer apartado, se comparan los tiempos de ejecución de los algoritmos implementados. En el segundo, se analiza la calidad de los diferentes métodos de segmentación desarrollados en este PCF. Para ello, se utilizarán diversas medidas, explicadas en el apartado 5.3. Primero, se comparan los métodos básicos cualitativa y cuantitativamente y, después, las dos mejoras implementadas de los algoritmos SG y KDE.

Para comparar las versiones básicas de los algoritmos implementados, se han ejecutado con sus parámetros óptimos (ver tablas del apartado 5.4).

5.5.1 Tiempos de ejecución

En primer lugar, se ha realizado una comparativa en coste computacional (tiempo de ejecución) de cada algoritmo con cada secuencia de prueba. Los tiempos se han expresando en segundos de procesamiento por imagen. Para obtener un tiempo medio de ejecución de cada algoritmo, se ha realizado la media de los tiempos que ha tardado en ejecutarse cada secuencia por cada algoritmo.

Los tiempos de procesamiento de cada secuencia por cada algoritmo de modelado de fondo implementado se muestran en la Tabla 5-11:

<i>TIEMPOS DE PROCESAMIENTO (s / Imagen)</i>								
ID SECUENCIA	DIFERENCIA	MEDIANA	HISTOGRAMAS	SG	MoG	GAMMA	KDE	COMBINADO
S1	0.0153	0.0153	13.3929	0.0212	5.29	0.012	1.07	13.435
S2	0.0251	0.0840	13.8902	0.0306	4.98	0.017	1.121	13.956
S3	0.0057	0.0121	13.7242	0.0269	8.3839	0.015	0.523	13.929
S4	0.006	0.0108	13.5353	0.0292	3.5975	0.011	1.081	13.786
S5	0.0152	0.0152	13.4074	0.0292	3.5975	0.011	2.975	13.677
S6	0.0112	0.0161	13.5050	0.0295	4.1171	0.012	4.357	13.601
S7	0.0146	0.0156	14.0642	0.0292	3.8724	0.011	4.292	14.592
MEDIA DE TIEMPOS DE EJECUCIÓN	0,0133	0,0241	13,6456	0,0279	4,8341	0,0127	2,2027	13,8537

Tabla 5-11. Tiempos de procesamiento en segundos por imagen de cada método de segmentación basado en modelado de fondo.

Para comparar visualmente los tiempos de procesamiento de cada imagen por cada método, se han realizado dos diagramas de barras con las medias de los tiempos más altos y los tiempos más bajos obtenidos al aplicar cada método en cada secuencia de prueba.

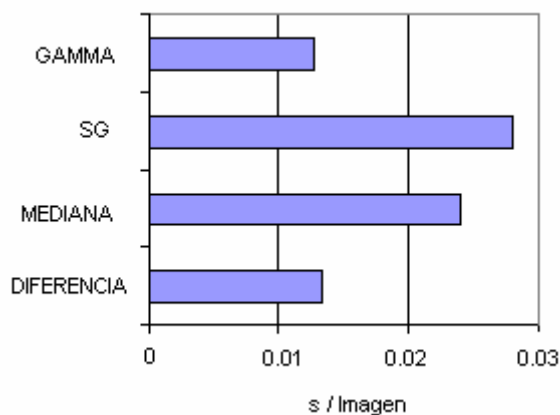


Figura 5-1: Tiempos medios de procesamiento expresados en segundos por imagen los métodos más rápidos de modelado de Fondo.

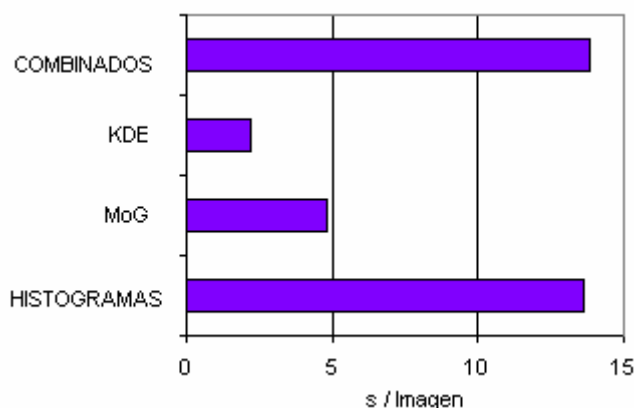


Figura 5-2: Tiempos medios de procesamiento expresados en segundos por imagen los métodos más lentos de modelado de Fondo.

Adicionalmente, en este proyecto se han mejorado dos de los algoritmos implementados: la SG mediante un filtro espacial y el KDE con un filtro espacio-temporal y dos colas de procesamiento. Para cada algoritmo con mejoras se han calculado los tiempos de procesamiento: En el caso de la mejora de la SG utilizando una ventana de 7×7 , el tiempo medio de ejecución es de 0.35 s/Imagen. Este tiempo incrementa de manera proporcional según el tamaño de la ventana de procesamiento que se utilice, aunque ese estudio no se ha realizado en este proyecto. En cuanto a las mejoras del KDE, los tiempos también se incrementan al añadir cada mejora: con la primera mejora (filtro vecinal) aumentan con respecto al KDE sin mejoras en 11 veces el tiempo de ejecución aproximadamente y, con la segunda mejora (dos colas de prioridad), el tiempo aumenta con

respecto al programa sin mejoras en el doble (estos datos se encuentran en la tabla posterior de resultados del KDE).

Es necesario mencionar que para el método del KDE las mejoras que se han realizado se han analizado independientemente para observar su funcionamiento con el método sin mejoras y poder compararlas, aunque las pruebas no se contemplan en este proyecto inicialmente se desarrollaron y se probaron ambas mejoras a la vez con KDE sin mejoras.

5.5.2 Calidad de la segmentación

A lo largo de este capítulo, se comparan y analizan los resultados obtenidos: en el primer apartado, de los métodos básicos implementados; en el segundo apartado, de la mejora de la SG; y en el tercer apartado, de las dos mejoras del KDE implementadas.

5.5.2.1 Comparativa de métodos básicos

Inicialmente, se han comparado todos los algoritmos sin mejoras con sus parámetros óptimos con cada secuencia de prueba. Para facilitar la comparativa se ofrecen tanto comparativas cualitativas como cuantitativas.

5.5.2.1.1 Análisis cuantitativo

En el caso de las cuantitativas se han obtenido en cada caso:

- Gráficas de porcentajes (expresados en tanto por uno) por cada imagen de cada secuencia de prueba de las siguientes medidas:
 - Positivos incorrectos: son los píxeles de valor uno detectados incorrectamente, es decir, detección del objeto en movimiento (FG) incorrecta.
 - Negativos incorrectos: son los píxeles de valor cero detectados incorrectamente, es decir, detección del fondo (BG) incorrecta.
- Tablas y diagramas de barras de las medidas de calidad (descritas en el apartado 5.3): precisión en cero (**P0**), precisión en uno (**P1**), *recall* en cero (**R0**), *recall* en uno (**R1**), *fscore* en cero (**FS0**) y *fscore* en uno (**FS1**). Todas estas medidas son las calculadas a todos los píxeles de cada imagen de las secuencias de prueba.

Estos resultados comparativos se muestran, para un conjunto de secuencias representativo, en las siguientes secciones.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S1:**

Características de las secuencias S1: Interior y Unimodal baja.

Como se puede observar en la **Figura 5-3**, para todos los modelos básicos implementados existe una fase inicial, hasta la imagen 100 de la secuencia, en la que el error de detección es nulo debido a que la imagen es unimodal e interior y no aparece ningún objeto en la escena (el error de FG es nulo). Los modelos de BGS que mejores resultados obtienen en la detección de objeto son los que producen menores porcentajes de error en el FG, que son los métodos paramétricos: Gamma (en verde), MoG (en amarillo) y SG (rosa). El porcentaje de error en los píxeles de objeto (de valor uno, del FG) es inferior al 0.1 en todas las imágenes de la secuencia en el caso del Gama y el MoG. Los métodos sencillos y no paramétricos detectan un error del objeto en movimiento en torno al 0.5 (KDE representado con el color negro) hasta el 0.3 (Diferencia de Frentes verde oscuro).

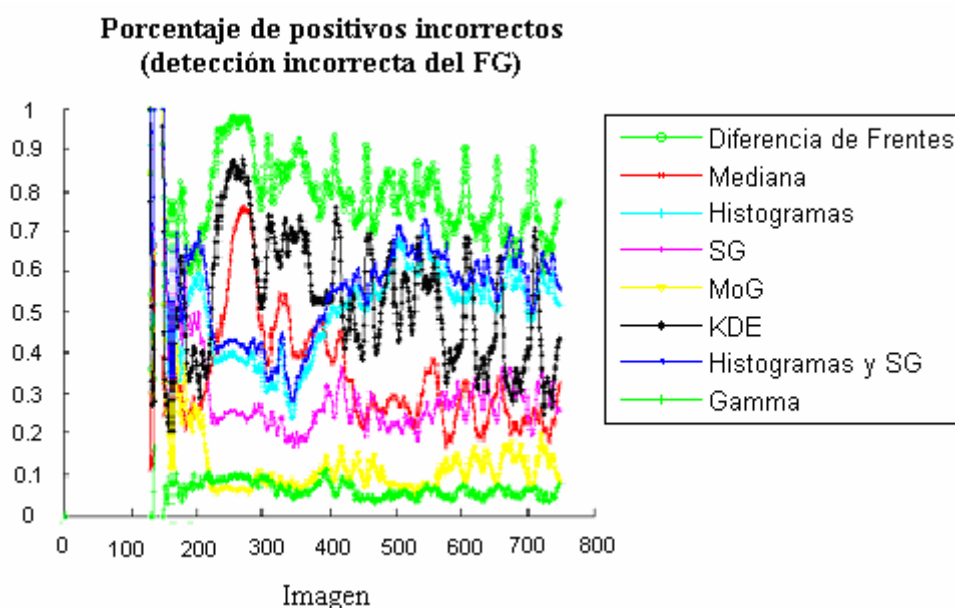


Figura 5-3: Gráfica de errores en la detección del ForeGround.

En cuanto a la **Figura 5-4**, destaca que la mayoría de los métodos detectan correctamente los píxeles de fondo (con valor cero) ya que el porcentaje de error de fondo para todas las imágenes está por debajo del 0.1. Sin embargo, el método KDE detecta mucho ruido en todas las imágenes como se puede observar en la gráfica inferior (alcanza 0.1 de error en BG). Por este motivo, el método KDE se implementa con las mejoras estudiadas, ya que son necesarias para eliminar el elevado porcentaje de ruido que se genera en el BG con ese método en todas las imágenes.

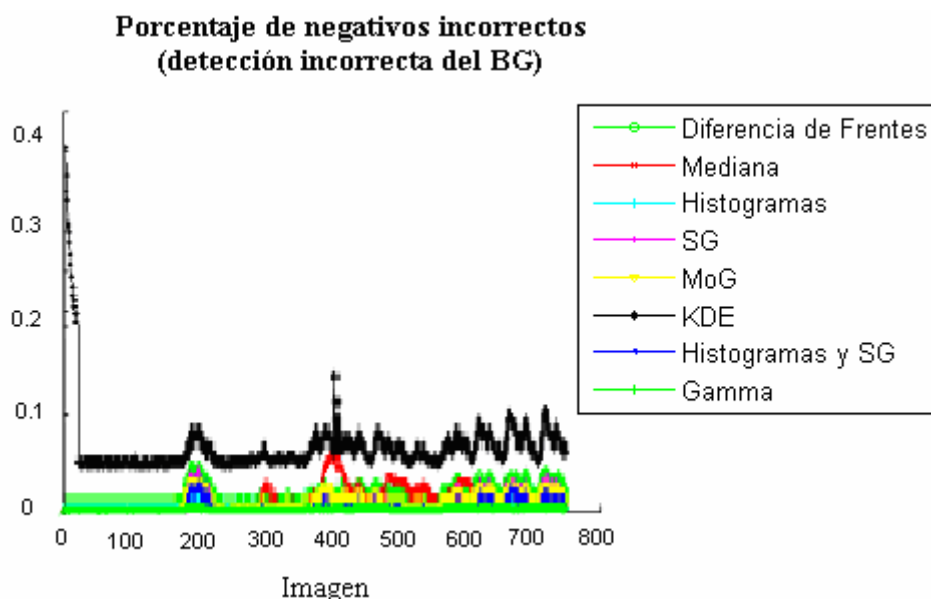


Figura 5-4: Gráfica de errores en la detección del BackGround.

SECUENCIA S1: (INTERIOR, UNIMODAL BAJA)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	92.18	58.90	98.34	22.15	95.17	32.20
MEDIANA	96.50	86.62	98.90	66.51	97.68	72.25
HISTOGRAMAS	94.82	89.35	99.37	49.37	97.04	63.61
SG	97.12	89.78	99.11	72.46	98.10	80.20
MoG	98.78	90.22	98.97	88.67	98.88	89.44
KDE	94.31	38.99	91.92	48.23	93.10	43.12
COMBINADO	94.39	91.03	99.53	44.77	96.89	60.02
GAMMA	99.32	97.79	99.78	93.66	99.54	95.68

Tabla 5-12 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S1.

A la vista de los resultados de la Tabla 5.12, que agrupa los indicadores de calidad, y de los diagramas de la **Figura 5-3** y **Figura 5-4**, los métodos que ofrecen mejores resultados globales del procesamiento de todas las imágenes de la secuencia de tipo unimodal S1 analizada son el método Gamma, MoG y SG. Esto se debe a que estos modelos son capaces de construir un fondo en base a unos parámetros (que los métodos básicos y no paramétricos no utilizan) y actualizarlo correctamente para poder determinar con un alto porcentaje de acierto los píxeles de FG y BG.

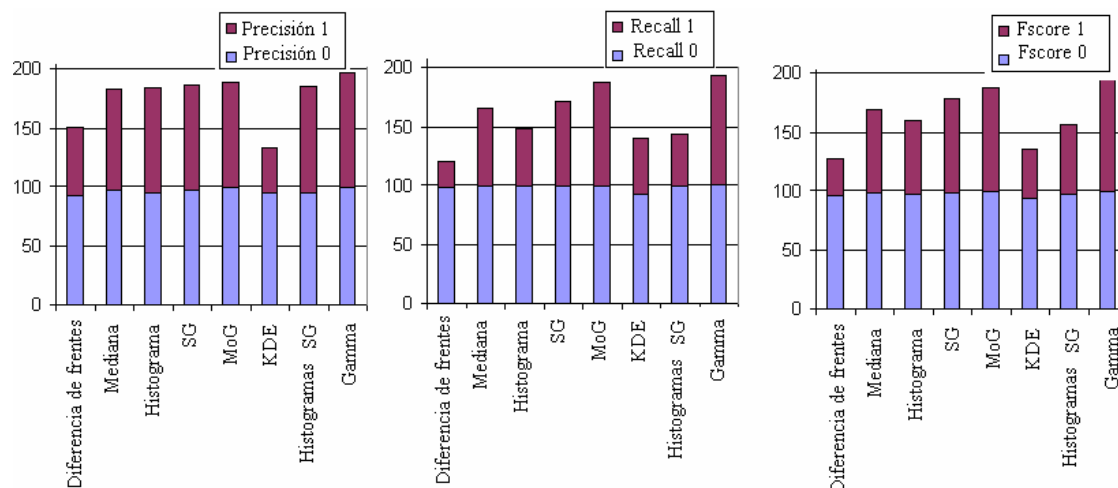


Figura 5-5: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S1.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S2:**

Características de la secuencias S2: Exterior y Multimodal alto.

Las gráficas siguientes representan los porcentajes (expresados en tanto por uno) de positivos incorrectos y negativos incorrectos obtenidos al ejecutar todos los modelos de segmentación de objetos implementados por cada imagen de una de las secuencias S2.

Para la secuencia de tipo multimodal alto S2, los resultados son diferentes con respecto a los anteriores en los que se procesaba una secuencia unimodal baja. Como se puede observar en la gráfica **Figura 5-6**, los métodos básicos de segmentación de objetos Histogramas (línea de color azul claro) y Mediana (línea de color rojo) y el método no paramétrico KDE (de color negro) ofrecen mejores resultados en la detección del objeto (píxeles de valor uno) que los métodos paramétricos. Por el contrario, estos métodos también detectan un exceso de píxeles de objeto (de valor uno) en los píxeles de fondo (de valor cero), lo cual significa que producen peores resultados en la detección del fondo, como puede observarse en la gráfica **Figura 5-7**.

Es importante señalar que la extracción de un fondo cualitativamente correcto se realiza únicamente con los modelos paramétricos y no paramétricos (según se puede ver más adelante en la Tabla 5-13), ya que los modelos básicos no actualizan el fondo adecuadamente, y en cada instante, el fondo que representan es incorrecto, a pesar de que la detección del objeto sea mejor con los métodos básicos.

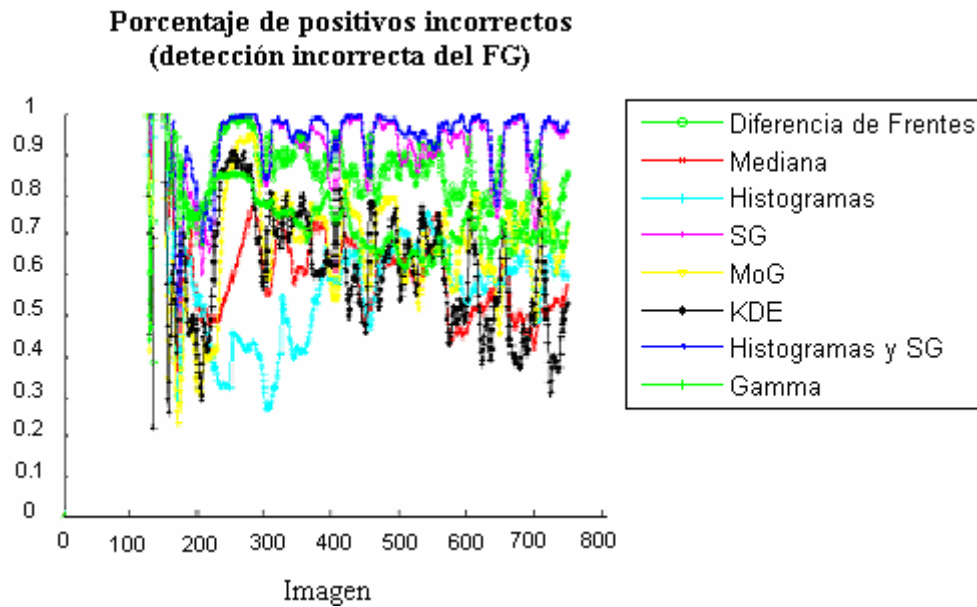


Figura 5-6: Gráfica de errores en la detección del ForeGround.

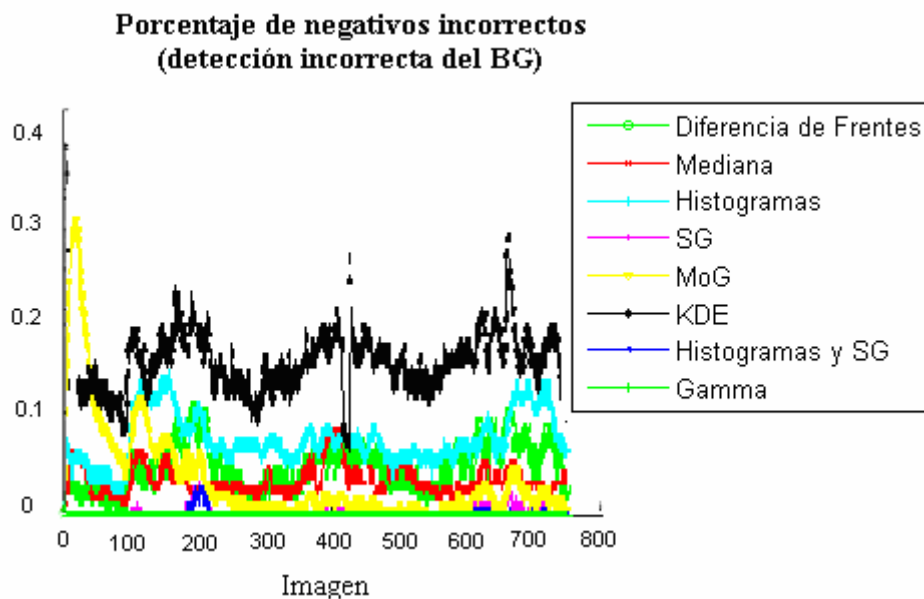


Figura 5-7: Gráfica de errores en la detección del BackGround.

La Tabla 5-13 que se muestra a continuación expone los resultados de precisión, *recall* y *fscore* de la ejecución de cada método de segmentación de objetos que se ha desarrollado sobre la secuencia S2.

En los resultados que se encuentran en la Tabla 5-13 y diagramas **Figura 5-6** y **Figura 5-7** se observa que en cómputo global, los resultados mejores en detección de fondo son los realizados por el método Gamma, lo cual parece lógico puesto que es el único método que utiliza por cada imagen

un eventanado 3 x 3 que reduce los errores de detección en píxeles aislados. Y posteriormente, producen resultados parecidos los métodos básicos de segmentación Mediana e Histogramas.

A la vista de los resultados podemos observar que los métodos básicos producen mejores porcentajes de detección, en general, que los paramétricos (salvo el método Gamma debido al eventanado, como ya se ha mencionado). Esto se debe a las características de las secuencia S2; secuencias muy complejas, es decir, exteriores de tipo multimodal alto, donde los píxeles de fondo presentan más variación por cada imagen que los propios objetos, son altamente complejas de analizar por cualquier algoritmo de segmentación, por este motivo, los métodos sencillos, que no construyen distribuciones que pueden ser erróneas sino que representan el fondo mediante cálculos más sencillos, ofrecen finalmente, en este caso, mejores resultados.

SECUENCIA: S2 (EXTERIOR, MULTIMODAL ALTO)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	91.41	30.19	96.16	15.51	93.72	20.49
MEDIANA	93.96	59.85	97.01	41.69	95.45	49.15
HISTOGRAMAS	94.05	44.79	94.14	44.40	94.09	44.60
SG	91.03	70.62	99.63	8.25	95.13	14.78
MoG	93.03	52.16	96.83	32.28	94.89	39.87
KDE	93.23	25.58	87.63	39.7	90.30	31.11
HIST + SG	90.80	77.56	99.82	5.61	95.10	10.48
GAMMA	92.84	99.99	99.99	27.92	96.28	43.65

Tabla 5-13 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S2.

Finalmente, se han realizado los diagramas de barras **Figura 5-8** con los resultados de la Tabla 5-13 superior para poder comparar los métodos de segmentación de objetos fácilmente. De izquierda a derecha, el primer diagrama de barras representa la precisión en ceros (P0) y en unos (P1), el segundo el *recall* en ceros (R0) y en unos (R1) y el tercero el *fscore* en ceros (FS0) y en unos (FS1).

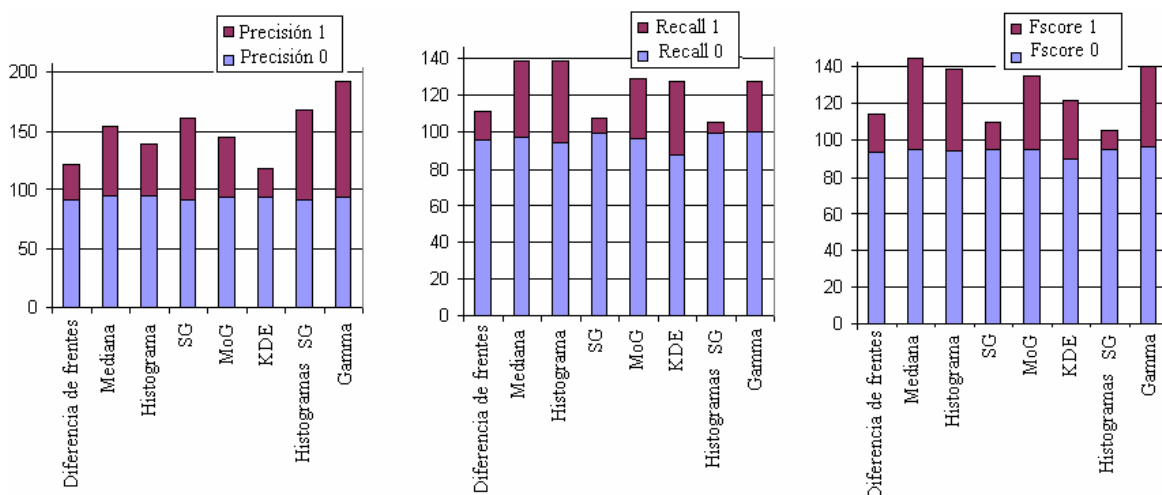


Figura 5-8: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S2.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S3:**

Características de la secuencias: Exterior y Multimodal medio.

Las gráficas **Figura 5-9** y **Figura 5-10** representan el porcentaje (expresados en tanto por uno) de positivos incorrectos y negativos incorrectos obtenidos al ejecutar todos los modelos de segmentación de objetos implementados por cada imagen de la secuencia S3.

Como se observa en la gráfica **Figura 5-9**, que presenta los resultados del análisis de la secuencia S3 de tipo multimodal medio, los mejores porcentajes en la detección del objeto son los que producen los métodos MoG (amarillo), Mediana (rojo) y KDE (negro) los cuales varían en función del método de 0.1 a 0.5 en la detección incorrecta del FG. No obstante, estos métodos también producen peores resultados en la detección del fondo, como se aprecia en la gráfica inferior **Figura 5-10**.

Estos resultados indican que en el procesamiento de secuencias de alta complejidad es necesario lograr un compromiso en la detección del objeto y fondo correctas para lograr una detección adecuada (sin demasiado ruido y sin perder el objeto en movimiento). Una forma de resolver este problema sería mediante técnicas de pre-post procesado.

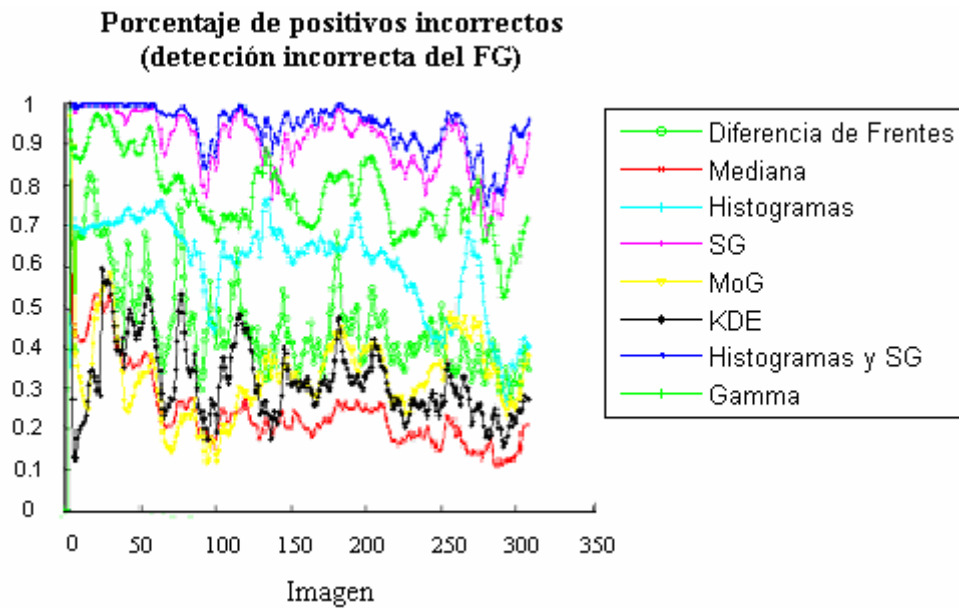


Figura 5-9: Gráfica de errores en la detección del ForeGround.

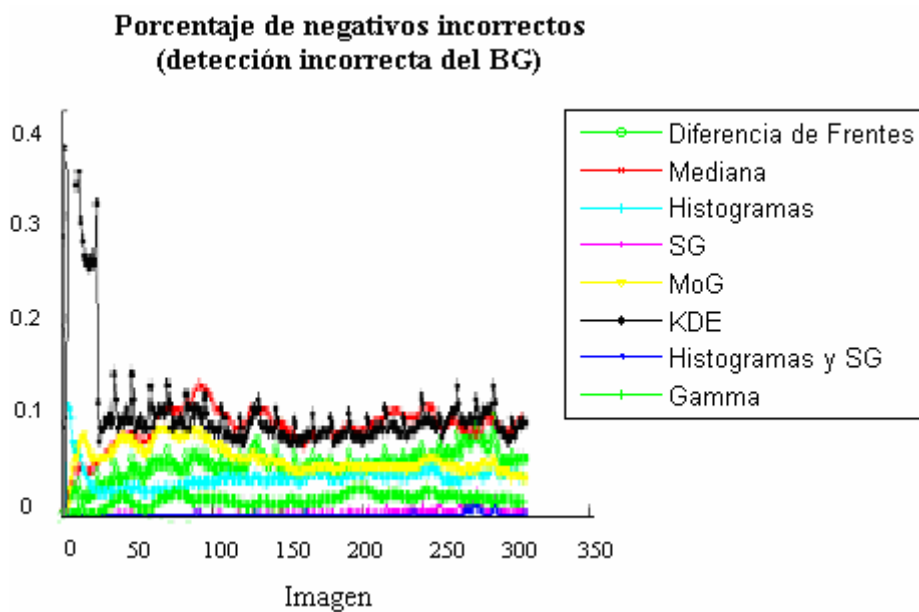


Figura 5-10: Gráfica de errores en la detección del BackGround.

La Tabla 5-14 que se muestra a continuación expone los resultados de precisión, *recall* y *fscore* de la ejecución de cada método de segmentación de objetos que se ha desarrollado sobre la secuencia S3.

SECUENCIA S3: (EXTERIOR, MULTIMODAL MEDIO)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	95.61	48.96	94.58	54.55	95.10	51.56
MEDIANA	97.52	44.14	90.84	75.81	94.06	55.79
HISTOGRAMAS	94.35	48.46	95.96	39.77	95.14	43.68
SG	92.04	66.78	99.53	9.81	95.64	17.11
MoG	96.74	51.48	93.99	66.82	95.34	58.15
KDE	96.73	37.43	89.09	68.39	92.75	48.38
HIST + SG	91.74	76.10	99.83	5.85	95.62	10.84
GAMMA	93.01	54.60	99.19	22.78	95.54	32.16

Tabla 5-14 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S3.

Finalmente, se han realizado los siguientes diagramas de barras con los resultados de la tabla superior (Tabla 5-14) para poder comparar los métodos de segmentación de objetos fácilmente. De izquierda a derecha, el primer diagrama de barras representa la precisión en ceros (P0) y en unos (P1), el segundo el *recall* en ceros (R0) y en unos (R1) y el tercero el *fscore* en ceros (FS0) y en unos (FS1).

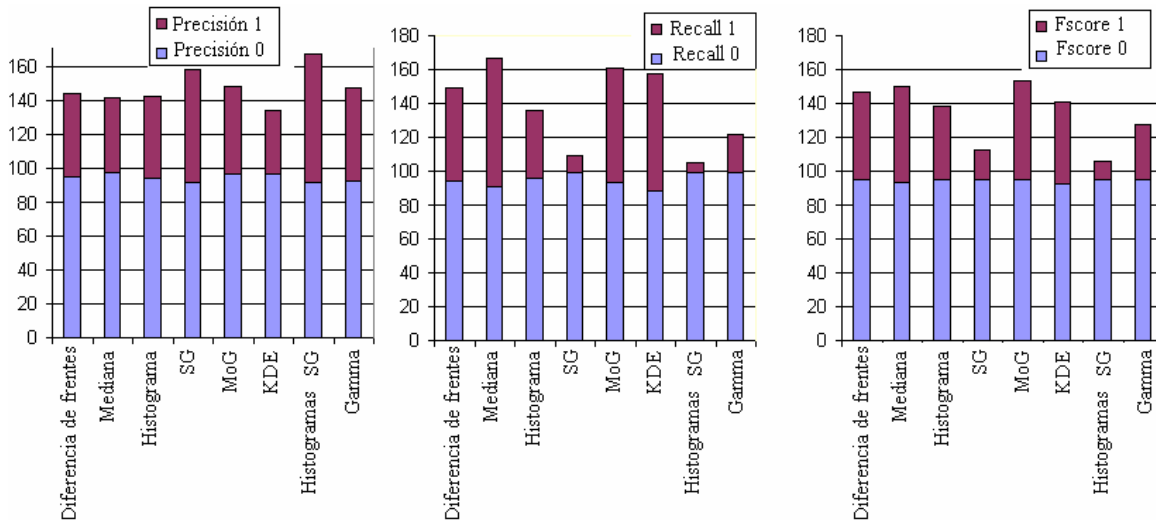


Figura 5-11: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S3.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S4:**

Características de la secuencias: Exterior y Multimodal medio.

Las gráficas siguientes representan los porcentajes (expresados en tanto por uno) de detección incorrecta en los píxeles de objeto **Figura 5-12** y detección incorrecta en los píxeles de fondo

Figura 5-13 obtenidos al ejecutar todos los modelos de segmentación de objetos implementados por cada imagen de la secuencia S4.

La secuencia S4 es multimodal (con un alto grado de cambio de los píxeles de fondo debido al movimiento de árboles y arbustos que pertenecen al fondo), esto provoca que los métodos paramétricos diseñados para modelar fondos multimodales como MoG (amarillo en las gráficas inferiores), y Gamma (verde oscuro en las gráficas inferiores) detecten correctamente un alto porcentaje de positivos (FG), en torno al 0.7, y también detecten correctamente un alto porcentaje de negativos (BG) alrededor del 0.9. Sin embargo, los métodos básicos como Diferencia de Frentes (verde claro en las gráficas inferiores) e Histogramas (azul claro en las gráficas inferiores) detectan incorrectamente positivos (FG) pero correctamente negativos (BG), es decir, generan máscaras de frente con un exceso de píxeles erróneos de objeto. Este hecho se debe a que estos métodos no están diseñados para modelar fondos multimodales como es este caso.

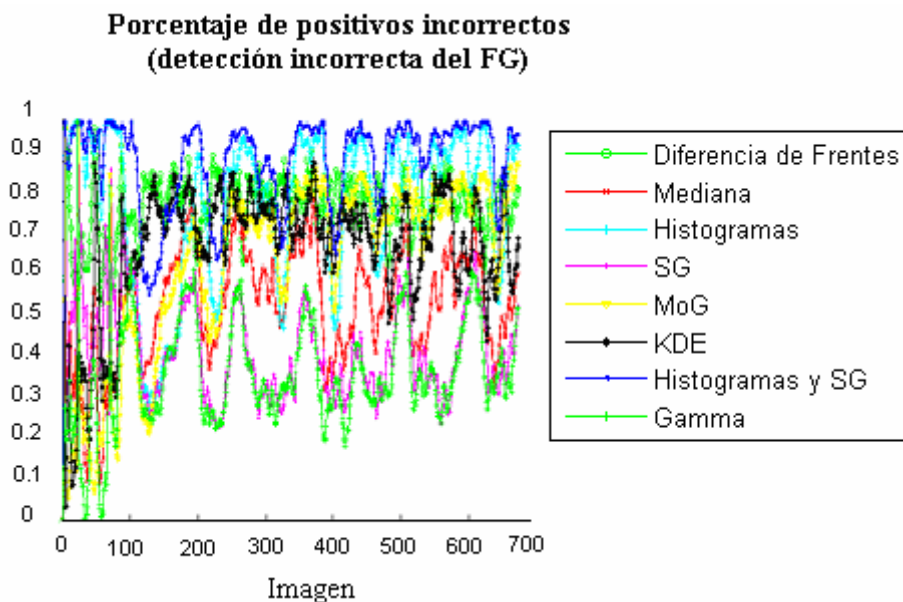


Figura 5-12: Gráfica de errores en la detección del ForeGround.

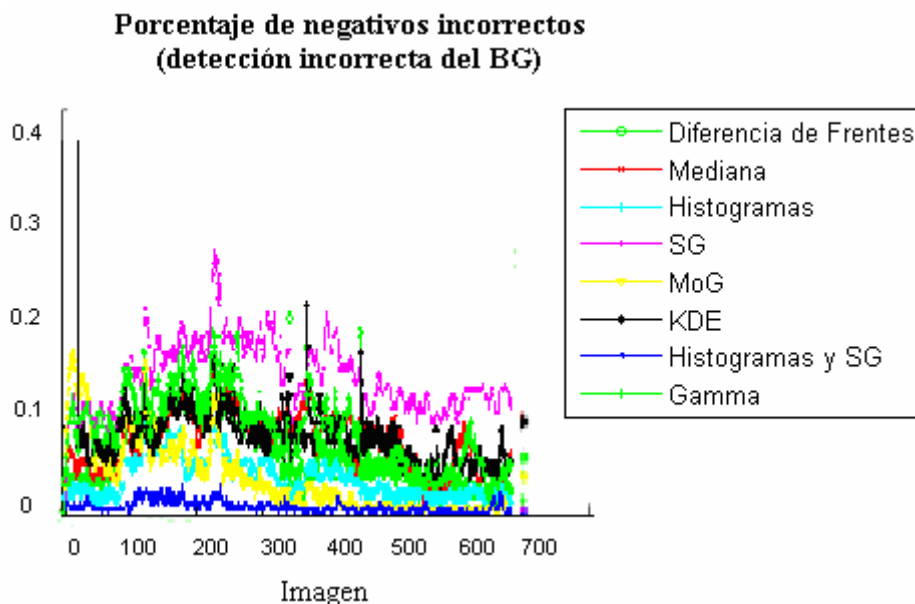


Figura 5-13: Gráfica de errores en la detección del BackGround.

La tabla Tabla 5-15 que se muestra a continuación expone los resultados de precisión, *recall* y *fscore* de la ejecución de cada método de segmentación de objetos que se ha desarrollado sobre la secuencia S4.

SECUENCIA S4: (EXTERIOR, MULTIMODAL MEDIO)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	88.21	30.97	92.95	20.29	90.52	24.52
MEDIANA	91.12	44.07	91.55	42.74	91.33	43.39
HISTOGRAMAS	88.81	45.29	95.75	22.56	92.15	30.12
SG	92.85	36.12	83.86	58.60	88.13	44.69
MoG	89.70	50.94	95.56	29.55	92.54	37.40
KDE	88.74	28.45	89.29	27.33	89.02	27.88
HIST + SG	87.64	63.13	99.05	10.34	93.01	17.77
GAMMA	93.74	48.82	89.97	61.43	91.81	54.40

Tabla 5-15 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S4.

Finalmente, se han realizado los siguientes diagramas de barras con los resultados de la tabla superior (Tabla 5-15) para poder comparar los métodos de segmentación de objetos fácilmente. De izquierda a derecha, el primer diagrama de barras representa la precisión en ceros (P0) y en unos (P1), el segundo el *recall* en ceros (R0) y en unos (R1) y el tercero el *fscore* en ceros (FS0) y en unos (FS1).

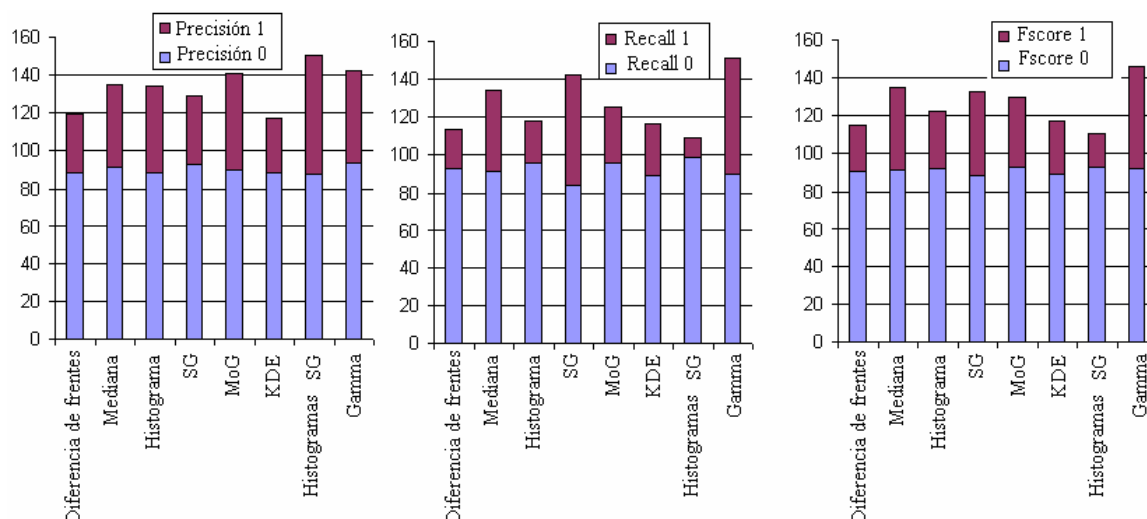


Figura 5-14: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S4.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S5:**

Características de las secuencias S5: Exterior y Multimodal medio.

Las gráficas siguientes (**Figura 5-15** y **Figura 5-16**) representan los porcentajes (expresados en tanto por uno) de detección incorrecta en los píxeles de objeto y detección incorrecta en los píxeles de fondo obtenidos al ejecutar todos los modelos de segmentación de objetos implementados por cada imagen de la secuencia S5.

La secuencia S5 se caracteriza como multimodal medio aunque los píxeles de fondo presentan una variación muy leve y además los píxeles que varían son pocos, por este motivo los resultados obtenidos son similares a los presentados anteriormente con la secuencia S1 (unimodal) . Como se aprecia en la gráficas **Figura 5-15** y **Figura 5-16**, los métodos que aportan mejores detecciones en el objeto en movimiento (FG) y en el fondo (BG) son los modelos Gamma (de color verde), MoG (de color amarillo) y SG (de color rosa) con un porcentaje de error en objeto del 0.2 aproximadamente y 0.1 aproximadamente de error en el fondo.

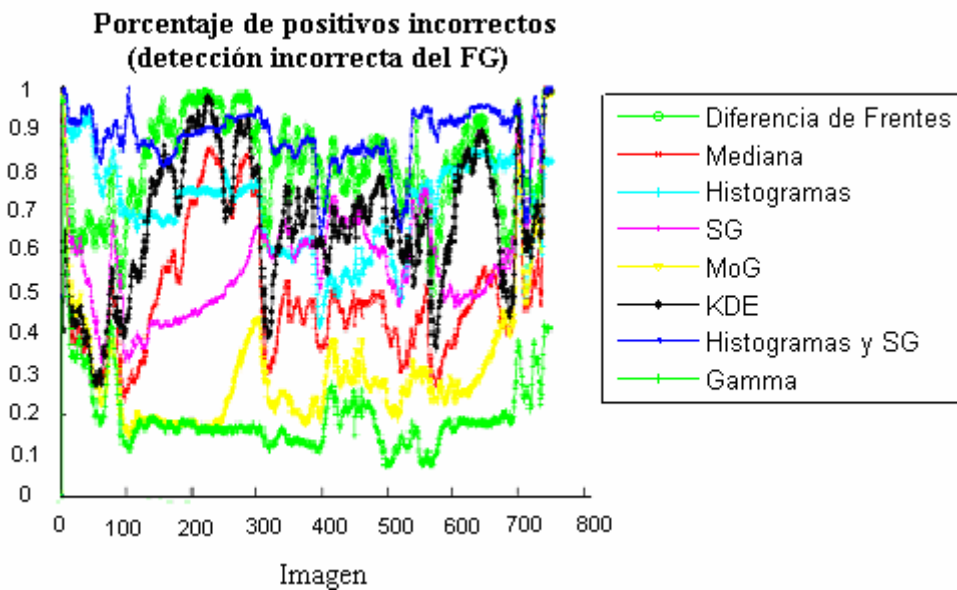


Figura 5-15: Gráfica de errores en la detección del ForeGround.

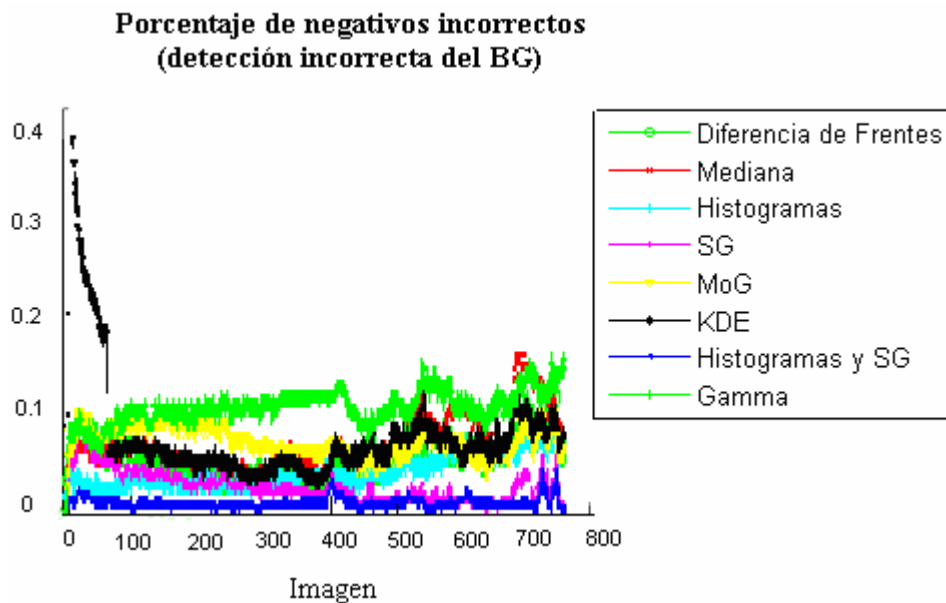


Figura 5-16: Gráfica de errores en la detección del BackGround.

La Tabla 5-16 que se muestra a continuación expone los resultados de precisión, *recall* y *fscore* de la ejecución de cada método de segmentación de objetos que se ha desarrollado sobre la secuencia S5.

SECUENCIA S5: (EXTERIOR, MULTIMODAL MEDIO)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	89.62	32.87	94.37	20.15	91.94	24.98
MEDIANA	93.45	50.87	93.07	52.38	93.26	51.62
HISTOGRAMAS	91.01	50.82	95.93	30.73	93.41	38.29
SG	92.56	68.01	97.24	42.90	94.84	52.62
MoG	95.78	57.21	92.82	70.15	94.28	63.03
KDE	91.08	37.08	92.08	34.11	91.58	35.54
HIST + SG	89.25	62.86	98.96	12.88	93.85	21.39
GAMMA	97.20	51.35	89.46	81.20	93.17	62.92

Tabla 5-16 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S5.

Finalmente, se han realizado los siguientes diagramas de barras (Tabla 5-17) con los resultados de la Tabla 5-16 para poder comparar los métodos de segmentación de objetos fácilmente. De izquierda a derecha, el primer diagrama de barras representa la precisión en ceros (P0) y en unos (P1), el segundo el *recall* en ceros (R0) y en unos (R1) y el tercero el *fscore* en ceros (FS0) y en unos (FS1).

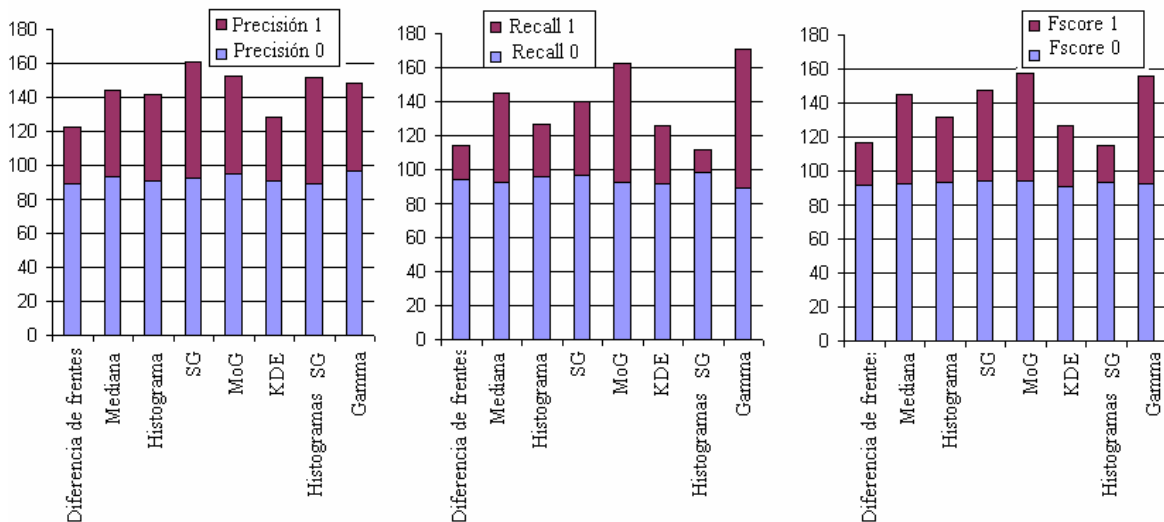


Figura 5-17: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S5.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S6:**

Características de la secuencias: Exterior y Unimodal.

Las gráficas **Figura 5-18** y **Figura 5-19** representan los porcentajes (expresados en tanto por uno) de detección incorrecta en los píxeles de objeto (**Figura 5-18**) y detección incorrecta en los píxeles de fondo (**Figura 5-19**) obtenidos al ejecutar todos los modelos de segmentación de objetos implementados por cada imagen de la secuencia S6.

La secuencia S6 es una secuencia unimodal, por este motivo, los modelos paramétricos Gamma (de color verde en las gráficas inferiores), MoG (de color amarillo en las gráficas inferiores) y SG (de color rosa en las gráficas inferiores) detectan un alto porcentaje de píxeles de objeto (FG) y también detectan correctamente un gran porcentaje de píxeles de fondo (BG), ya que como se puede observar en las figuras **Figura 5-18** y **Figura 5-19** la detección incorrecta en FG y BG está próxima a 0.

A la vista de los resultados, los métodos paramétricos producen muy buenos resultados en la detección del objeto en fondos de píxeles con poca variación (unimodales y multimodales medios) representando el fondo con los parámetros de una distribución.

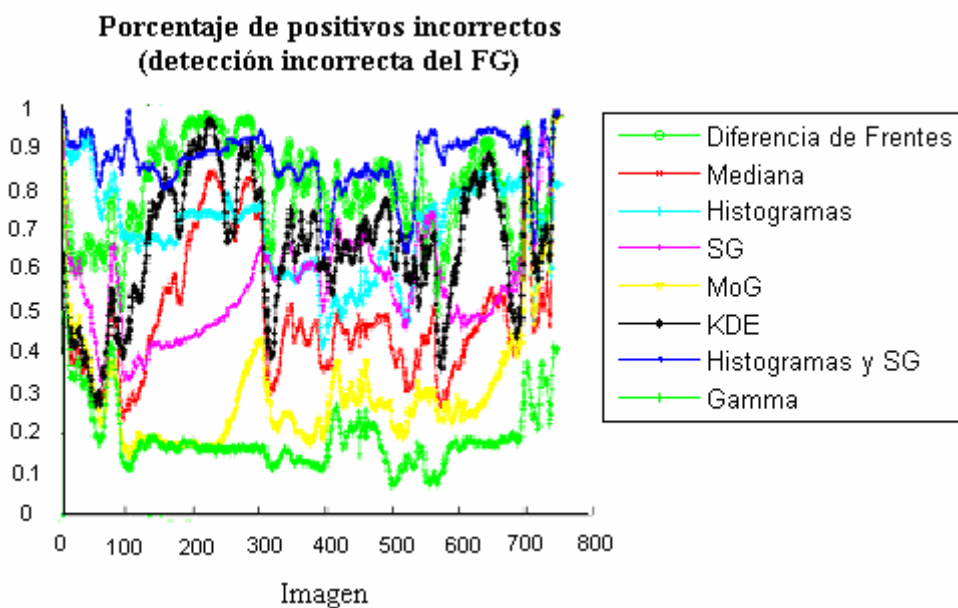


Figura 5-18: Gráfica de errores en la detección del ForeGround.

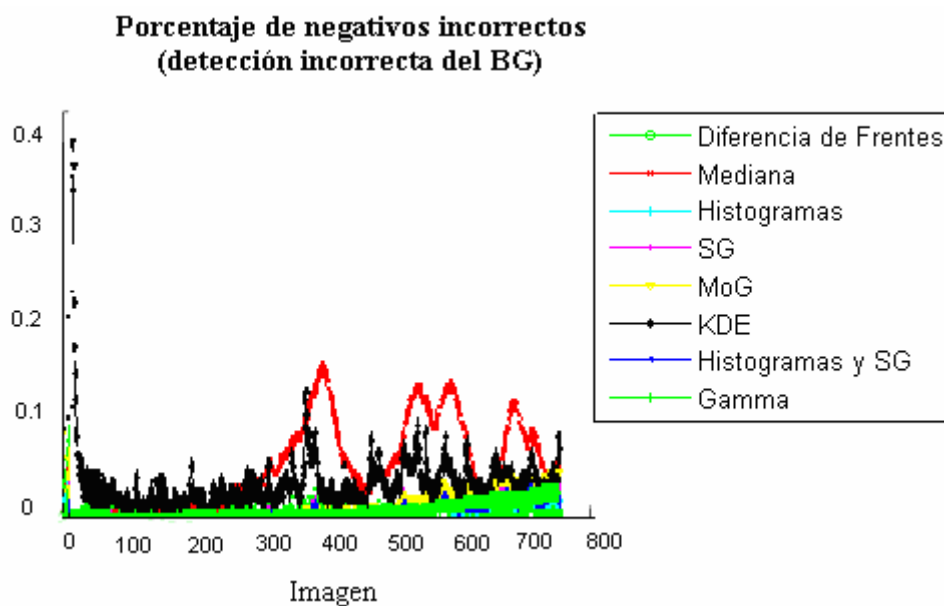


Figura 5-19: Gráfica de errores en la detección del BackGround.

La Tabla 5-17 que se muestra a continuación expone los resultados de precisión, *recall* y *fscore* de la ejecución de cada método de segmentación de objetos que se ha desarrollado sobre la secuencia S6.

SECUENCIA S6: (EXTERIOR, UNIMODAL)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	93.46	68.26	99.07	22.27	96.18	33.59
MEDIANA	98.53	61.69	95.34	84.08	96.91	71.16
HISTOGRAMAS	97.76	90.22	99.27	74.55	98.51	81.64
SG	98.83	87.06	98.84	86.98	98.84	87.02
MoG	99.38	85.78	98.62	93.11	99.01	89.29
KDE	94.66	49.55	96.46	38.99	95.55	43.64
HIST + SG	97.37	88.81	99.21	70.02	98.28	78.31
GAMMA	99.60	90.39	99.09	95.58	99.34	92.91

Tabla 5-17 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S6.

Finalmente, se han realizado los siguientes diagramas de barras (**Figura 5-20**) con los resultados de la tabla superior (Tabla 5-17) para poder comparar los métodos de segmentación de objetos fácilmente. De izquierda a derecha, el primer diagrama de barras representa la precisión en ceros (P0) y en unos (P1), el segundo el *recall* en ceros (R0) y en unos (R1) y el tercero el *fscore* en ceros (FS0) y en unos (FS1).

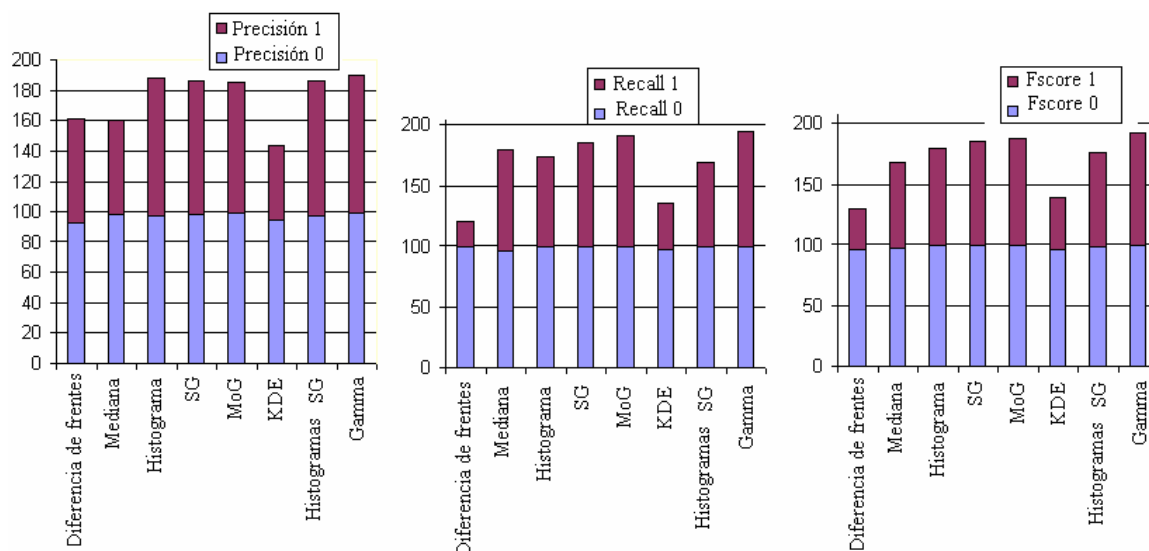


Figura 5-20: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S6.

➤ **SECUENCIA DEL CONJUNTO DE PRUEBA S7:**

Características de las secuencias: Exterior y Unimodal.

Las gráficas **Figura 5-21** y **Figura 5-22** representan los porcentajes (expresados en tanto por uno) de detección incorrecta en los píxeles de objeto **Figura 5-21** y detección incorrecta en los píxeles de fondo **Figura 5-22** obtenidos al ejecutar todos los modelos de segmentación de objetos implementados por cada imagen de la secuencia S7.

En cuanto a la secuencia S7 es necesario mencionar antes de analizar las gráficas que, en primer lugar, es una secuencia unimodal (prácticamente sin variación en los píxeles de fondo) y, en segundo lugar, el número de píxeles de fondo en las imágenes de la secuencia es muy superior al de los de objeto, es decir, el objeto es muy pequeño con respecto al fondo. Por este motivo, y como se puede observar con mayor claridad en la Tabla 5-18, el porcentaje de fondo incorrectamente detectado es muy pequeño en todos los métodos de segmentación de objetos implementados, se encuentra en torno al 0.1. No obstante, los métodos que ofrecen mejores resultados en la detección del objeto (FG) son el Gamma (de color verde en las gráficas inferiores), MoG (de color amarillo en las gráficas inferiores) y KDE (de color negro en las gráficas inferiores), tal y como ocurría en los resultados del procesamiento de otras secuencias de tipo unimodal de bajo.

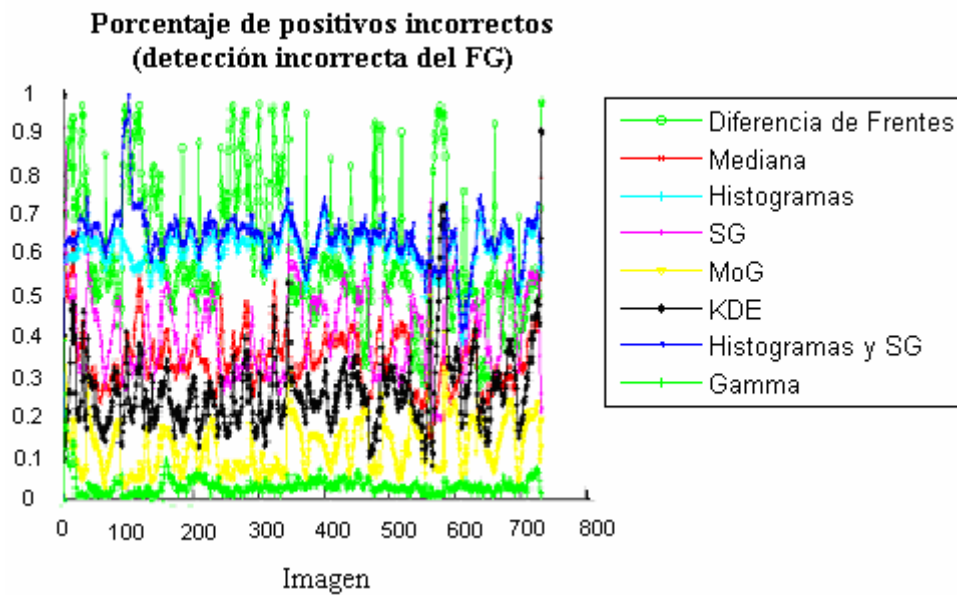


Figura 5-21: Gráfica de errores en la detección del BackGround.

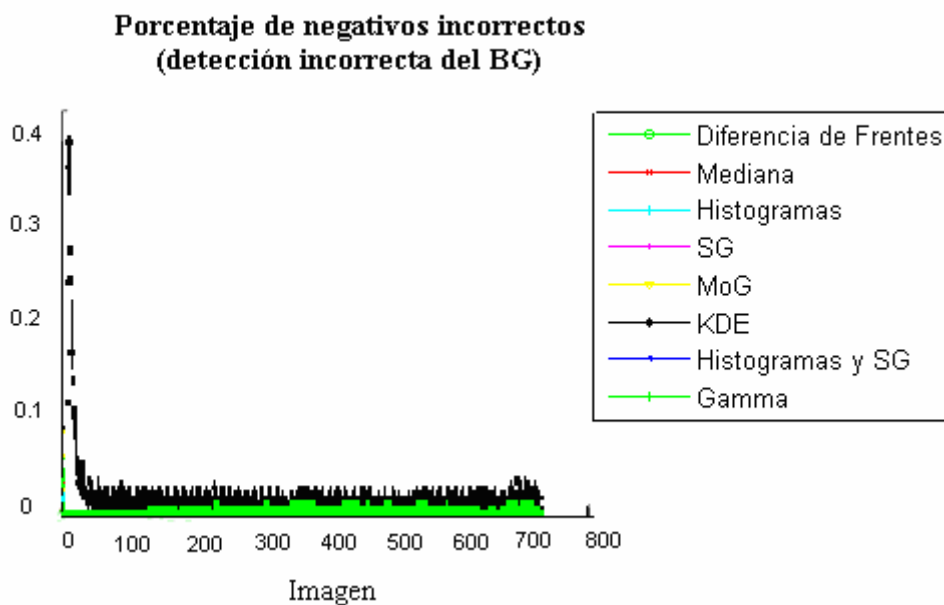


Figura 5-22: Gráfica de errores en la detección del BackGround.

La Tabla 5-18 que se muestra a continuación expone los resultados de precisión, *recall* y *fscore* de la ejecución de cada método de segmentación de objetos que se ha desarrollado sobre la secuencia S7.

SECUENCIA S7: (EXTERIOR, UNIMODAL)						
Método de BS	P0	P1	R0	R1	FS0	FS1
FD	99.21	71.99	99.75	44.07	99.48	54.67
MEDIANA	99.53	89.95	99.89	66.93	99.71	76.75
HISTOGRAMAS	99.14	73.53	99.80	39.15	99.47	51.09
SG	99.38	60.75	99.48	56.69	99.43	58.65
MoG	99.80	85.58	99.79	86.49	99.80	86.03
KDE	99.62	33.04	97.88	74.02	98.74	45.69
HIST + SG	99.10	87.10	99.92	36.37	99.51	51.31
GAMMA	99.95	65.11	99.26	97.18	99.61	77.97

Tabla 5-18 . Resultados comparativos de todos los métodos ejecutados sobre la secuencia S7.

Finalmente, se han realizado los siguientes diagramas de barras **Figura 5-23** con los resultados de la tabla superior (Tabla 5-18) para poder comparar los métodos de segmentación de objetos fácilmente. De izquierda a derecha, el primer diagrama de barras representa la precisión en ceros (P0) y en unos (P1), el segundo el *recall* en ceros (R0) y en unos (R1) y el tercero el *fscore* en ceros (FS0) y en unos (FS1).

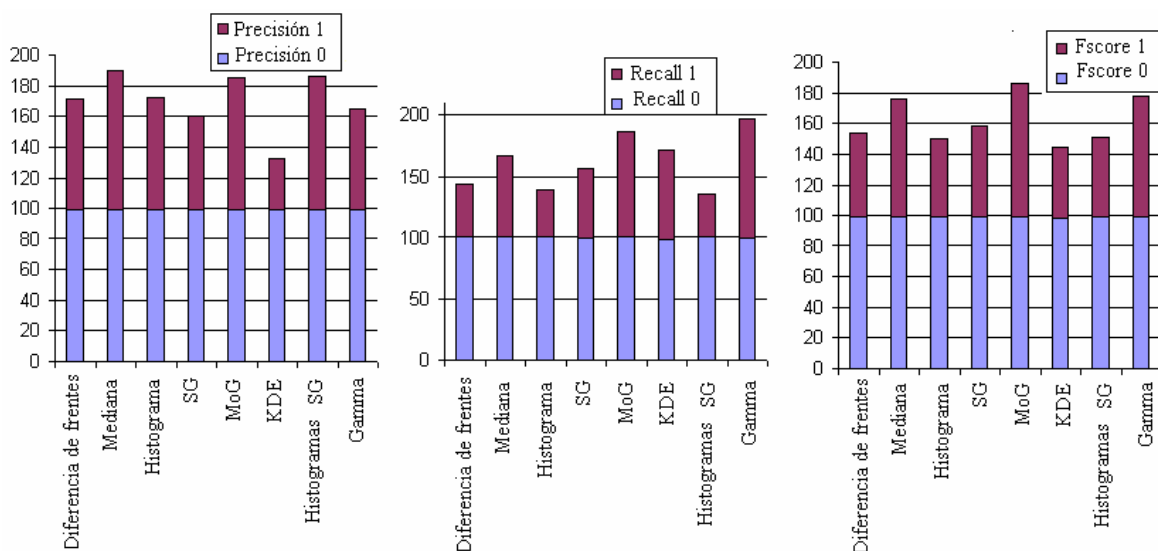


Figura 5-23: Diagramas de barras (de izquierda a derecha) de: precisión, recall, fscore en la secuencia S7.

5.5.2.1.2 Análisis cualitativo

Finalmente, se han seleccionado dos de las secuencias de los videos de prueba de características diferentes para mostrar las imágenes y el frente que se han obtenido al ejecutar los diferentes algoritmos que se han implementado. Las imágenes que se observan en la siguiente figura (**Figura**

5-24) han sido capturadas en un instante de tiempo en el que podían observarse con claridad las diferencias de los frentes obtenidos por los distintos algoritmos de segmentación.

Comparando los frentes de las dos secuencias se puede destacar la cantidad de píxeles de fondo erróneamente detectados como frente de la secuencia multimodal (situada a la izquierda de la **Figura 5-24**) frente a la claridad de detección del objeto en movimiento de la secuencia unimodal (situada a la derecha de la **Figura 5-24**). También se pueden visualizar claramente los resultados expuestos de forma numérica en el apartado anterior (5.5.2.1.1): en las secuencias multimodales muy complejas, los algoritmos sencillos detectan mejor el objeto a costa de perder píxeles de fondo frente a los algoritmos paramétricos y no paramétricos, que funcionan mejor en secuencias multimodales medias y unimodales detectando con gran precisión los píxeles de fondo y objeto.





















METODO MODELADO DE FONDO	IMAGEN DE FONDO MULTIMODAL instante 204	FRENTES instante 204	IMAGEN DE FONDO UNIMODAL instante 279	FRENTES instante 279
DIFERENCIA DE FERNTES				
MEDIANA				
HISTOGRAMAS				
SG				
MoG				



Figura 5-24: Imagen y Frentes obtenidos con los distintos métodos de segmentación de objetos sobre dos de las secuencias del banco de pruebas: Baile 2(multimodal) y Hambre (unimodal).

5.5.2.1.3 Conclusiones

En este apartado se exponen de manera resumida los parámetros óptimos y rangos de variación de cada algoritmo en función de las características de cada secuencia de prueba y, se compara la calidad de los segmentadores implementados en función de las características de las secuencias de prueba.

Una vez comparados los métodos de segmentación de objetos basados en modelado de fondo para cada secuencia de prueba, podemos concluir cuáles son los parámetros más críticos y los rangos de variación en los que se deben encontrar los parámetros de cada algoritmo de segmentación para ofrecer buenos resultados, en función de las características de las secuencias, que son los siguientes:

- Secuencias multimodales, con un elevado porcentaje de ruido en el fondo, requieren:
 - Los métodos básicos: umbrales altos que eviten detectar movimientos del fondo como parte del objeto. Para el método de Diferencia de Frentes se debería tomar un umbral elevado aproximadamente en un rango [300-500], para el caso de la Mediana [50-100] y para Histogramas [0.7-0.8]. Variar los parámetros en los rangos indicados para cada método no afecta en gran medida los resultados finales.
 - Los métodos paramétricos (Gamma, SG y MoG): factores de actualización (alphas) altos en torno a $1 \cdot 10^{-2}$ que actualicen el fondo con un porcentaje elevado de cada imagen y un porcentaje menor del fondo anterior para no incorporar los movimientos del fondo rápidamente, y desviaciones elevadas en un rango de

[30:40] que soporten la multimodalidad del fondo. Siendo ambos parámetros decisivos en los resultados finales, ya que la elección de un valor alpha un orden de magnitud diferente a la indicada o una desviación 20 veces menor o mayor a la del rango propuesto puede generar muy malos resultados. La elección del parámetro umbral de discriminación entre objeto y fondo es poco influyente en los resultados si los anteriores parámetros han sido elegidos correctamente, por ejemplo, se puede fijar a cualquier valor alrededor del 0.75.

- Los métodos no paramétricos (KDE): el parámetro más importante es la probabilidad del Kernel que para este tipo de secuencias debe ser de [0.004-0.005].
- Secuencias multimodales medias y bajas, con movimiento lento de objetos requieren:
 - Los métodos básicos: umbrales medios [20-60].
 - Los métodos paramétricos (Gamma, SG y MoG): factores de actualización (alphas) más bajos en torno a 10^{-4} - 10^{-3} y desviaciones altas en un rango de [20-30].
 - Los métodos no paramétricos (KDE): el parámetro más importante es la probabilidad de Kernel que para este tipo de secuencias debe ser mayor que en el caso anterior en torno a 0.007.
- Secuencias unimodales, con movimientos lentos en los objetos requieren:
 - Los métodos básicos: umbrales bajos [10-20].
 - Los métodos paramétricos (Gamma, SG y MoG): factores de actualización (alphas) muy bajos en torno a 10^{-6} - 10^{-5} para actualizar el fondo con un pequeño porcentaje de cada Imagen y un porcentaje elevado del fondo anterior y, desviaciones muy pequeñas [5-10] son suficientes para soportar la unimodalidad del fondo.
 - Los métodos no paramétricos (KDE): el parámetro más importante es la probabilidad de Kernel que para las secuencias unimodales pueden encontrarse alrededor de 0.009. los métodos más adecuados en función de la modalidad del fondo.

Como se ha señalado en los resultados anteriores, en secuencias de fondo unimodal o multimodal bajo y medio, los métodos que detectan correctamente el objeto y el fondo son: Gamma, MoG y SG.

Para modelar fondos complejos multimodales, los métodos que ofrecen mejores resultados manteniendo un compromiso de detección correcta de objeto y fondo son: Gamma, SG, Mediana e Histogramas. Estos últimos resultados que podrían no ser los esperados al inicio del proyecto, son totalmente coherentes si tenemos en cuenta la complejidad de los fondos multimodales altos que se han utilizado para probar los métodos; aún hoy, no se han desarrollado algoritmos de sustracción de

fondo que no detecten movimientos del fondo como parte del fondo cuando estos píxeles se mueven más que los píxeles pertenecientes a los objetos en movimiento.

5.5.2.2 *SG vs SG con mejoras*

La mejora del método SG tiene como objetivo eliminar píxeles en movimiento pertenecientes al fondo (por ejemplo, debidos a movimientos de árboles). Esta mejora (explicada en detalle en la Sección 3.2.1.1) consiste en aplicar a las máscaras del frente, calculadas con el método SG sin mejoras, un *filtro* espacial mediante una ventana 5 x 5 alrededor de los píxeles detectados como frente. Si el píxel detectado como frente posee una varianza inferior a la de los píxeles de alrededor (ventana 5 x 5), se considera que el píxel se ha movido a uno de sus píxeles vecinos y se marca como fondo.

Para comparar el método SG con y sin mejora, se han obtenido los indicadores de calidad descritos en la Sección 5.3, para las secuencias de prueba. Los resultados obtenidos al aplicar la SG con y sin mejora se muestran a continuación en la Tabla 5-19:

<i>GAUSSIANA SIMPLE SIN/CON MEJORAS</i>							
<i>ID SECUENCIA</i>	<i>TIPO</i>	<i>P0</i>	<i>P1</i>	<i>R0</i>	<i>R1</i>	<i>FS0</i>	<i>FS1</i>
S1	SIN FILTRO	98.7	87.8	99.7	55.2	99.2	77.3
	CON FILTRO	99.1	89.1	99.8	69	99.4	77.7
S2	SIN FILTRO	97.9	73.3	99.7	29.4	98.8	41.9
	CON FILTRO	97.6	83.9	99.9	16.3	98.8	27.3
S3	SIN FILTRO	95.2	40.1	95.2	39.9	95.2	40.1
	CON FILTRO	93.4	67.3	99.5	11.8	96.3	20.2
S4	SIN FILTRO	92.3	63.7	99.5	88	95.8	15.4
	CON FILTRO	91.7	78.7	99.9	7	95.6	1.4
S5	SIN FILTRO	95.2	59.6	95.8	56.1	95.5	57.8
	CON FILTRO	93.9	84.6	99.1	41.7	96.4	55.8
S6	SIN FILTRO	99.5	97.1	99.9	86	99.7	91.2
	CON FILTRO	99.4	98.2	99.9	83.7	99.6	90.4
S7	SIN FILTRO	99.6	69.1	99.6	62.5	99.6	65.6
	CON FILTRO	99.2	87.2	99.9	39.2	99.6	54.1

Tabla 5-19. Parámetros óptimos para el método Gaussiana simple con y sin filtro

Para comparar con mayor claridad los resultados de los algoritmos SG sin y con mejora (sin y con filtro) expuestos en la Tabla 5-19, se han realizado los diagramas de precisión y *recall*. Los diagramas de barras obtenidos son los siguientes:

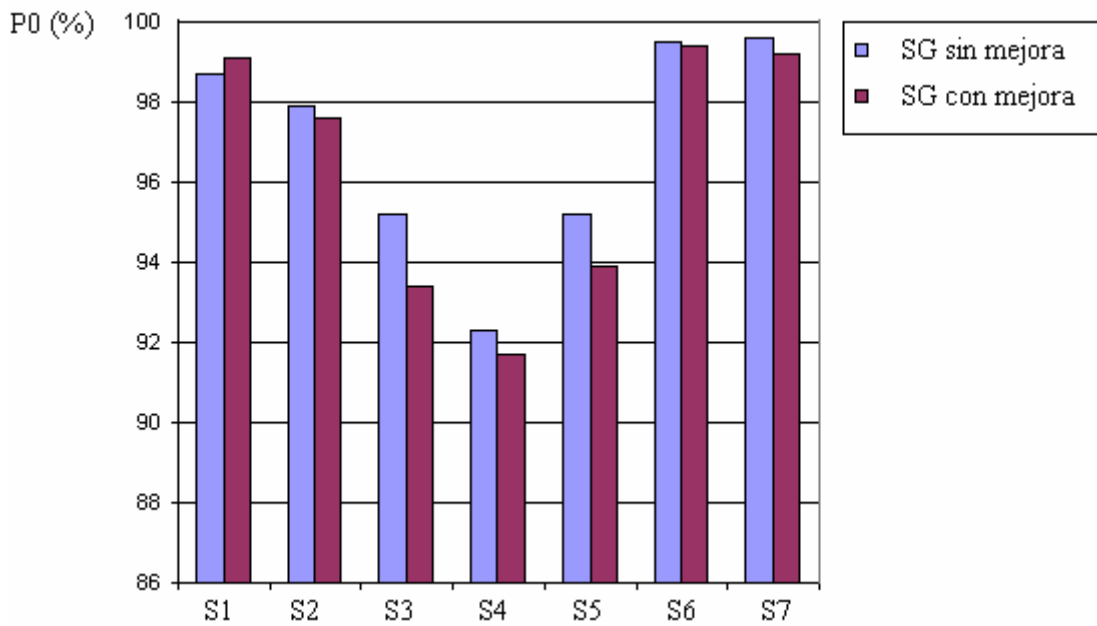


Figura 5-25: Diagrama de barras de la precisión en cero expresada en % por cada secuencia procesada por cada método SG con y sin mejora.

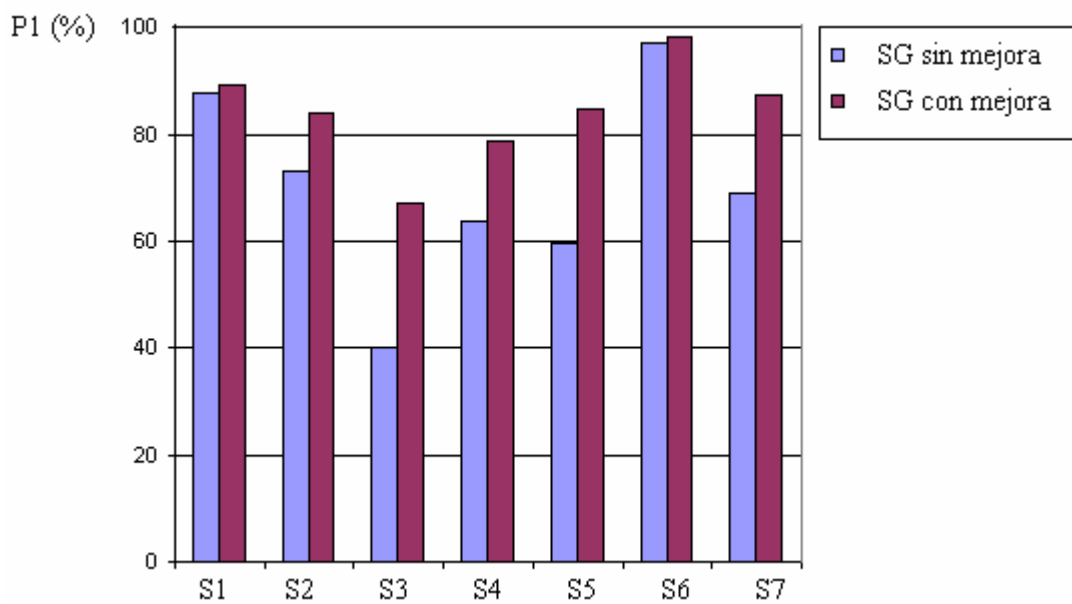


Figura 5-26: Diagrama de barras de la precisión en uno expresada en % por cada secuencia procesada por cada método SG con y sin mejora.

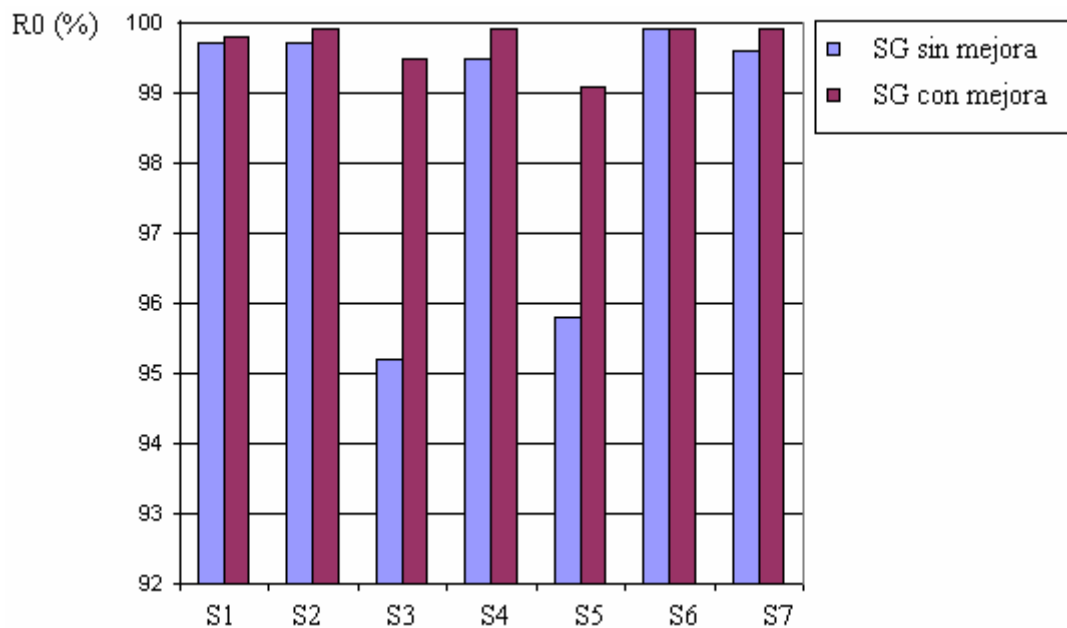


Figura 5-27: Diagrama de barras del *recall* en ceros expresada en % por cada secuencia procesada por cada método SG con y sin mejora.

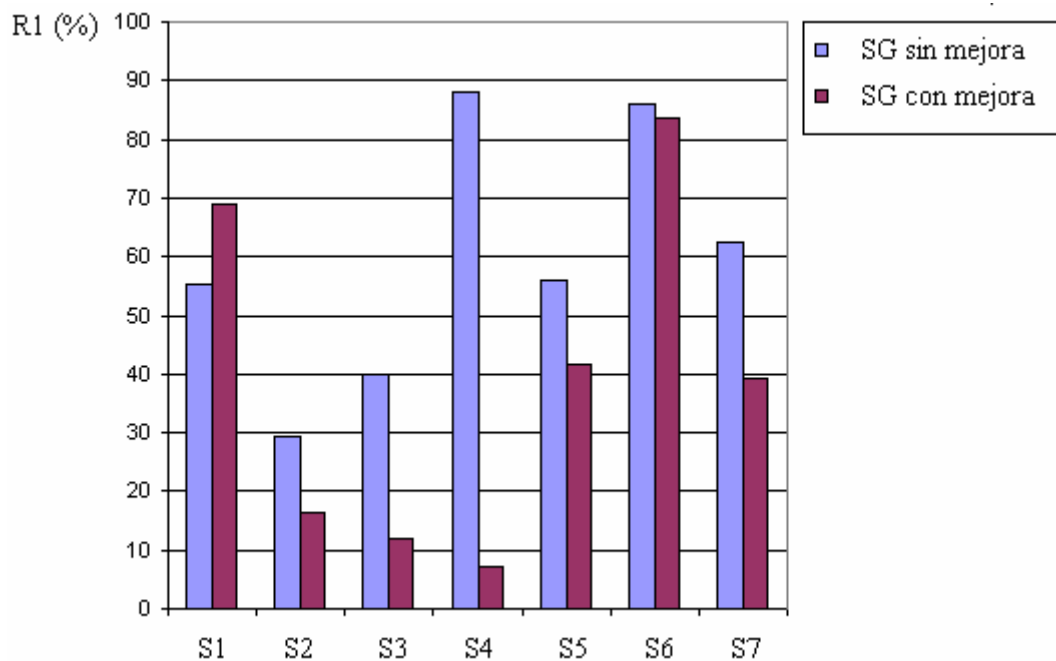


Figura 5-28: Diagrama de barras del *recall* en unos expresada en % por cada secuencia procesada por cada método SG con y sin mejora.

A la vista de los diagramas de barras, podemos concluir que el **método SG con mejora** ofrece **mayores porcentajes en precisión** en la detección del objeto ya que se aciertan más píxeles con valor uno del total que detecta el sistema con valor uno. También presenta **mayor recall** en el fondo debido a que se aciertan más píxeles con valor cero con respecto al total de píxeles de valor cero del *ground truth*. Esto significa que el método con mejora de la SG clasifica mejor los píxeles

del fondo que se mueven como píxeles de fondo (de valor igual a cero). Y la precisión en unos es mayor porque el total de píxeles con valor uno detectados, aún siendo mayor que el *ground-truth*, es menor que el del método sin mejoras.

El método de **SG sin mejoras**, posee mayores porcentajes de **precisión en BG** porque, aunque clasifica correctamente menos píxeles de fondo, una mayor proporción de los que detecta lo son realmente. Y presenta mayores porcentajes de **recall en FG** porque se aciertan más píxeles con valor uno con respecto al total de píxeles con valor uno del *ground truth* que con el método sin mejoras. Lo cual se debe a una detección mejor del objeto en movimiento.

Para aclarar los resultados anteriores, se muestran las siguientes figuras y fórmulas de las máscaras a la salida del algoritmo SG con y sin mejora obtenidas en el instante 61 de la ejecución sobre la secuencia de prueba Silla.

Como se aprecia en la **Figura 5-29**, SG sin mejoras clasifica como frente píxeles en movimiento del fondo (lo que resulta en mayor recall pero baja precisión para el frente), mientras que SG con mejora clasifica correctamente el fondo, aunque a costa de perder píxeles de frente (lo que resulta en menor recall pero mayor precisión para el frente).

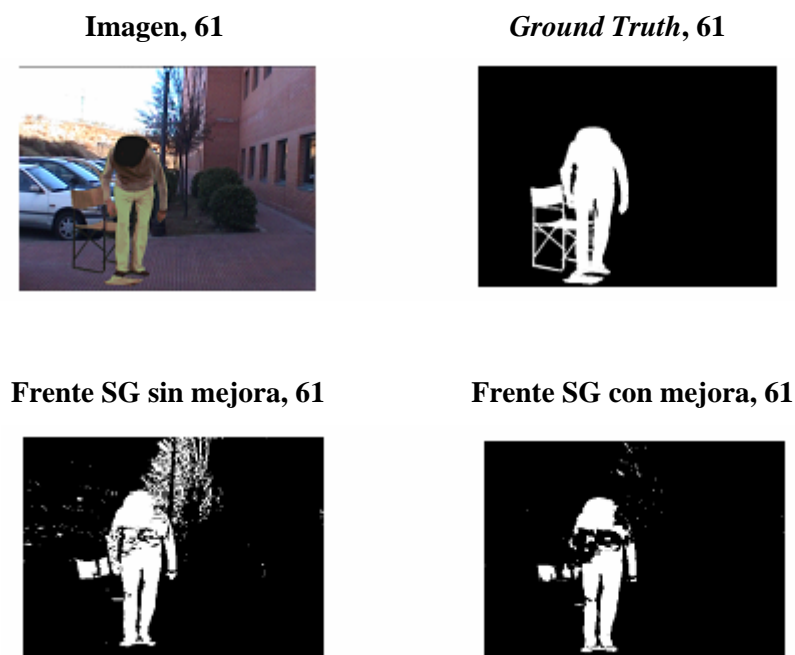
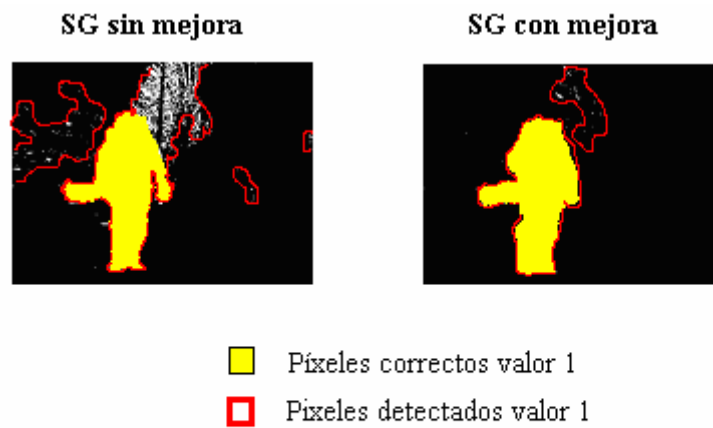
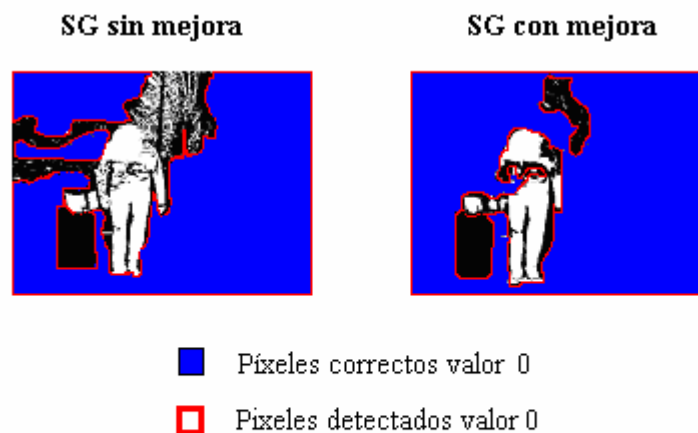


Figura 5-29: Imagen y Máscaras de Frente del algoritmo SG sin y con mejora capturadas en el instante 61.



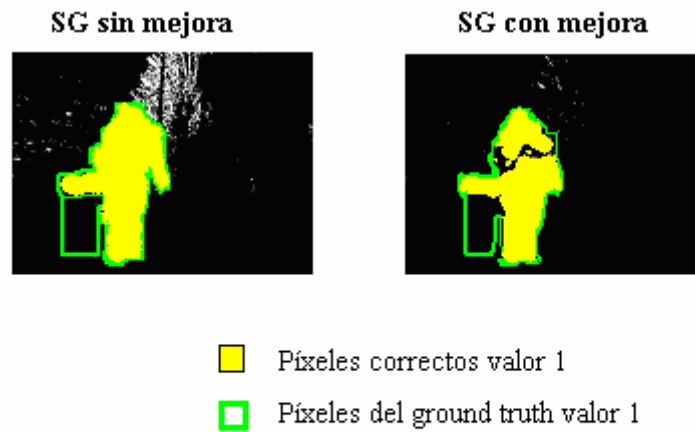
$$\text{Precision}(1) = \frac{\text{píxeles correctos valor 1}}{\text{píxeles detectados valor 1}}$$

Figura 5-30: Máscaras de Frente del algoritmo SG sin y con mejora divididas en píxeles correctos y detectados de valor 1.



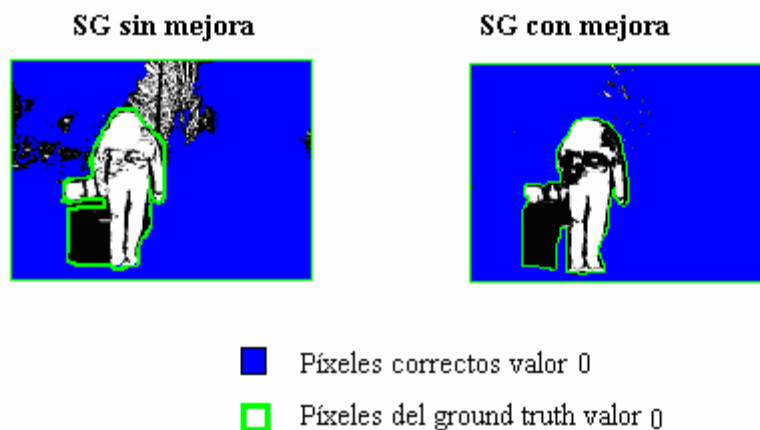
$$\text{Precision}(0) = \frac{\text{píxeles correctos valor 0}}{\text{píxeles detectados valor 0}}$$

Figura 5-31: Máscaras de Frente del algoritmo SG sin y con mejora divididas en píxeles correctos y detectados de valor 0.



$$\text{Recall}(1) = \frac{\text{píxeles correctos valor 1} \quad \color{yellow}{\square}}{\text{píxeles ground truth valor 1} \quad \color{green}{\square}}$$

Figura 5-32: Máscaras de Frente del algoritmo SG sin y con mejora divididas en píxeles correctos y verdaderos (*ground truth*) de valor 1.



$$\text{Recall}(0) = \frac{\text{píxeles correctos valor 0} \quad \color{blue}{\square}}{\text{píxeles ground truth valor 0} \quad \color{green}{\square}}$$

Figura 5-33: Máscaras de Frente del algoritmo SG sin y con mejora divididas en píxeles correctos y verdaderos (*ground truth*) de valor 0.

5.5.2.3 KDE vs KDE con mejoras

En el algoritmo KDE se han implementado dos mejoras la primera mejora (ver Sección 3.3.1.1), consiste en aplicar un filtro espacial (en una ventana 5x5) por cada píxel detectado como frente. La segunda mejora (ver Sección 3.3.1.2) consiste en utilizar dos colas FIFO de procesamiento (una con mecanismo selectivo y otra a ciegas) sobre las que calcular los píxeles pertenecientes al fondo y el frente. A través de la primera mejora se pretende reducir el elevado porcentaje de ruido presente en las escenas procesadas con el KDE sin mejora, y mediante la segunda mejora, se intenta incrementar la precisión en ceros y unos del método KDE sin mejora e incorporar al modelo de fondo los nuevos objetos que aparecen en la escena

Los resultados obtenidos al aplicar el método KDE sin mejoras (Tipo 1) y con mejoras (Tipo 2 para el filtrado espacial y Tipo 3 para las colas de procesamiento) se muestran para cada secuencia definida por el identificador de secuencia (S1,S2,..) en la siguiente tabla (Tabla 5-20). Estos resultados recogen las medidas definidas en el apartado 5.3 : precisión en BG (P0, precisión en cero), precisión en FG (P1, precisión en uno), *Recall* en BG (R0, recall en cero), *Recall* en FG (R1, recall en uno), *Fscore* en FG (FS1, *Fscore* en unos), *Fscore* en BG (FS0, *Fscore* en ceros), todas ellas expresadas en tanto por cien.

<i>KDE CON Y SIN MEJORAS</i>								
<i>ID SECUENCIAS</i>	<i>TIPO</i>	<i>Tiempo</i>	<i>P0</i>	<i>P1</i>	<i>R0</i>	<i>R1</i>	<i>FS0</i>	<i>FS1</i>
S1	1	1.07	98.7	14.5	90.5	55.3	94.4	22.3
	2	12.699	97.6	9.5	94.7	18.9	96.1	12.6
	3	3.89	97.7	8.5	92.3	24.6	94.9	12.7
S2	1	1.121	96.5	10.9	87.1	33.4	91.5	16.5
	2	21.5693	95.5	6.1	94.1	7.9	94.8	6.9
	3	4.9 89	95.6	5.6	89.8	12.7	92.6	7.8
S3	1	0.523	96.5	33.7	88.4	64.7	92.3	44.3
	2	22.6870	93.4	38.5	95.4	31.8	94.7	34.9
	3	2.9232	93.7	25.9	91.9	31.5	92.8	28.5
S4	1	1.081	93.8	13.4	84.8	29.4	89.1	18.4
	2	18.0864	92.8	9.5	91.9	01.9	92.3	10.2
	3	2.4964	92.8	8.9	86.9	15.9	89.8	11.4
S5	1	2.975	93.5	29.1	87.8	45.1	90.6	35.3
	2	33.8136	92.1	25.6	90.4	29.8	91.2	27.6
	3	8.4965	92.8	27.2	88.8	37.9	90.7	31.6
S6	1	4.357	97.2	23.6	96.8	25.9	97.1	24.8
	2	17.4196	96.8	15.8	97.2	14.2	96.9	14.9
	3	9.8707	97.2	23.6	96.8	25.9	97.1	24.7
S7	1	4.292	99.6	26.5	97.8	65.1	98.7	37.7
	2	19.5876	99.2	19.5	98.3	34.3	98.7	24.8
	3	12.1311	99.2	18.8	98.1	36.9	98.7	24.9

Tabla 5-20 . Parámetros óptimos para el método KDE sin mejoras, con la mejora del filtro y con la mejora de las dos colas.

Para comparar gráficamente los resultados obtenidos de los tres tipos de métodos realizados del KDE (expuestos en la tabla anterior) se muestra un diagrama de barras donde cada secuencia se expresa numerada con el nombre de S1, S2, etc.:

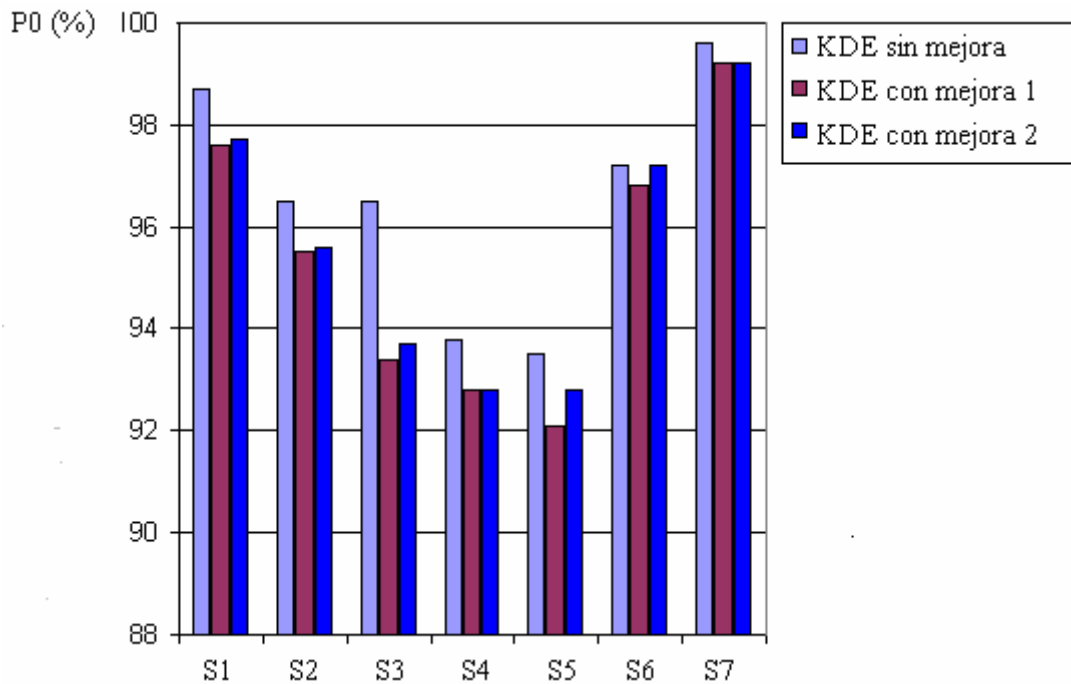


Figura 5-34: Diagrama de barras de la precisión en cero expresada en % por cada secuencia procesada por cada método KDE con y sin mejoras.

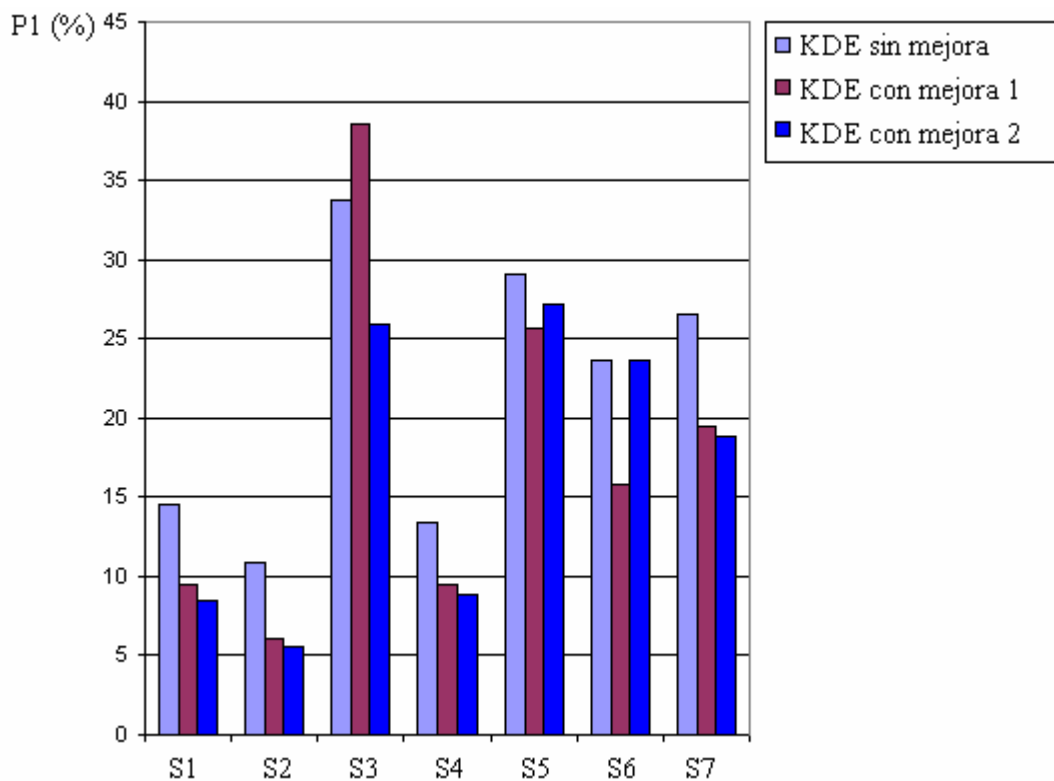


Figura 5-35: Diagrama de barras de la precisión en uno expresada en % por cada secuencia procesada por cada método KDE con y sin mejoras.

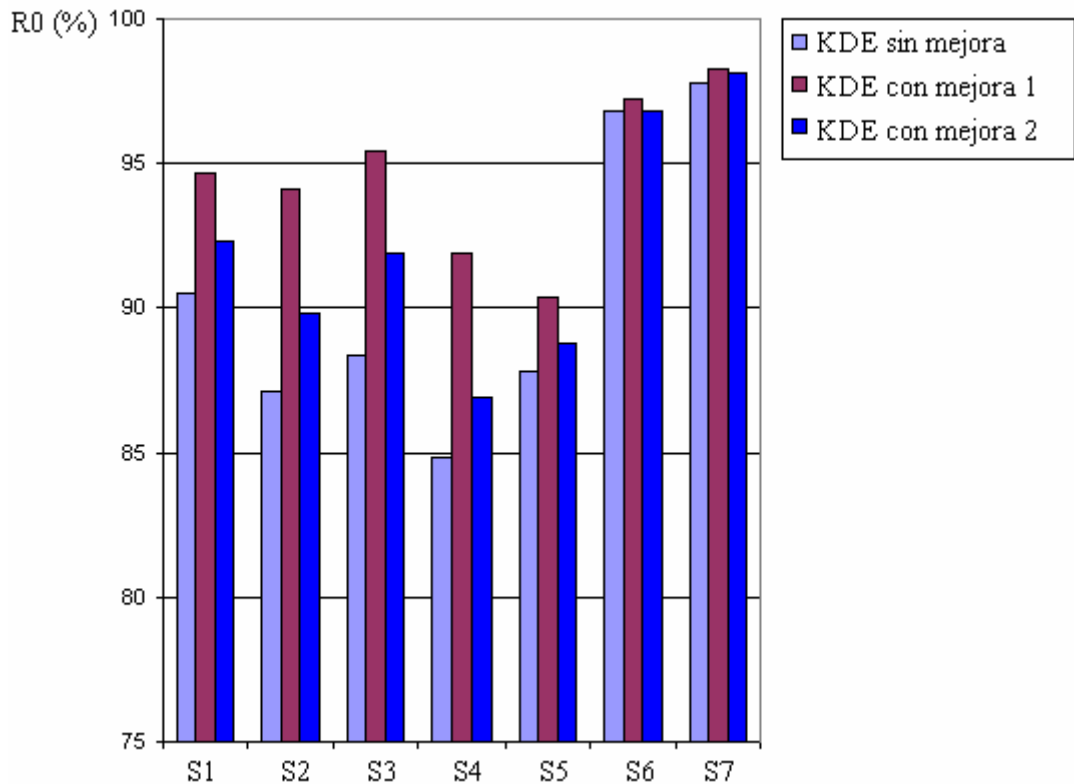


Figura 5-36: Diagrama de barras del *recall* en cero expresada en % por cada secuencia procesada por cada método KDE con y sin mejoras.

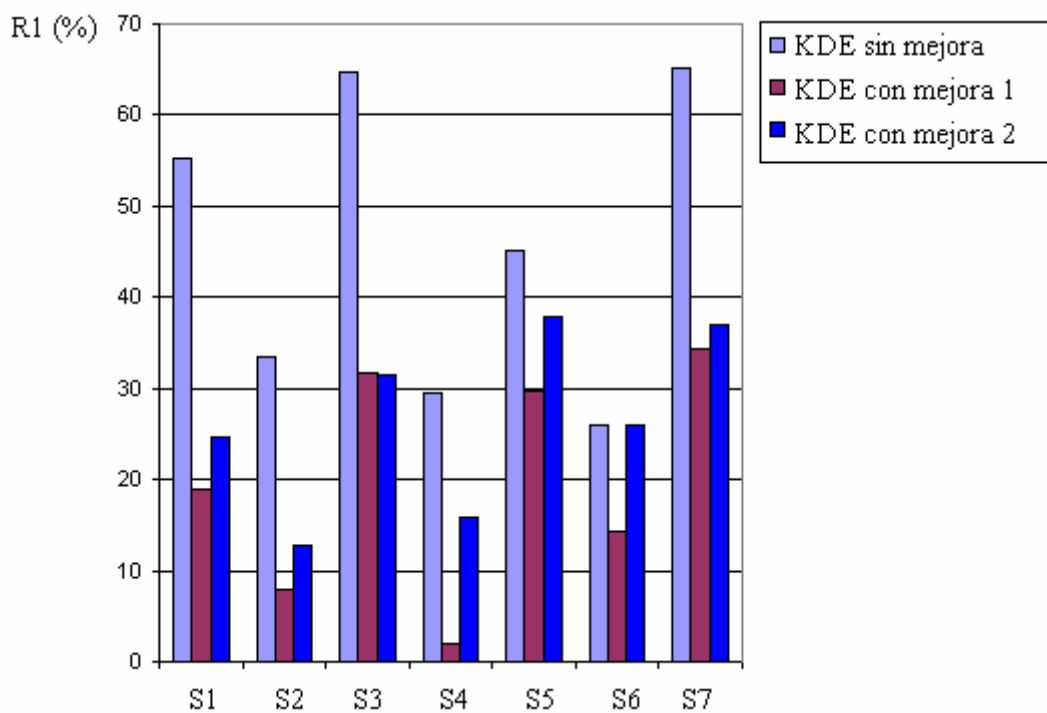


Figura 5-37: Diagrama de barras del *recall* en uno expresada en % por cada secuencia procesada por cada método KDE con y sin mejoras.

Los resultados que extraemos a la vista de las gráficas y las tablas se resumen en estas dos tablas:

En la Tabla 5-21 se comparan los resultados del KDE con y sin mejoras. Como se puede observar, con el algoritmo **sin mejoras** existe **mayor precisión en la detección del objeto**, es decir, mayor acierto en unos y hay menor error en el fondo, es decir, menos ceros erróneos respecto al total detectados (los pocos ceros que detectan se detectan correctamente). Sin embargo, con el algoritmo **con mejoras** se consigue acertar más elementos en movimiento del fondo(se **incrementa el recall en ceros**) pero se pierde *recall* en la máscara del frente (en los objetos en movimiento) y se equivocan menos unos (es decir, los pocos unos que detecta el algoritmo de las máscaras en movimiento son correctos).

<i>KDE sin mejora</i> ■	<i>KDE con mejora</i> ■ ■
+ TP (se aciertan + unos)	+ TN (se aciertan + ceros)
- FN (se equivocan - ceros)	- FP (se equivocan - unos)

Tabla 5-21 . Resultados comparativos del método KDE con y sin mejoras.

En la Tabla 5-22 se comparan las dos mejoras del KDE. A la vista de los resultados, podemos concluir que el método con la **mejora 1** (filtro espacio-temporal de ventana de tamaño 5 x 5) es **mejor en todos los aspectos de detección** (mejor precisión y *recall* en BG y FG) que la mejora 2 (dos colas de prioridad), salvo por el tiempo de procesamiento, ya que la mejora 1 tarda aproximadamente entre cuatro y cinco veces más en procesar cada imagen que la mejora 2. No obstante, es importante tener en cuenta que en estas dos mejoras no se han calculado los parámetros óptimos (que serían para la mejora 1, el tamaño de la ventana y para la mejora 2, los tamaños de las colas de procesamiento), sino que se han probado unos valores fijos para estos parámetros que son posiblemente poco óptimos, sólo para tener una idea de su funcionamiento.

<i>KDE con mejora 1</i> ■	<i>KDE con mejora 2</i> ■
+ TP (se aciertan + unos), + TN (se aciertan + ceros)	-TP,-TN (se aciertan menos unos y ceros)
- FP (se equivocan - unos), - FN (se equivocan - ceros)	+FP,+FN (se equivocan mas unos y ceros)

Tabla 5-22 . Resultados comparativos del método KDE con las mejoras: mejora 1 (filtro espacio-temporal) y mejora 2 (dos colas de procesamiento).

Para comprender mejor estos resultados veamos un ejemplo de las máscaras obtenidas con los tres métodos del KDE sobre de la secuencia Tennis 1 capturadas en el instante 188.

Como se aprecia en la **Figura 5-38**, KDE sin mejoras clasifica como frente píxeles en movimiento del fondo (lo que resulta en mayor recall pero baja precisión para el frente), mientras que KDE con la mejora 1, clasifica correctamente el fondo a costa de perder píxeles de frente (lo que resulta en menor recall pero mayor precisión para el frente). La mejora 2, clasifica mejor los píxeles de fondo a costa de perder píxeles de objeto con respecto a método KDE sin mejoras, pero clasifica peor fondo y objeto con respecto a la mejora 1 (por este motivo, son más bajos la precisión y el *recall* en FG y BG que la mejora 1).

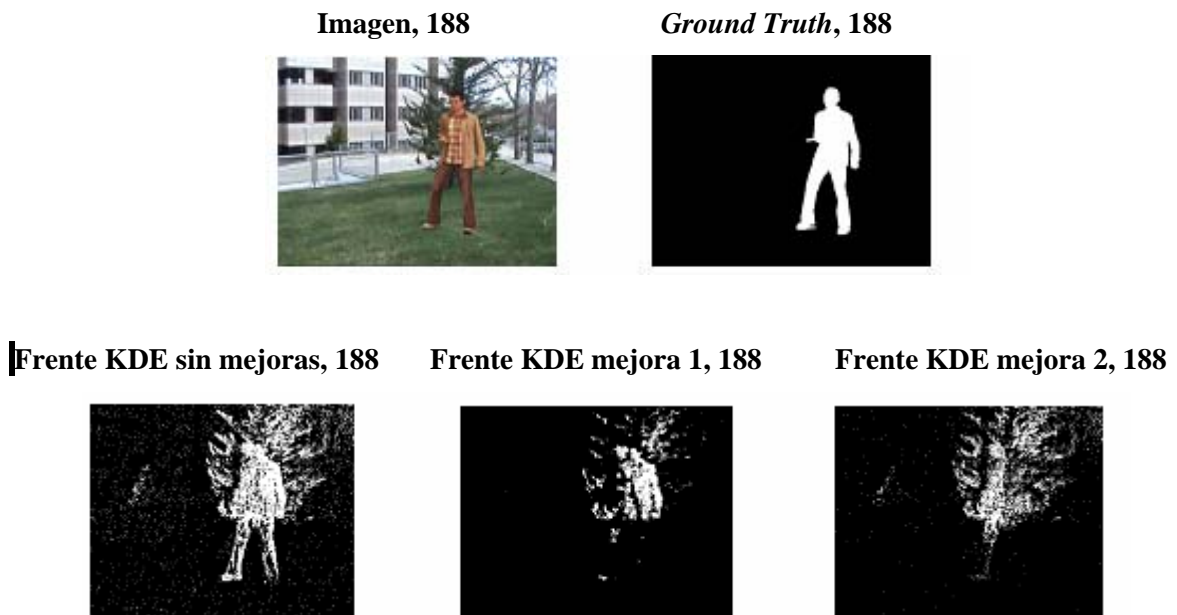


Figura 5-38: Imagen y Máscaras de Frente del algoritmo KDE sin y con mejoras capturadas en el instante 188.

6 Conclusiones y trabajo futuro

En este PFC se ha desarrollado un estudio de las técnicas de segmentación de objetos basadas en modelado de fondo con el objetivo de comparar estas técnicas y, en función de los requerimientos del sistema, elegir la técnica y los parámetros más adecuados y de esta manera contribuir al estado del arte.

Las técnicas de segmentación de objetos que se han estudiado se han clasificado en tres grupos: técnicas básicas de modelado de fondo (Diferencia de Frentes, Mediana e Histogramas), técnicas paramétricas (SG, MoG y Gamma) y técnicas no paramétricas (KDE). Además, se han desarrollado métodos de segmentación de objetos combinados a largo y corto plazo con dos de los métodos estudiados (Histogramas y SG) y se han mejorado dos de los métodos implementados (KDE y SG).

Finalmente, para evaluar y comparar los métodos se han utilizado diversas medidas objetivas y se han realizado gráficas, diagramas de barras y tablas para llegar a las conclusiones que exponemos a continuación.

La segmentación de objetos en movimiento sigue siendo un reto hoy en día en el procesamiento de secuencias de alta complejidad (secuencias clasificadas en este proyecto como multimodales altas) ya que los algoritmos de segmentación de objetos basados en modelado de fondo diseñados para modelar este tipo de secuencias, no son capaces de extraer un objeto en movimiento correcto si los píxeles de fondo presentan una variación más significativa que los del objeto.

Por otro lado, para realizar una detección del objeto adecuada y extraer un fondo correcto, un algoritmo de modelado de fondo requiere dos cuestiones importantes: la primera de ellas, es inicializar un modelo de fondo adaptado a la secuencia de video con el que comenzar el procesado y que no arrastre errores durante el análisis. La segunda cuestión, es un procesamiento capaz de actualizar el modelo de fondo con la multitud de variaciones a los que pueden estar expuestos los píxeles de fondo (por ejemplo, variaciones de la fuente de iluminación, movimientos de los píxeles de fondo en secuencias exteriores debido a árboles u otros objetos que deben discriminarse como parte del fondo u objetos que aparecen o desaparecen del fondo).

No obstante, en secuencias interiores (unimodales) o exteriores con bajo porcentaje de variación o *ruido* (clasificadas en este PFC como secuencias multimodales medias o bajas) los algoritmos de extracción de fondo paramétricos, por orden de máxima precisión: en primer lugar, el método Gamma, después MoG y, por último, SG, ofrecen muy buenos resultados ya que son capaces de

detectar correctamente (con un alto porcentaje de acierto en torno al 0.9) el objeto en movimiento y extraer un modelo de fondo adecuado.

Es muy importante estudiar los parámetros de cualquier modelo de segmentación de objetos que se utilice para que el modelo sea capaz de distinguir correctamente entre frente y fondo y, evitar la propagación de errores en el análisis de la secuencia de video a procesar.

Las operaciones de mejora que se realicen en una etapa de modelado de fondo no deben ser demasiado costosas computacionalmente para permitir una ejecución más o menos rápida (en función de los requerimientos del sistema). Ya que, como se ha observado en los resultados de las mejoras del método KDE, puede no ser apropiado incrementar excesivamente el coste computacional para mejorar levemente el porcentaje de aciertos; no debemos olvidar que el modelado de fondo es la primera etapa de análisis en multitud de sistemas y si esta etapa añade un excesivo coste computacional, el rendimiento global del sistema puede empeorar gravemente.

6.1 Trabajo futuro

La segmentación de objetos mediante modelado de fondo es un campo que ofrece multitud de posibilidades por investigar: el modelo matemático para representar el fondo de la escena, los métodos de inicializar y actualizar el modelo de fondo, así como las técnicas anteriores o posteriores al modelado que pueden mejorar los resultados obtenidos en la detección del objeto en movimiento y la extracción del fondo.

Una de las posibles líneas de trabajo futuro podría ir encaminada a mejorar las operaciones de inicialización y actualización del fondo de los algoritmos de modelado de fondo implementados.

Para mejorar el modelo de fondo inicial, se propone añadir al algoritmo de modelado de fondo, una etapa inicial que generase un fondo adecuado a la secuencia a procesar, por ejemplo, ejecutando el método de modelado que se vaya a utilizar sobre un conjunto de imágenes de la escena u otro modelo de segmentación de objetos más sencillo (Mediana, Histogramas o SG).

Para mejorar la actualización, se podría desarrollar una etapa de actualización de nivel superior que utilizase la información de las máscaras del frente y el fondo generados por el método de segmentación para decidir si los píxeles detectados como frente pertenecen realmente al frente o son píxeles en movimiento del fondo.

Por otro lado, todos los métodos de segmentación pueden mejorarse añadiendo etapas de pre-procesado para eliminar ruido de fondo y, etapas de post-procesado para regenerar las mascararas de objeto, por ejemplo, utilizando un análisis de componentes conexas [90][91].

También sería adecuado combinar métodos que pueden realizarse a largo plazo (KDE o Histogramas) con métodos sencillos a corto plazo (Gamma, SG o MoG) ya que como se ha observado en los resultados, la fusión de los métodos en secuencias de análisis complejas puede mejorar la detección del objeto y la extracción del fondo.

Una forma de encontrar los parámetros óptimos (con máximo porcentaje de acierto en la detección del objeto y fondo) sería diseñar una arquitectura similar a la utiliza en la evaluación de este PFC, añadiendo realimentación, de modo que, tras la ejecución del método de segmentación de objetos con un rango de parámetros fijo, pudiese calcular los parámetros óptimos y pasárselos de nuevo al método hasta encontrar los parámetros más adecuados de la secuencia de video analizada.

Los tiempos de procesamiento pueden mejorarse optimizando los códigos e implementándolos utilizando otros lenguajes de programación más óptimos para realizar aplicaciones en tiempo real y menos costosos que Matlab (utilizado en este PFC), como por ejemplo C/C++.

Referencias

- [1] Tsaig, Y. and Averbuch, A.: Automatic Segmentation of Moving Objects in Video Sequences: A Region Labeling Approach. *IEEE Trans. Circuits Syst. for Video Technol.* 12 (2002) 597-612.
- [2] Zhi Liu, Yu Luy y Zhaoyang Zhang. Real-time spatiotemporal segmentation of video objects in the H.264 compressed domain. *Journal of Visual Communication and Image Representation*, N°18, Pag.275-290, 2007.
- [3] J. Heikkila and O. Silven: A real-time system for monitoring of cyclists and pedestrians in: *Second IEEE Workshop on Visual Surveillance Fort Collins, Colorado (Jun.1999)* pp. 74-81.
- [4] Fan, J.P.[Jian-ping], Ji, Y.[Yongchang], Wu, L. D. [Li-De], Automatic Moving Object Extraction toward Content-Based Video Representation and Indexing, *JVCIR(12)*, N°3, September 2001, pp. 306-347.
- [5] Liang, L., Lu, H., Xue, X., Tan, Y.P.: Program segmentation for tv videos. In: *ISCAS, IEEE International Symposium on Circuits and Systems. Volume 2. (2005)* 1549–1552.
- [6] Background subtraction techniques: a review. *Systems, Man and Cybernetics M Piccardi - 2004 IEEE International Conference on*, 2004.
- [7] Cheung S. C., Kamath C.: Robust techniques for background subtraction in urban trafficvideo. *Video Communications and Image Processing. SPIE Electronic Imaging, San Jose, UCRL-JC-153846, UCRL-CONE-200706 (2004)*
- [8] Rita Cucchiara: *People Surveillance, VISMAL Palermo 2006.*
- [9] Bing Leng and Qionghai Dai: Video object segmentation based on accumulative frame difference: *PCE Picture Coding Sympoium (Nov. 2007).*
- [10] Zhen Tang, Zhenjiang Miao, Yanli Wan: Background Subtraction Using Running Gaussian Average and Frame Difference. *ICEC 2007.*
- [11] Wren, Azarbayejani, Darrell, Pentland, 1997. Pfinder: Real-time tracking of the human body. *PAMI.*
- [12] Stauffer, Grimson, 1999. Adaptive background mixture models for real-time tracking. *CVPR.*
- [13] 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.
- [14] Mittal A., Paragios N. Motion-based background subtraction using adaptive kernel density estimation. In *Proceedings of the Int. Conf. Comp Vision and Patt Recog (CVPR) 2004*; 302-309.
- [15] H. Dias, J. Rocha, P. Silva, C. Leao, “Distributed Surveillance System”, *Conference on Artificial Intelligence, 2005. EPIA 2005*, pp 257-261
- [16] M. Valera and S.A. Velastin, “Intelligent Distributed Surveillance Systems”, *IEEE Proc.-Vis. Image Signal Process.*, Vol. 152, No. 2, April 2005.
- [17] M. M. Chang, M. I. Sezan, and A. M. Tekalp, “An algorithm for simultaneous motion estimation and scene segmentation,” in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing*, vol. V, Adelaide, Australia, Apr. 1994, pp. 221–224.
- [18] Lienhart, R., Kuhmunch, C., Effelsberg, W.: On the detection and recognition of television commercials. In: *International Conference on Multimedia Computing and Systems*, pp. 509–516 (1997).
- [19] O. Sukmarg, K.R. Rao, Fast object detection and segmentation in MPEG compressed domain, in: *Proc. IEEE TENCON, Kuala Lumpur, Malaysia, 2000*, vol. 3, pp. 364–368.
- [20] Kentaro T, John K, Barry B, Brian M. Wallflower: Principles and Practice of Background Maintenance, *ICCV, Seventh Int. Conf. On Com Vision (ICCV'99) - 1999*; 1: 255-261.

- [21] Elgammal A, Duraiswami R, Harwood D, Davis LS. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. Proc. IEEE Jul. 2002; 90(7): 1151-1163.
- [22] Harville M., Gordon G., Woodfill J. Foreground Segmentation Using Adaptive Mixture Models in Color and Depth. event, IEEE Workshop on Detection and Recog of Events in Video (EVENT'01),2001; 3-12.
- [23] C. Stauffer and W.E.L. Grimson. Adaptive background mixture models for real-time tracking. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages II: 246–252, 1999.
- [24] T. Horprasert, D. Harwood, and L.S. Davis. A statistical approach for real-time robust background subtraction and shadow detection. In Proceedings of the Frame-Rate Workshop, 1999.
- [25] D. Gutches, M. Trajković, E. Cohen-Solal, D. Lyons, and A. K. Jain. A background model initialization algorithm for video surveillance. In IEEE International Conference on Computer Vision, 2001.
- [26] Cheung S-C, Kamath C. Robust techniques for background subtraction in urban traffic video. In S. Panchanathan and B. Vasudev, editors, Proc Elect Imaging: Visual Comm Image Proce 2004 (Part One) SPIE, 2004; 5308: 881-892.
- [27] Radke RJ, Andra S, Al-Kofahi O, Roysam B. Image Change Detection Algorithms: A systematic survey, image processing, IEEE Trans on March 2005, 14(3): 294-307.
- [28] Barron J, Fleet D, Beauchemin S. Performance of optical flow Techniques. Int J Comp Vision 1994; 12(1): 43-77.
- [29] Blake A, Isard M. Active Contours. Springer Verlag, 1999.
- [30] Wren CR, Azarbayejani A, Darrell T, Pentland A, Pfinder: Real time tracking of the human body. IEEE Trans Pat Anal Mach Intel 1997; 19(7): 780-785.
- [31] Hartley R., Zisserman A. Multiple view geometry in computer vision. Cambridge University Press, 2000.
- [32] Lillestrand R. Techniques for change detection. IEEE Trans. On Computers 1972; 21(7): 654-659.
- [33] Ulstad MS. An algorithm for estimating small scale differences between two digital images. Pat Recog 1973; 1(5): 323-333.
- [34] Dai X, Khorrani S. The effects of image misregistration on the accuracy of remotely sensed change detection. IEEE Trans. Geoscience Remote Sensing September 1998; 36(5): 1566-1577.
- [35] Pizarro O, Singh H. Toward large-area mosaicing for underwater scientific applications. IEEE J Oceanic Eng October 2003; 28(4):651-672.
- [36] Toth D., Aach T., Metzler V., Bayesian spatio-temporal motion detection under varying illumination illumination-invariant change detection. In Proc. of X EUSIPCO, Tampere, Finland, 2000; 3-7.
- [37] Toth TAD, Metzler V. Illumination-invariant change detection. In The 4th IEEE Southwest Symposium on Image Analysis and Interpretation, April 2000; 3.
- [38] Hager GD., Belhumeur PN. Efficient region tracking with parametric models of geometry and illumination. IEEE Trans. Patt Anal. Machine Intell 1998; 20(10): 1025-1039.
- [39] Neri A, Colonnese S, Russo G, Talone P. Automatic moving object and background separation. Signal Processing 1998; 66(2): 219-232.
- [40] Skifstad K, Jain R. Illumination independent change detection for real world image sequences. Comp Vision, Graphics Image Process June 1989; 46: 387-399.
- [41] Hotter M, Mester R, Muller F. Detection and description of moving objects by stochastic modelling and analysis of complex scenes. Signal Process: Image Comm 1996; 8: 281-293.
- [42] Aach T, Kaup A, Mester R. Statistical model-based change detection in moving video. Signal Process 1993; 31: 165-180.
- [43] Liu SC, Fu CW, Chang S. Statistical change detection with moments under time-varying illumination. IEEE Transactions on Image Processing 1993; 7(9): 1258-1268.

- [44] Mech R, Wollborn M. A noise robust method for 2D shape estimation of moving objects in video sequences considering a moving camera. *Signal Process* 1998; 66(2): 203-217.
- [45] Cavallaro A, Ebrahimi T. Video object extraction based on adaptive background and statistical change detection. In *Proc. of SPIE Visual Comm Image Process (VCIP) 2001*; 465-475.
- [46] Cutler R, Davis L. View-based detection and analysis of periodic motion in: *Int Conf Patt Recog Brisbane, Australia (Aug.) 1998*; 1: 495-500.
- [47] Stauffer C, Grimson W. Learning Patts of activity using realtime Tracking. In *IEEE Trans. on Patt Anal Mach Intel Aug 2000*; 22:747-57.
- [48] Haritaoglu I., Harwood D., Davis LS. W4: Who? when? where? what? a real time system for detecting and tracking people in. *Third Face and Gesture Recog Conf. (Apr.) 1998*; 222-227.
- [49] Harville M. A framework for high-level feedback to adaptive, perpixel, mixture-of-gaussian background models. *European Conf. Comp Vision 2002*; 3: 543-560.
- [50] Magee DR. Tracking multiple vehicles using foreground, background, and motion models. In *Proceedings of the Statistical Methods in Video Processing Workshop, (Copenhagen, Denmark), June 2002*; 7-12.
- [51] 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.
- [52] Cavallaro A, Ebrahimi T. Change detection based on color edges, circuits and systems. *The 2001 IEEE Int Symposium, May 2001*; 2: 141-144.
- [53] Costantini R., Ramponi G., Bracamonte J., et al. Countering illumination variations in a video urveillance environment. In *Proc. of SPIE, Electronic Imaging Conf., San Jose, USA, 2001*; 4304: 85- 97.
- [54] Pless R., Larson J., Siebers S., Westover B. Evaluation of local models of dynamic background. In *Proceedings IEEE Conf Comp Vision Patt Recog, 2003*; 2: 73-78, (Madison,WI).
- [55] Boulton T., Ali Erkin A., Lewis P, et al. Frame-rate omnidirectional surveillance and tracking of camuaged and occluded targets. In *Proceedings Second IEEE Workshop on Visual Surveillance June 1999*; 48-55. (Fort Collins, CO).
- [56] Eveland C, Konolige K, Bolles RC. Background Modeling for Segmentation of Video-Rate Stereo Sequences. *CVPR, 1998 IEEE Comp Society Conf. on Comp Vision and Patt Recog (CVPR'98) 1998*; 266-271.
- [57] Ivanov Y, Bobick A, Liu J. Fast lighting independent background subtraction. *Int. J Comp Vision 2000*; 37(2): 199-207.
- [58] Gordon G., Darrell T., Harville M., Woodfill J. Background Estimation and Removal Based on Range and Color. *CVPR, 1999 IEEE Comp Society Conf. on Comp Vision and Patt Recog (CVPR'99) 1999*; 2: 2459.
- [59] Cristani M., Bicego M., Murino V. Multi-level background initialization using Hidden Markov Models. In *First ACM SIGMM Int. workshop on Video surveillance 2003*; 11-20.
- [60] Ralph Ewerth, Bernd Freisleben: Frame difference normalization: an approach to reduce error rates of cut detection algorithms for MPEG videos. *ICIP (2) 2003*: 1009-1012.
- [61] Koller D., Weber J., Huang T., Malik J., Ogasawara G., Rao B. Russell S. Toward robust automatic traffic scene analysis in realtime. in *Proc. Int. Conf. Patt Recog, 1994*; 126-131.
- [62] Boulton T, Micheals R, Gao X, Eckmann M. Into the woods: Visual surveillance of non-cooperative camouflaged targets in complex outdoor settings. In *Proceedings of the IEEE, October 2001*; 1382- 1402.
- [63] Karmann K-P, Brandt AV, Gerl R. Moving object segmentation based on adaptive reference images. In *Signal Processing V: Theories and Application. Amsterdam. The Netherlands: Elsevier, 1990*.
- [64] Monnet A., Mittal A., Paragios N., Ramesh V. Background modeling and subtraction of dynamic scenes. *Proc Int. Conf. Comput. Vision, Nice, France, 2003*; 2:1305-1312.

- [65] Gordon G., Darrell T., Harville M., Woodfill J. Background Estimation and Removal Based on Range and Color. CVPR, 1999 IEEE Comp Society Conf. on Comp Vision and Patt Recog (CVPR'99) 1999; 2: 2459.
- [66] Jabri S., Duric Z., Wechsler H., Rosenfeld A. Detection and Location of People in Video Images Using Adaptive Fusion of Color and Edge Information. ICPR, 15th Int. Conf. on Patt Recog (ICPR'00) 2000; 4: 4627-4630.
- [67] Stauffer C, Grimson W. Learning Patts of activity using real time Tracking. In IEEE Trans. on Patt Anal Mach Intel Aug 2000; 22:747-57.
- [68] Javed O., Shafique K., Shah M. A hierarchical approach to robust background subtraction using color and gradient information. motion. Workshop.
- [69] Stenger B., Ramesh V., Paragios N., Coetsee F., Buhmann J. M. Topology Free Hidden Markov Models: Application to Background Modeling. ICCV, Eighth Int. Conf. on Computer Vision (ICCV'01) 2001; 1: 294-301.
- [70] Kim K, Chalidabhongse TH, Harwood D, Davis LS. Real-time foreground-background segmentation using codebook model. Real-Time Imaging 2005; 11(3): 172-185.
- [71] Butler D., Sridharan S., Bove VM Jr. Real-time Adaptive Background Segmentation. Acoustics, Speech, and Signal Processing. 2003. Proceedings. (ICASSP '03). 2003 IEEE Int. Conf. on April 2003; 3: 349-52.
- [72] Stauffer C., Grimson WEL. Adaptive Background Mixture Models for Real-Time Tracking. CVPR, 1999 IEEE Comput Society Conf. on Comp Vision and Patt Recog (CVPR'99) 1999; 2: 246-252.
- [73] Karmann K-P, Brandt A. Moving object Recog using and adaptive background memory. In Time-Varying Image Processing and Moving Object Recog. V. Cappellini, Ed. 2, Elsevier Science Publishers B.V. 1990; 289-307.
- [74] Horprasert T, Harwood D., Davis LS. A statistical approach for real-time robust background subtraction and shadow detection. Proc. IEEE Int'l Conf. Comp Vision, Frame Rate Workshop 1999; 1-19.
- [75] Long W, Yang YH. Stationary background generation: An alternative to the difference of two images, Patt Recog 1990; 23(12): 1351-1359.
- [76] Gloyer B., Aghajan HK, Siu KY, Kailath T. Video-based freeway monitoring system using recursive vehicle tracking. In Proc. Of IS&T-SPIE Symposium on Electronic Imaging: Image and Video Processing 1995.
- [77] Fuentes L., Velastin S. From tracking to advanced surveillance. In Proceedings of IEEE Int. Conf on Image Processing, (Barcelona, Spain), Sept 2003.
- [78] McFarlane N, Schofield C. Segmentation and tracking of piglets in images, Machine Vision Appl 1995; 8(3): 187-193.
- [79] Migdal J., Eric W., Grimson L. Background Subtraction Using Markov Thresholds, wacv-motion. IEEE Workshop on Motion and Video Computing (WACV/MOTION'05) 2005; 2: 58-65.
- [80] Jain R., Kasturi R., Schunk GB. Machine Vision. McGRAWHILL Int. Editions, 1995.
- [81] Durucan E, Ebrahimi T. Robust and illumination invariant change detection based on linear dependence for surveillance applications. In Proc. of X EUSIPCO, Tampere, Finland, 2000; 1041-1044.
- [82] Hotter M, Mester R, Muller F. Detection and description of moving objects by stochastic modelling and analysis of complex scenes. Signal Proces: Image Comm 1996; 8: 281-293.
- [83] K. McKoen, R. Navarro-Prieto, B. Duc, E. Durucan, F. Ziliani, and T. Ebrahimi, "Evaluation of video segmentation methods for surveillance applications," Proc. EUSIPCO, vol. II, pp. 1045-1048, Sep. 2000.
- [84] P. Correia and F. Pereira, "Objective evaluation of relative segmentation quality," Proc. ICIP, vol. 2, pp. 308-311, Sep. 2000.
- [85] M. Caramma, R. Lancini, and M. Marconi, "A perceptual PSNR based on the utilization of a linear model of HVS, motion vectors and DFT-3D," Proc. EUSIPCO, vol. IV, pp. 2185-2188, Sep. 2000.
- [86] Fabrizio Tiburzi, Marcos Escudero, Jesús Bescós y José M. Martínez, "A Corpus for Motion-based Video-object Segmentation", IEEE International Conference on Image

- Processing (Workshop on Multimedia Information Retrieval), ICIP'2008, SanDiego (USA), Oct. 2008. (accepted for presentation).
- [87] Varcheie, P.D.Z.[Parisa Darvish Zadeh], Sills-Lavoie, M.[Michael], Bilodeau, G.A.[Guillaume-Alexandre], "An Efficient Region-Based Background Subtraction Technique", CRV08(71-78).
- [88] F. El Baf, T. Bouwmans, B. Vachon, "Comparison of Background Subtraction Methods for a Multimedia Application", 14th International Conference on systems, Signals and Image Processing, IWSSIP 2007, pages 385-388, Maribor, Slovenia, June 2007.
- [89] Parks, Donovan H.; Fels, Sidney S., "Evaluation of Background Subtraction Algorithms with Post-Processing," IEEE Fifth International Conference on Advanced Video and Signal Based Surveillance, AVSS '08, pp.192-199, 1-3 Sept. 2008.
- [90] S.B. Gray, "Local Properties of Binary Images in Two Dimensions," *IEEE Transactions on Computers*, vol. 20, no. 5, pp. 551-561, May, 1971.
- [91] Fu Chang, Chun-Jen Chen, Chi-Jen Lu, A linear-time component-labeling algorithm using contour tracing technique, *Computer Vision and Image Understanding*, v.93 n.2, p.206-220, February 2004.
- [92] Monari, E.[Eduardo], Pasqual, C.[Charlotte], Fusion of background estimation approaches for motion detection in non-static backgrounds, AVSBS07(347-352).

Glosario

BGS	BackGround Subtraction techniques
FG	Foreground
BG	Background
SG	Simple Gaussian
MoG	Mixture of Gaussian
KDE	Kernel Density Estimation
HMM	Hidden Markov Model

Anexos

Anexo A

En este anexo se presentan unas tablas que contienen los resultados del análisis de los parámetros más significativos de los métodos de segmentación implementados.

Es necesario aclarar que para los métodos SG, MoG y KDE para los cuales se han encontrado varios parámetros óptimos, sólo hemos incluido en este anexo las tablas de los parámetros más significativos, debido a que el proceso de búsqueda de los parámetros óptimos ha requerido múltiples pruebas e incluir estas tablas conllevaba a aumentar considerablemente la memoria. Además las conclusiones extraídas de estas tablas se hayan resumidas en la memoria.

DIFERENCIA S1											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	4967032	451348	956243	24625697	0,963	0,083	0,832	0,321	0,893	0,132	1,025
40	2131295	363615	1043976	27461434	0,963	0,146	0,928	0,258	0,945	0,186	1,132
60	1431594	333821	1073770	28161135	0,963	0,189	0,952	0,237	0,957	0,21	1,168
80	983803	308738	1098853	28608926	0,963	0,239	0,967	0,219	0,965	0,229	1,194
100	500463	268749	1138842	29092266	0,962	0,349	0,983	0,191	0,973	0,247	1,22
120	500463	268749	1138842	29092266	0,962	0,349	0,983	0,191	0,973	0,247	1,22
140	371368	252508	1155083	29221361	0,962	0,405	0,987	0,179	0,975	0,249	1,223
160	284755	238036	1169555	29307974	0,962	0,455	0,99	0,169	0,976	0,247	1,222
450	489527	268749	1138842	29103202	0,962	0,354	0,983	0,191	0,973	0,248	1,221
500	186515	213380	1194211	29406214	0,961	0,534	0,994	0,152	0,977	0,236	1,213
550	127008	184145	1223446	29465721	0,96	0,592	0,996	0,131	0,978	0,214	1,192
600	105064	161187	1246404	29487665	0,959	0,605	0,996	0,115	0,978	0,193	1,17
650	97108	148186	1259405	29495621	0,959	0,604	0,997	0,105	0,978	0,179	1,157
700	91060	136845	1270746	29501669	0,959	0,6	0,997	0,097	0,977	0,167	1,145
750	85968	126592	1280999	29506761	0,958	0,596	0,997	0,09	0,977	0,156	1,134
800	81586	117589	1290002	29511143	0,958	0,59	0,997	0,084	0,977	0,146	1,124

DIFENCIA S2											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
50	3387456	361292	1046299	26205273	0,962	0,096	0,886	0,257	0,922	0,14	1,062
100	2474993	297805	1109786	27117736	0,961	0,107	0,916	0,212	0,938	0,142	1,08
150	2052859	266823	1140768	27539870	0,96	0,115	0,931	0,19	0,945	0,143	1,088
200	1729304	241458	1166133	27863425	0,96	0,123	0,942	0,172	0,951	0,143	1,094
250	1594361	230609	1176982	27998368	0,96	0,126	0,946	0,164	0,953	0,143	1,096
300	1366425	210956	1196635	28226304	0,959	0,134	0,954	0,15	0,957	0,141	1,098
350	1268771	202241	1205350	28323958	0,959	0,137	0,957	0,144	0,958	0,141	1,099

Anexos

400	1100660	186631	1220960	28492069	0,959	0,145	0,963	0,133	0,961	0,139	1,099
450	1003649	179805	1227786	28589080	0,959	0,152	0,966	0,128	0,962	0,139	1,101
500	961214	173375	1234216	28631515	0,959	0,153	0,968	0,123	0,963	0,136	1,099
550	900410	167345	1240246	28692319	0,959	0,157	0,97	0,119	0,964	0,135	1,099
600	844306	161749	1245842	28748423	0,958	0,161	0,971	0,115	0,965	0,134	1,099
650	792681	156466	1251125	28800048	0,958	0,165	0,973	0,111	0,966	0,133	1,099
700	745482	151434	1256157	28847247	0,958	0,169	0,975	0,108	0,966	0,131	1,098
750	701253	146679	1260912	28891476	0,958	0,173	0,976	0,104	0,967	0,13	1,097
800	660574	141972	1265619	28932155	0,958	0,177	0,978	0,101	0,968	0,128	1,096

DIFERENCIA S3											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	3815230	507324	770939	12221067	0,941	0,117	0,762	0,397	0,842	0,181	1,023
40	2650645	432231	846032	13385652	0,941	0,14	0,835	0,338	0,884	0,198	1,083
60	2282666	401077	877186	13753631	0,94	0,149	0,858	0,314	0,897	0,202	1,099
80	1997014	373203	905060	14039283	0,939	0,157	0,875	0,292	0,906	0,205	1,111
100	1576873	325662	952601	14459424	0,938	0,171	0,902	0,255	0,92	0,205	1,124
120	1576873	325662	952601	14459424	0,938	0,171	0,902	0,255	0,92	0,205	1,124
140	1416636	305010	973253	14619661	0,938	0,177	0,912	0,239	0,924	0,203	1,128
160	1279214	286572	991691	14757083	0,937	0,183	0,92	0,224	0,929	0,202	1,13
180	1128177	268756	1009507	14908120	0,937	0,192	0,93	0,21	0,933	0,201	1,134
200	1055983	254588	1023675	14980314	0,936	0,194	0,934	0,199	0,935	0,197	1,132
220	1055983	254588	1023675	14980314	0,936	0,194	0,934	0,199	0,935	0,197	1,132
240	963622	241299	1036964	15072675	0,936	0,2	0,94	0,189	0,938	0,194	1,132
260	879652	228817	1049446	15156645	0,935	0,206	0,945	0,179	0,94	0,192	1,132
280	879652	228817	1049446	15156645	0,935	0,206	0,945	0,179	0,94	0,192	1,132
300	807967	217073	1061190	15228330	0,935	0,212	0,95	0,17	0,942	0,188	1,131
320	807967	217073	1061190	15228330	0,935	0,212	0,95	0,17	0,942	0,188	1,131

DIFERENCIA S4											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	1613054	1014539	1682830	26689897	0,941	0,386	0,943	0,376	0,942	0,381	1,323
40	1052465	872303	1825066	27250486	0,937	0,453	0,963	0,323	0,95	0,377	1,327
60	893765	812676	1884693	27409186	0,936	0,476	0,968	0,301	0,952	0,369	1,321
80	780458	759250	1938119	27522493	0,934	0,493	0,972	0,281	0,953	0,358	1,311
100	626107	667962	2029407	27676844	0,932	0,516	0,978	0,248	0,954	0,335	1,289
120	626107	667962	2029407	27676844	0,932	0,516	0,978	0,248	0,954	0,335	1,289
140	569333	628721	2068648	27733618	0,931	0,525	0,98	0,233	0,955	0,323	1,277
160	520209	593488	2103881	27782742	0,93	0,533	0,982	0,22	0,955	0,311	1,266

DIFERENCIA S5											
---------------	--	--	--	--	--	--	--	--	--	--	--

Anexos

τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	1773848	614971	1433969	16809532	0,921	0,257	0,905	0,3	0,913	0,277	1,19
40	1114897	507020	1541920	17468483	0,919	0,313	0,94	0,247	0,929	0,276	1,206
60	966089	466852	1582088	17617291	0,918	0,326	0,948	0,228	0,933	0,268	1,201
80	861618	432864	1616076	17721762	0,916	0,334	0,954	0,211	0,935	0,259	1,194
100	717738	377709	1671231	17865642	0,914	0,345	0,961	0,184	0,937	0,24	1,178
120	717738	377709	1671231	17865642	0,914	0,345	0,961	0,184	0,937	0,24	1,178
140	663747	354874	1694066	17919633	0,914	0,348	0,964	0,173	0,938	0,231	1,17
160	617332	334580	1714360	17966048	0,913	0,351	0,967	0,163	0,939	0,223	1,162

DIFERENCIA S6											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
6	667598	119528	438757	14326117	0,97	0,152	0,955	0,214	0,963	0,178	1,141
8	686120	119533	438752	14307595	0,97	0,148	0,954	0,214	0,962	0,175	1,137
10	244989	109215	449070	14748726	0,97	0,308	0,984	0,196	0,977	0,239	1,216
12	244989	109215	449070	14748726	0,97	0,308	0,984	0,196	0,977	0,239	1,216
14	244989	109215	449070	14748726	0,97	0,308	0,984	0,196	0,977	0,239	1,216
16	114264	101522	456763	14879451	0,97	0,47	0,992	0,182	0,981	0,262	1,243
18	114264	101522	456763	14879451	0,97	0,47	0,992	0,182	0,981	0,262	1,243
20	104928	101520	456765	14888787	0,97	0,492	0,993	0,182	0,981	0,266	1,247
40	45955	89827	468458	14947760	0,97	0,662	0,997	0,161	0,983	0,259	1,242
60	36649	85073	473212	14957066	0,969	0,699	0,998	0,152	0,983	0,25	1,233
80	32046	81008	477277	14961669	0,969	0,717	0,998	0,145	0,983	0,241	1,225
100	28131	73896	484389	14965584	0,969	0,724	0,998	0,132	0,983	0,224	1,207
120	28131	73896	484389	14965584	0,969	0,724	0,998	0,132	0,983	0,224	1,207
140	27156	70760	487525	14966559	0,968	0,723	0,998	0,127	0,983	0,216	1,199
160	26384	67916	490369	14967331	0,968	0,72	0,998	0,122	0,983	0,208	1,191

DIFERENCIA S7											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	134240	151983	194585	30519512	0,994	0,531	0,996	0,439	0,995	0,48	1,475
40	46376	129837	216731	30607376	0,993	0,737	0,998	0,375	0,996	0,497	1,492
60	41277	121635	224933	30612475	0,993	0,747	0,999	0,351	0,996	0,477	1,473
80	38999	114450	232118	30614753	0,992	0,746	0,999	0,33	0,996	0,458	1,453
100	36117	102908	243660	30617635	0,992	0,74	0,999	0,297	0,995	0,424	1,419
120	36117	102908	243660	30617635	0,992	0,74	0,999	0,297	0,995	0,424	1,419
140	34895	98138	248430	30618857	0,992	0,738	0,999	0,283	0,995	0,409	1,405
160	33688	93812	252756	30620064	0,992	0,736	0,999	0,271	0,995	0,396	1,391

MEDIANA S1											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma

10	190946	491249	183507	22980698	0,992	0,72	0,992	0,728	0,992	0,724	1,716
20	50399	420494	254262	23121245	0,989	0,893	0,998	0,623	0,993	0,734	1,728
30	39729	359902	314854	23131915	0,987	0,901	0,998	0,533	0,992	0,67	1,662
40	32031	296295	378461	23139613	0,984	0,902	0,999	0,439	0,991	0,591	1,582
50	23641	228917	445839	23148003	0,981	0,906	0,999	0,339	0,99	0,494	1,484
60	16507	154375	520381	23155137	0,978	0,903	0,999	0,229	0,989	0,365	1,354
70	11766	117226	557530	23159878	0,976	0,909	0,999	0,174	0,988	0,292	1,28
80	8988	95508	579248	23162656	0,976	0,914	1	0,142	0,987	0,245	1,233
90	6554	74612	600144	23165090	0,975	0,919	1	0,111	0,987	0,197	1,184

MEDIANA S2											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	5944307	1073570	334021	23544742	0,986	0,153	0,798	0,763	0,882	0,255	1,137
20	3805468	938879	468712	25683581	0,982	0,198	0,871	0,667	0,923	0,305	1,228
30	2571838	826299	581292	26917211	0,979	0,243	0,913	0,587	0,945	0,344	1,288
40	1777691	723493	684098	27711358	0,976	0,289	0,94	0,514	0,957	0,37	1,327
50	1236961	624469	783122	28252088	0,973	0,335	0,958	0,444	0,965	0,382	1,347
60	856881	533863	873728	28632168	0,97	0,384	0,971	0,379	0,971	0,382	1,352
70	585122	448839	958752	28903927	0,968	0,434	0,98	0,319	0,974	0,368	1,341
80	392226	368991	1038600	29096823	0,966	0,485	0,987	0,262	0,976	0,34	1,316
90	257294	295883	1111708	29231755	0,963	0,535	0,991	0,21	0,977	0,302	1,278

MEDIANA S3											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	2653450	811629	465334	13280467	0,966	0,234	0,833	0,636	0,895	0,342	1,237
20	1192775	570613	706350	14741142	0,954	0,324	0,925	0,447	0,939	0,375	1,315
30	608465	401353	875610	15325452	0,946	0,397	0,962	0,314	0,954	0,351	1,305
40	334842	279179	997784	15599075	0,94	0,455	0,979	0,219	0,959	0,295	1,254
50	193478	191827	1085136	15740439	0,936	0,498	0,988	0,15	0,961	0,231	1,192
60	116091	129516	1147447	15817826	0,932	0,527	0,993	0,101	0,962	0,17	1,132
70	70384	85775	1191188	15863533	0,93	0,549	0,996	0,067	0,962	0,12	1,081
80	43256	55049	1221914	15890661	0,929	0,56	0,997	0,043	0,962	0,08	1,042
90	26092	35501	1241462	15907825	0,928	0,576	0,998	0,028	0,962	0,053	1,015

MEDIANA S4											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	1714977	1163928	544377	17105358	0,969	0,404	0,909	0,681	0,938	0,507	1,446
20	1072213	874240	834065	17748122	0,955	0,449	0,943	0,512	0,949	0,478	1,427
30	743933	642019	1066286	18076402	0,944	0,463	0,96	0,376	0,952	0,415	1,367
40	541543	463279	1245026	18278792	0,936	0,461	0,971	0,271	0,953	0,342	1,295
50	400521	332212	1376093	18419814	0,93	0,453	0,979	0,194	0,954	0,272	1,226

Anexos

60	298352	235501	1472804	18521983	0,926	0,441	0,984	0,138	0,954	0,21	1,164
70	221201	162323	1545982	18599134	0,923	0,423	0,988	0,095	0,955	0,155	1,11
80	163907	106896	1601409	18656428	0,921	0,395	0,991	0,063	0,955	0,108	1,063
90	120798	69037	1639268	18699537	0,919	0,364	0,994	0,04	0,955	0,073	1,028

MEDIANA S5											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	1158858	1121975	922984	17324823	0,949	0,492	0,937	0,549	0,943	0,519	1,462
20	726824	911404	1133555	17756857	0,94	0,556	0,961	0,446	0,95	0,495	1,445
30	512566	755337	1289622	17971115	0,933	0,596	0,972	0,369	0,952	0,456	1,408
40	362858	624747	1420212	18120823	0,927	0,633	0,98	0,306	0,953	0,412	1,365
50	265166	513315	1531644	18218515	0,922	0,659	0,986	0,251	0,953	0,364	1,317
60	193482	416018	1628941	18290199	0,918	0,683	0,99	0,203	0,953	0,313	1,266
70	136351	330616	1714343	18347330	0,915	0,708	0,993	0,162	0,952	0,263	1,215
80	97231	256110	1788849	18386450	0,911	0,725	0,995	0,125	0,951	0,214	1,165
90	68890	195860	1849099	18414791	0,909	0,74	0,996	0,096	0,951	0,17	1,12

MEDIANA S6											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	58675	461879	96119	14831647	0,99356106	0,887	0,996	0,828	0,995	0,856	1,851
20	36540	430632	127366	14853782	0,99149825	0,922	0,998	0,772	0,995	0,84	1,835
30	25167	398903	159095	14865155	0,98941079	0,941	0,998	0,715	0,994	0,812	1,806
40	17267	366212	191786	14873055	0,98726930	0,955	0,999	0,656	0,993	0,778	1,771
50	11704	330729	227269	14878618	0,98495494	0,966	0,999	0,593	0,992	0,735	1,727
60	7936	290682	267316	14882386	0,98235503	0,973	0,999	0,521	0,991	0,679	1,67
70	5768	245286	312712	14884554	0,97942314	0,977	1	0,44	0,989	0,606	1,596
80	4175	195237	362761	14886147	0,97621069	0,979	1	0,35	0,988	0,516	1,503
90	2846	140072	417926	14887476	0,97269422	0,98	1	0,251	0,986	0,4	1,386

MEDIANA S7											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	35827	189480	122505	25986908	0,995	0,841	0,999	0,607	0,997	0,705	1,702
20	20547	137081	174904	26002188	0,993	0,87	0,999	0,439	0,996	0,584	1,58
30	12980	100545	211440	26009755	0,992	0,886	1	0,322	0,996	0,473	1,468
40	8469	73307	238678	26014266	0,991	0,896	1	0,235	0,995	0,372	1,368
50	6025	53434	258551	26016710	0,99	0,899	1	0,171	0,995	0,288	1,283
60	4745	40123	271862	26017990	0,99	0,894	1	0,129	0,995	0,225	1,22
70	3697	30907	281078	26019038	0,989	0,893	1	0,099	0,995	0,178	1,173
80	2741	23523	288462	26019994	0,989	0,896	1	0,075	0,994	0,139	1,134
90	1630	16960	295025	26021105	0,989	0,912	1	0,054	0,994	0,103	1,097

GENERACIÓN DE HISTOGRAMAS S1											
<i>U</i>	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	125997	328703	346053	23149327	0,985	0,723	0,995	0,487	0,99	0,582	1,572
8,00E+05	149334	345930	328826	23125990	0,986	0,698	0,994	0,513	0,99	0,591	1,581
8,50E+05	223891	362803	311953	23051433	0,987	0,618	0,99	0,538	0,989	0,575	1,564
9,00E+05	532204	379722	295034	22743120	0,987	0,416	0,977	0,563	0,982	0,479	1,461
9,50E+05	2606926	403821	270935	20668398	0,987	0,134	0,888	0,598	0,935	0,219	1,154

GENERACIÓN DE HISTOGRAMAS S2											
<i>U</i>	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	1328505	164349	213851	19755055	0,989	0,11	0,937	0,435	0,962	0,176	1,138
8,00E+05	1618766	176564	201636	19464794	0,99	0,098	0,923	0,467	0,955	0,162	1,118
8,50E+05	1996753	187311	190889	19086807	0,99	0,086	0,905	0,495	0,946	0,146	1,092
9,00E+05	2518685	197194	181006	18564875	0,99	0,073	0,881	0,521	0,932	0,127	1,06
9,50E+05	3483853	209545	168655	17599707	0,991	0,057	0,835	0,554	0,906	0,103	1,009

GENERACIÓN DE HISTOGRAMAS S3											
<i>U</i>	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	726888	667728	610535	15309409	0,962	0,479	0,955	0,522	0,958	0,5	1,458
8,00E+05	1017318	720801	557462	15018979	0,964	0,415	0,937	0,564	0,95	0,478	1,428
8,50E+05	1450425	765362	512901	14585872	0,966	0,345	0,91	0,599	0,937	0,438	1,375
9,00E+05	2131773	804754	473509	13904524	0,967	0,274	0,867	0,63	0,914	0,382	1,296
9,50E+05	3510584	854804	423459	12525713	0,967	0,196	0,781	0,669	0,864	0,303	1,167

GENERACIÓN DE HISTOGRAMAS S4											
<i>U</i>	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	396028	429309	1286781	18520202	0,935	0,52	0,979	0,25	0,957	0,338	1,294
8,00E+05	504718	499926	1216164	18411512	0,938	0,498	0,973	0,291	0,955	0,367	1,323
8,50E+05	742329	568082	1148008	18173901	0,941	0,434	0,961	0,331	0,951	0,375	1,326
9,00E+05	1544566	635193	1080897	17371664	0,941	0,291	0,918	0,37	0,93	0,326	1,256
9,50E+05	3279786	707186	1008904	15636444	0,939	0,177	0,827	0,412	0,879	0,248	1,127

GENERACIÓN DE HISTOGRAMAS S5											
<i>U</i>	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	284155	379950	1668990	18299225	0,916	0,572	0,985	0,185	0,949	0,28	1,229
8,00E+05	382747	426168	1622772	18200633	0,918	0,527	0,979	0,208	0,948	0,298	1,246
8,50E+05	552882	480018	1568922	18030498	0,92	0,465	0,97	0,234	0,944	0,312	1,256
9,00E+05	938533	549778	1499162	17644847	0,922	0,369	0,949	0,268	0,935	0,311	1,246
9,50E+05	2283282	651018	1397922	16300098	0,921	0,222	0,877	0,318	0,899	0,261	1,598

GENERACIÓN DE HISTOGRAMAS S6											
------------------------------	--	--	--	--	--	--	--	--	--	--	--

Anexos

U	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	5194	188166	370119	14988521	0,976	0,973	1	0,337	0,988	0,501	1,488
8,00E+05	6482	193796	364489	14987233	0,976	0,968	1	0,347	0,988	0,511	1,499
8,50E+05	10891	202631	355654	14982824	0,977	0,949	0,999	0,363	0,988	0,525	1,513
9,00E+05	33482	208719	349566	14960233	0,977	0,862	0,998	0,374	0,987	0,521	1,509
9,50E+05	211389	213093	345192	14782326	0,977	0,502	0,986	0,382	0,982	0,434	1,415

GENERACIÓN DE HISTOGRAMAS S7											
U	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
7,50E+05	26108	69324	242661	25996627	0,991	0,726	0,999	0,222	0,995	0,34	1,335
8,00E+05	28991	79020	232965	25993744	0,991	0,732	0,999	0,253	0,995	0,376	1,371
8,50E+05	38419	91552	220433	25984316	0,992	0,704	0,999	0,293	0,995	0,414	1,409
9,00E+05	65497	109627	202358	25957238	0,992	0,626	0,997	0,351	0,995	0,45	1,445
9,50E+05	364443	131224	180761	25658292	0,993	0,265	0,986	0,421	0,989	0,325	1,314

SG S1												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	0,000001	9956164	1282861	124730	19636565	0,994	0,114	0,664	0,911	0,796	0,203	0,999
3	0,000001	795389	1160942	246649	28797340	0,992	0,593	0,973	0,825	0,982	0,69	1,672
6	0,000001	69956	1022259	385332	29522773	0,987	0,936	0,998	0,726	0,992	0,818	1,81
9	0,000001	47090	866862	540729	29545639	0,982	0,948	0,998	0,616	0,99	0,747	1,737
12	0,000001	39784	748608	658983	29552945	0,978	0,95	0,999	0,532	0,988	0,682	1,67
15	0,000001	32821	622657	784934	29559908	0,974	0,95	0,999	0,442	0,986	0,604	1,59
18	0,000001	25528	478789	928802	29567201	0,97	0,949	0,999	0,34	0,984	0,501	1,485

SG S2												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
35	0,0100	78947	216286	1191305	29513782	0,961	0,733	0,997	0,154	0,979	0,254	1,2329
40	0,0100	61165	199842	1207749	29531564	0,961	0,766	0,998	0,142	0,979	0,24	1,2185
45	0,0100	48322	184440	1223151	29544407	0,96	0,792	0,998	0,131	0,979	0,225	1,2038
50	0,0100	38862	169960	1237631	29553867	0,96	0,814	0,999	0,121	0,979	0,21	1,1892
55	0,0100	31598	155945	1251646	29561131	0,959	0,832	0,999	0,111	0,979	0,196	1,1743
60	0,0100	25745	142504	1265087	29566984	0,959	0,847	0,999	0,101	0,979	0,181	1,1595
65	0,0100	21207	129473	1278118	29571522	0,959	0,859	0,999	0,092	0,979	0,166	1,1447

SG S3												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	0,001	5271611	1187630	637850	13535229	0,955	0,184	0,72	0,651	0,821	0,287	1,108
5	0,001	2053130	949069	876411	16753710	0,95	0,316	0,891	0,52	0,92	0,393	1,313
10	0,001	882842	627152	1198328	17923998	0,937	0,415	0,953	0,344	0,945	0,376	1,321
15	0,001	408410	381476	1444004	18398430	0,927	0,483	0,978	0,209	0,952	0,292	1,244

Anexos

20	0.001	195357	218837	1606643	18611483	0,921	0,528	0,99	0,12	0,954	0,195	1,149
25	0.001	97419	119881	1705599	18709421	0,916	0,552	0,995	0,066	0,954	0,117	1,071
30	0.001	51493	60990	1764490	18755347	0,914	0,542	0,997	0,033	0,954	0,063	1,017

SG S4												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	0.05	55387	126308	1589782	18860843	0,922	0,695	0,997	0,074	0,958	0,133	1,091
25	0.05	50656	120896	1595194	18865574	0,922	0,705	0,997	0,07	0,958	0,128	1,086
30	0.05	46798	116176	1599914	18869432	0,922	0,713	0,998	0,068	0,958	0,124	1,082
35	0.05	43539	112308	1603782	18872691	0,922	0,721	0,998	0,065	0,958	0,12	1,078
40	0.05	40681	108723	1607367	18875549	0,922	0,728	0,998	0,063	0,958	0,117	1,075
45	0.05	38227	105536	1610554	18878003	0,921	0,734	0,998	0,061	0,958	0,113	1,072
50	0.05	36094	102561	1613529	18880136	0,921	0,74	0,998	0,06	0,958	0,111	1,069

SG S5												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	0,0005	2068565	1356699	692241	16514815	0,96	0,396	0,889	0,662	0,923	0,496	1,419
5	0,0005	776489	1148588	900352	17806891	0,952	0,597	0,958	0,561	0,955	0,578	1,533
10	0,0005	523164	971256	1077684	18060216	0,944	0,65	0,972	0,474	0,958	0,548	1,506
15	0,0005	367830	782427	1266513	18215550	0,935	0,68	0,98	0,382	0,957	0,489	1,446
20	0,0005	246011	598910	1450030	18337369	0,927	0,709	0,987	0,292	0,956	0,414	1,37
25	0,0005	162204	432687	1616253	18421176	0,919	0,727	0,991	0,211	0,954	0,327	1,281
30	0,0005	104186	295702	1753238	18479194	0,913	0,739	0,994	0,144	0,952	0,242	1,194

SG S6												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	0,000001	918888	817632	85672	18810128	0,995	0,471	0,953	0,905	0,974	0,619	1,593
3	0,000001	66939	807251	96053	19662077	0,995	0,923	0,997	0,894	0,996	0,908	1,904
6	0,000001	18948	793865	109439	19710068	0,994	0,977	0,999	0,879	0,997	0,925	1,922
9	0,000001	11665	774346	128958	19717351	0,994	0,985	0,999	0,857	0,996	0,917	1,913
12	0,000001	8878	758687	144617	19720138	0,993	0,988	1	0,84	0,996	0,908	1,904
15	0,000001	7323	746112	157192	19721693	0,992	0,99	1	0,826	0,996	0,901	1,897
18	0,000001	6555	734643	168661	19722461	0,992	0,991	1	0,813	0,996	0,893	1,889

SG S7												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	0,00001	1603414	291754	54814	29050338	0,998	0,154	0,948	0,842	0,972	0,26	1,233
5	0,00001	106669	218705	127863	30547083	0,996	0,672	0,997	0,631	0,996	0,651	1,647
10	0,00001	75574	151586	194982	30578178	0,994	0,667	0,998	0,437	0,996	0,528	1,524
15	0,00001	49488	104717	241851	30604264	0,992	0,679	0,998	0,302	0,995	0,418	1,413
20	0,00001	27002	68362	278206	30626750	0,991	0,717	0,999	0,197	0,995	0,309	1,304

Anexos

25	0,00001	10124	41242	305326	30643628	0,99	0,803	1	0,119	0,995	0,207	1,202
30	0,00001	4942	24971	321597	30648810	0,99	0,835	1	0,072	0,995	0,133	1,127

MoG S1												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
24	0.00000500	43659	205664	67829	20315168	0,997	0,825	0,998	0,752	0,997	0,787	1,784
26	0.00000500	38977	204623	68870	20319850	0,997	0,840	0,998	0,748	0,997	0,791	1,789
28	0.00000500	37025	203586	69907	20321802	0,997	0,846	0,998	0,744	0,997	0,792	1,789
30	0.00000500	36405	202613	70880	20322422	0,997	0,848	0,998	0,741	0,997	0,791	1,788
32	0.00000500	35974	201729	71764	20322853	0,996	0,849	0,998	0,738	0,997	0,789	1,787
34	0.00000500	35646	200860	72633	20323181	0,996	0,849	0,998	0,734	0,997	0,788	1,785
38	0.00000500	35068	199188	74305	20323759	0,996	0,850	0,998	0,728	0,997	0,785	1,782

MoG S2												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	0.01000000	1634778	396420	1011171	27957951	0,965	0,195	0,945	0,282	0,955	0,231	1,185
25	0.01000000	1317330	370949	1036642	28275399	0,965	0,220	0,955	0,264	0,960	0,240	1,200
30	0.01000000	1094958	352292	1055299	28497771	0,964	0,243	0,963	0,250	0,964	0,247	1,210
35	0.01000000	924889	339316	1068275	28667840	0,964	0,268	0,969	0,241	0,966	0,254	1,220
40	0.01000000	793991	328968	1078623	28798738	0,964	0,293	0,973	0,234	0,969	0,260	1,229
45	0.01000000	39963	93395	1314196	29552766	0,957	0,700	0,999	0,066	0,978	0,121	1,099
50	0.01000000	35936	89649	1317942	29556793	0,957	0,714	0,999	0,064	0,978	0,117	1,095

MoG S3												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
10	0.00500000	1879477	817452	460811	14156820	0,968	0,303	0,883	0,640	0,924	0,411	1,335
15	0.00500000	1347983	742344	535919	14688314	0,965	0,355	0,916	0,581	0,940	0,441	1,380
20	0.00500000	1041005	682724	595539	14995292	0,962	0,396	0,935	0,534	0,948	0,455	1,403
25	0.00500000	837501	632405	645858	15198796	0,959	0,430	0,948	0,495	0,953	0,460	1,414
30	0.00500000	693671	592732	685531	15342626	0,957	0,461	0,957	0,464	0,957	0,462	1,419
35	0.00500000	589112	556020	722243	15447185	0,955	0,486	0,963	0,435	0,959	0,459	1,418
40	0.00500000	510167	525475	752788	15526130	0,954	0,507	0,968	0,411	0,961	0,454	1,415

MoG S4												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
5	0.00500000	1982167	1350580	365510	16934063	0,979	0,405	0,895	0,787	0,935	0,535	1,470
10	0.00500000	1347564	1237708	478382	17568666	0,973	0,479	0,929	0,721	0,951	0,575	1,526
15	0.00500000	1105182	1155022	561068	17811048	0,969	0,511	0,942	0,673	0,955	0,581	1,536
20	0.00500000	962155	1086906	629184	17954075	0,966	0,530	0,949	0,633	0,958	0,577	1,535
25	0.00500000	860362	1025961	690129	18055868	0,963	0,544	0,955	0,598	0,959	0,570	1,528
30	0.00500000	779507	972787	743303	18136723	0,961	0,555	0,959	0,567	0,960	0,561	1,521

Anexos

35	0.00500000	708314	924127	791963	18207916	0,958	0,566	0,963	0,539	0,960	0,552	1,512
----	------------	--------	--------	--------	----------	-------	-------	-------	-------	-------	-------	-------

MoG S5												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
5	0.00100000	2004925	1508428	540512	16578455	0,968	0,429	0,892	0,736	0,929	0,542	1,471
10	0.00100000	1484570	1418900	630040	17098810	0,964	0,489	0,920	0,693	0,942	0,573	1,515
15	0.00100000	1281704	1360400	688540	17301676	0,962	0,515	0,931	0,664	0,946	0,580	1,526
20	0.00100000	1167805	1320160	728780	17415575	0,960	0,531	0,937	0,644	0,948	0,582	1,530
25	0.00100000	1087692	1284798	764142	17495688	0,958	0,542	0,941	0,627	0,950	0,581	1,531
30	0.00100000	1026169	1252721	796219	17557211	0,957	0,550	0,945	0,611	0,951	0,579	1,530
35	0.00100000	976216	1226738	822202	17607164	0,955	0,557	0,947	0,599	0,951	0,577	1,528

MoG S6												
σ	α	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
6	0.00000100	110150	493376	64909	14883565	0,996	0,817	0,993	0,884	0,994	0,849	1,843
6	0.00000500	110168	493377	64908	14883547	0,996	0,817	0,993	0,884	0,994	0,849	1,843
6	0.00001000	110080	493377	64908	14883635	0,996	0,818	0,993	0,884	0,994	0,849	1,844
6	0.00005000	107331	493372	64913	14886384	0,996	0,821	0,993	0,884	0,994	0,851	1,846
6	0.00010000	103685	493357	64928	14890030	0,996	0,826	0,993	0,884	0,994	0,854	1,848
6	0.00050000	80866	481106	77179	14912849	0,995	0,856	0,995	0,862	0,995	0,859	1,854
6	0.00100000	43071	377918	180367	14950644	0,988	0,898	0,997	0,677	0,993	0,772	1,764

GAMMA S1												
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma	
1	51295	627841	36301	23027283	0,998	0,924	0,998	0,945	0,998	0,935	1,933	
2	18842	610531	53611	23059736	0,998	0,97	0,999	0,919	0,998	0,944	1,942	
3	15128	596612	67530	23063450	0,997	0,975	0,999	0,898	0,998	0,935	1,933	

GAMMA S2												
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma	
1	1106473	811929	584719	28289839	0,98	0,423	0,962	0,581	0,971	0,49	1,461	
2	631611	682116	714532	28764701	0,976	0,519	0,979	0,488	0,977	0,503	1,48	
3	381691	580185	816463	29014621	0,973	0,603	0,987	0,415	0,98	0,492	1,472	

GAMMA S3												
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma	
1	3159448	951692	307734	12688326	0,976	0,231	0,801	0,756	0,88	0,354	1,234	
2	2358117	862027	397399	13489657	0,971	0,268	0,851	0,684	0,907	0,385	1,292	
3	1887629	794325	465101	13960145	0,968	0,296	0,881	0,631	0,922	0,403	1,325	

GAMMA S4												
----------	--	--	--	--	--	--	--	--	--	--	--	--

Anexos

τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	303226	319328	1381873	18420533	0,93	0,513	0,984	0,188	0,956	0,275	1,231
2	198961	178341	1522860	18524798	0,924	0,473	0,989	0,105	0,956	0,172	1,127
3	134922	111307	1589894	18588837	0,921	0,452	0,993	0,065	0,956	0,114	1,07

GAMMA S5											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	2050939	1625706	409564	16338751	0,976	0,442	0,888	0,799	0,93	0,569	1,499
2	1802813	1578526	456744	16586877	0,973	0,467	0,902	0,776	0,936	0,583	1,519
3	1674623	1544846	490424	16715067	0,971	0,48	0,909	0,759	0,939	0,588	1,527

GAMMA S6											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	39905	506147	45930	14752658	0,997	0,927	0,997	0,917	0,997	0,922	1,919
2	31138	504879	47198	14761425	0,997	0,942	0,998	0,915	0,997	0,928	1,925
3	25292	503957	48120	14767271	0,997	0,952	0,998	0,913	0,998	0,932	1,93

GAMMA S7											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
1	150192	301701	9694	25769453	1	0,668	0,994	0,969	0,997	0,791	1,787
2	140782	294318	17077	25778863	0,999	0,676	0,995	0,945	0,997	0,789	1,785
3	134613	287787	23608	25785032	0,999	0,681	0,995	0,924	0,997	0,784	1,781

KDE S1											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
5,00E+03	1683732	376259	1031332	27908997	0,964	0,183	0,943	0,267	0,954	0,217	1,171
7,00E+03	1921757	441848	965743	27670972	0,966	0,187	0,935	0,314	0,950	0,234	1,185
9,00E+03	2087222	489994	917597	27505507	0,968	0,190	0,929	0,348	0,948	0,246	1,194
1,10E+04	2506853	532132	875459	27085876	0,969	0,175	0,915	0,378	0,941	0,239	1,181
1,30E+04	2997653	567745	839846	26595076	0,969	0,159	0,899	0,403	0,933	0,228	1,161
1,50E+04	3289641	595736	811855	26303088	0,970	0,153	0,889	0,423	0,928	0,225	1,153
L	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
20	2517529	521076	886515	27075200	0,968	0,171	0,915	0,370	0,941	0,234	1,175
40	2917336	571467	836124	26675393	0,970	0,164	0,901	0,406	0,934	0,233	1,168
60	3209093	608762	798829	26383636	0,971	0,159	0,892	0,432	0,929	0,233	1,162
80	3471790	637675	769916	26120939	0,971	0,155	0,883	0,453	0,925	0,231	1,156

KDE S2											
τ	fp	tp	fn	tn	Prec0	Prec1	Rec0	Rec1	Fs0	Fs1	Suma
5,00E+03	3309359	429759	977832	26283370	0,964	0,115	0,888	0,305	0,925	0,167	1,092
7,00E+03	4191831	491949	915642	25400898	0,965	0,105	0,858	0,349	0,909	0,162	1,070

9,00E+03	4882221	537081	870510	24710508	0,966	0,099	0,835	0,382	0,896	0,157	1,053
1,10E+04	5578791	576343	831248	24013938	0,967	0,094	0,811	0,409	0,882	0,152	1,035
1,30E+04	6161698	609349	798242	23431031	0,967	0,090	0,792	0,433	0,871	0,149	1,020
1,50E+04	6606883	635387	772204	22985846	0,967	0,088	0,777	0,451	0,862	0,147	1,009
<i>L</i>	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
20	3826667	470516	937075	25766062	0,965	0,109	0,871	0,334	0,915	0,165	1,080
40	4135881	494324	913267	25456848	0,965	0,107	0,860	0,351	0,910	0,164	1,074
60	4552409	514582	893009	25040320	0,966	0,102	0,846	0,366	0,902	0,159	1,061
80	4956103	528669	878922	24636626	0,966	0,096	0,833	0,376	0,894	0,153	1,048

KDE3											
τ	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
5,00E+03	1887994	385380	892883	14148303	0,941	0,170	0,882	0,301	0,911	0,217	1,128
7,00E+03	2473205	458657	819606	13563092	0,943	0,156	0,846	0,359	0,892	0,218	1,110
9,00E+03	2964385	509614	768649	13071912	0,944	0,147	0,815	0,399	0,875	0,214	1,090
1,10E+04	3489877	549866	728397	12546420	0,945	0,136	0,782	0,430	0,856	0,207	1,063
1,30E+04	3954722	580535	697728	12081575	0,945	0,128	0,753	0,454	0,839	0,200	1,038
1,50E+04	4305322	603945	674318	11730975	0,946	0,123	0,732	0,472	0,825	0,195	1,020
<i>L</i>	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
20	2437486	376276	901987	13598811	0,938	0,134	0,848	0,294	0,891	0,184	1,075
40	3060059	390015	888248	12976238	0,936	0,113	0,809	0,305	0,868	0,165	1,033
60	3593353	407895	870368	12442944	0,935	0,102	0,776	0,319	0,848	0,155	1,002
80	4063664	417569	860694	11972633	0,933	0,093	0,747	0,327	0,829	0,145	0,974

KDE S4											
τ	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
5,00E+03	1116755	778950	937140	17799475	0,950	0,411	0,941	0,454	0,945	0,431	1,377
7,00E+03	1342599	927408	788682	17573631	0,957	0,409	0,929	0,540	0,943	0,465	1,408
9,00E+03	1519789	1027131	688959	17396441	0,962	0,403	0,920	0,599	0,940	0,482	1,422
1,10E+04	1779561	1102463	613627	17136669	0,965	0,383	0,906	0,642	0,935	0,480	1,414
1,30E+04	2014152	1159984	556106	16902078	0,968	0,365	0,894	0,676	0,929	0,474	1,404
1,50E+04	2173641	1203994	512096	16742589	0,970	0,356	0,885	0,702	0,926	0,473	1,398
<i>L</i>	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
20	2188316	1111432	604658	16727914	0,965	0,337	0,884	0,648	0,923	0,443	1,366
40	2717693	1201544	514546	16198537	0,969	0,307	0,856	0,700	0,909	0,426	1,336
60	3188465	1278082	438008	15727765	0,973	0,286	0,831	0,745	0,897	0,413	1,310
80	3657503	1319534	396556	15258727	0,975	0,265	0,807	0,769	0,883	0,394	1,277

KDE S5											
τ	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
5,00E+03	1301720	539558	1509382	17281660	0,920	0,293	0,930	0,263	0,925	0,277	1,202

7,00E+03	1522963	631274	1417666	17060417	0,923	0,293	0,918	0,308	0,921	0,300	1,221
9,00E+03	1672810	693866	1355074	16910570	0,926	0,293	0,910	0,339	0,918	0,314	1,232
1,10E+04	1925858	746881	1302059	16657522	0,928	0,279	0,896	0,365	0,912	0,316	1,228
1,30E+04	2157619	790887	1258053	16425761	0,929	0,268	0,884	0,386	0,906	0,317	1,222
1,50E+04	2286667	824947	1223993	16296713	0,930	0,265	0,877	0,403	0,903	0,320	1,222
<i>L</i>	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
20	1923434	790004	1258936	16659946	0,930	0,291	0,896	0,386	0,913	0,332	1,245
40	2105982	859406	1189534	16477398	0,933	0,290	0,887	0,419	0,909	0,343	1,252
60	2251312	922291	1126649	16332068	0,935	0,291	0,879	0,450	0,906	0,353	1,259
80	2426118	968843	1080097	16157262	0,937	0,285	0,869	0,473	0,902	0,356	1,258

KDE S6											
τ	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
5,00E+03	289502	112663	445622	14704213	0,971	0,280	0,981	0,202	0,976	0,235	1,210
7,00E+03	303723	121858	436427	14689992	0,971	0,286	0,980	0,218	0,975	0,248	1,223
9,00E+03	310512	127776	430509	14683203	0,972	0,292	0,979	0,229	0,975	0,256	1,232
1,10E+04	389305	132356	425929	14604410	0,972	0,254	0,974	0,237	0,973	0,245	1,218
1,30E+04	438373	135949	422336	14555342	0,972	0,237	0,971	0,244	0,971	0,240	1,211
1,50E+04	461621	138572	419713	14532094	0,972	0,231	0,969	0,248	0,971	0,239	1,210
<i>L</i>	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
20	403923	126207	432078	14589792	0,971	0,238	0,973	0,226	0,972	0,232	1,204
40	422819	131862	426423	14570896	0,972	0,238	0,972	0,236	0,972	0,237	1,209
60	444041	140029	418256	14549674	0,972	0,240	0,970	0,251	0,971	0,245	1,216
80	467730	144912	413373	14525985	0,972	0,237	0,969	0,260	0,971	0,248	1,218

KDE S7											
τ	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
1,00E+04	581884	158883	187685	30071868	0,994	0,214	0,981	0,458	0,987	0,292	1,280
1,50E+04	949030	180726	165842	29704722	0,994	0,160	0,969	0,521	0,982	0,245	1,226
2,00E+04	1040898	194207	152361	29612854	0,995	0,157	0,966	0,560	0,980	0,246	1,226
2,50E+04	1377180	205077	141491	29276572	0,995	0,130	0,955	0,592	0,975	0,213	1,187
3,00E+04	1458790	212592	133976	29194962	0,995	0,127	0,952	0,613	0,973	0,211	1,184
<i>L</i>	<i>fp</i>	<i>tp</i>	<i>fn</i>	<i>tn</i>	<i>Prec0</i>	<i>Prec1</i>	<i>Rec0</i>	<i>Rec1</i>	<i>Fs0</i>	<i>Fs1</i>	<i>Suma</i>
20	676821	180599	165969	29976931	0,994	0,211	0,978	0,521	0,986	0,300	1,286
40	623688	200393	146175	30030064	0,995	0,243	0,980	0,578	0,987	0,342	1,330
60	595629	212492	134076	30058123	0,996	0,263	0,981	0,613	0,988	0,368	1,356
80	589895	221779	124789	30063857	0,996	0,273	0,981	0,640	0,988	0,383	1,371

PRESUPUESTO (ejemplo)

1) Ejecución Material

- Compra de ordenador personal (Software incluido)..... 2.000 €
- Material de oficina 150 €
- Total de ejecución material 2.150 €

2) Gastos generales

- 16 % sobre Ejecución Material 352 €

3) Beneficio Industrial

- 6 % sobre Ejecución Material 132 €

4) Honorarios Proyecto

- 640 horas a 15 € / hora 9600 €

5) Material fungible

- Gastos de impresión 60 €
- Encuadernación 200 €

6) Subtotal del presupuesto

- Subtotal Presupuesto 12060 €

7) I.V.A. aplicable

- 16% Subtotal Presupuesto 1929.6 €

8) Total presupuesto

- Total Presupuesto 13939,6 €

Madrid, Mes de 2009

El Ingeniero Jefe de Proyecto

Fdo.: Sonsoles Herrero Martín
Ingeniero Superior de Telecomunicación

PLIEGO DE CONDICIONES

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de análisis comparativo de técnicas de segmentación de secuencias de video basadas en modelado de fondo desarrollado en este PFC. 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.