

UNIVERSIDAD AUTÓNOMA DE MADRID
ESCUELA POLITÉCNICA SUPERIOR



SEGMENTACIÓN
PERSONA-FONDO USANDO
INFORMACIÓN DE
SEGMENTACIÓN FRENTE-FONDO

Raúl Lara Arranz

PROYECTO FIN DE CARRERA
INGENIERÍA DE TELECOMUNICACIÓN

Departamento de Tecnología Electrónica y de las Comunicaciones
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Diciembre 2014

SEGMENTACIÓN PERSONA-FONDO USANDO INFORMACIÓN DE SEGMENTACIÓN FRENTE-FONDO

Raúl Lara Arranz

Tutor: Álvaro García Martín

Ponente: José María Martínez Sánchez



**Video Processing and Understanding Lab
Departamento de Ingeniería Informática
Escuela Politécnica Superior
Universidad Autónoma de Madrid
Diciembre 2014**

**Trabajo parcialmente financiado por el gobierno español bajo el proyecto
TEC2011-25995 (EventVideo)**



Resumen

En la actualidad el procesamiento y análisis de vídeo está en auge, esto es debido principalmente a la masiva instalación de cámaras de video para múltiples cometidos. Por este motivo, la algoritmia relativa al procesamiento y análisis de vídeo ha adquirido suma importancia. Actualmente los algoritmos de segmentación Persona-Fondo no están muy extendidos, por el contrario los segmentadores Frente-Fondo han sido desarrollados mucho más que los primeros.

Por este motivo, el objetivo principal de este proyecto es mejorar la segmentación Persona-Fondo usando información de segmentación Frente-Fondo puesto que ambas segmentaciones tienen objetivos distintos. Por eso tras un amplio análisis del Estado del Arte, en el caso de la segmentación Persona-Fondo el algoritmo disponible en el laboratorio VPULab, se estudió el funcionamiento de los segmentadores elegidos en ambos casos. Tras el estudio, se desarrollaron diversas opciones de unir las informaciones de ambas segmentaciones para mejorar en primer lugar la segmentación Persona-Fondo. Una vez se obtuvo un método óptimo, se aplicó el mismo a la segmentación Frente-Fondo para obtener también una mejora de la misma.

Una vez se obtuvieron los métodos óptimos, se procedió a la evaluación de las segmentaciones por independiente, posteriormente se evaluó un primer método sin obtener resultados óptimos. Tras esto, se evaluaron los dos métodos óptimos desarrollados en el proyecto usando un dataset conformado por vídeos del Estado del Arte, obteniendo resultados satisfactorios de ambos métodos.

Abstract

Nowadays the video processing and analysis is in full developing, this is mainly due to the massive installation of video cameras for multiple tasks. For this reason, the algorithms of the video processing and analysis have gained importance. Currently People-Background segmentation algorithms are not spread out, however the Background-Foreground segmentation has been developed more than the firsts one.

Therefore, the main objective of this project is to enhance the People-Background segmentation using Background-Foreground segmentation information as both segmentations have different objectives. So after an extensive analysis of the State of Art, in the case of People-Background segmentation the algorithm available in the VPULab, the functioning of the segmentation was studied in both cases. After that, various options of attaching the information to improve the People-Background segmentation was developed. Once an optimal method was obtained, it was applied to the Foreground-Background segmentation to obtain also an improvement.

Once the optimal methods were obtained, an evaluation was made for both kinds of segmentation independently. Later one of the method was evaluated obtaining not enough improvement. After this, both optimal methods were evaluated with a dataset made by videos of the State of Art, obtaining good results.

Agradecimientos

En primer lugar, quería dar las gracias a mi tutor Álvaro García, por darme la oportunidad de realizar este Proyecto Fin de Carrera en el VPULab por su apoyo y dedicación durante este tiempo. No quiero olvidarme de los jefes del grupo y de los compañeros del grupo Pencho, Marcos, Rafa, Luis y Diego, los cuales me han hecho amenas y entretenidas las horas de laboratorio.

Pero sin duda alguna, tengo que darles las gracias a las personas más importantes que me han llevado hasta aquí. A mi MADRE y mi PADRE, gracias por darme la oportunidad de estar donde estoy ahora mismo, gracias por ser mi apoyo, mis fuerzas y por hacerme como soy. Mucha parte de esta carrera es vuestra porque vuestro esfuerzo la ha hecho posible. A mi HERMANO, que sin el esto no hubiera sido posible dado que es único, el mejor. Gracias a los tres por no haberme fallado nunca y estar siempre hay, soy quien soy por vosotros.

Darle las gracias a mi compañera de prácticas y pareja Pili, por todo lo que hemos compartido tanto en la universidad como fuera de ella. Muchas gracias a mis compañeros de clase, a los compañeros de Bolonia que pese a no compartir clase, se han compartido horas de comer y laboratorios, a los integrantes del grupo RFCAS. No me quiero olvidar de mis amigos de toda la vida de Sanse los cuales siempre han estado ahí. Muchas Gracias a todos aunque no diga vuestro nombre cada uno sabe quién es.

Raúl Lara Arranz

Diciembre de 2014

Índice general

Resumen	v
Abstract	VII
Agradecimientos	IX
1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.3. Medios	3
1.4. Estructura de la memoria	3
2. Estado del Arte	5
2.1. Introducción	5
2.2. Segmentación Persona-Fondo	5
2.3. Segmentación Frente-Fondo	8
2.3.1. Introducción	9
2.3.2. Problemática en Background Subtraction	11
2.3.3. Clasificación del Estado del Arte de Background Subtraction	12
2.3.4. Conclusiones	21
2.4. Conclusiones	22
3. Algoritmo	23
3.1. Introducción	23
3.2. Algoritmo Persona-Fondo	24
3.3. Algoritmo Frente-Fondo	28
3.3.1. Introducción	28
3.3.2. Librería BGSLibrary	28
3.3.3. Algoritmo SGMM-SOD	30
3.4. Algoritmo Propuesto.	31
3.4.1. Primera Aproximación	31
3.4.2. Segunda Aproximación	34
3.4.3. Tercera Aproximación	37
3.4.4. Conclusiones	38

4. Evaluación	41
4.1. Introducción	41
4.2. Dataset	42
4.3. Métrica	44
4.4. Resultados	46
4.4.1. Segmentación Persona-Fondo	46
4.4.2. Segmentación Frente-Fondo	49
4.4.3. Primera Aproximación	52
4.4.4. Segunda y Tercera Aproximación	56
4.4.5. Segunda Aproximación	59
4.4.6. Tercera Aproximación	63
4.4.7. Conclusiones	68
5. Conclusiones y Trabajo Futuro	71
5.1. Conclusiones	71
5.2. Trabajo Futuro	72
Bibliografía	74
A. Acrónimos	83
B. Descripción algoritmos Fondo-Frente	85
C. Ejemplo de Configuración Segmentador Frente-Fondo	95
D. Resultados Adicionales a la Evaluación	99
D.1. Segmentadores Frente-Fondo	100
D.2. Resultados Segunda Aproximación	107
D.3. Resultados Tercera Aproximación	114
E. Presupuesto	121
F. Pliego De Condiciones	123

Índice de figuras

2.1. Modelo Holístico o persona entera y Modelo Por Partes al doble de resolución	6
2.2. Modelo Independiente y Modelo Dependiente	6
2.3. Imagen Original, Mapa de Confianza en escala de grises o color y Máscara Final	8
2.4. Esquema resultado de una segmentación en las dos clases Fondo y Frente (Imagen obtenida dataset VPU UAM)	9
2.5. Esquema BGS básico de una secuencia de vídeo	11
2.6. Esquema Clasificación Métodos de Segmentación Frente-Fondo	13
2.7. Máscara Salida Segmentador Frente-Fondo para dos Segmentadores Distintos	21
3.1. Ejemplo visual parámetros algoritmo segmentación Persona-Fondo.	26
3.2. Desarrollo del algoritmo Persona-Fondo utilizado para este proyecto	27
3.3. Ejemplo ejecución de la Primera Aproximación	32
3.4. Mejoras visuales de los segmentadores Frente-Fondo de la primera aproximación	33
3.5. Máscaras de Salida de la ejecución de la Primera Aproximación	33
3.6. Gráfica Precisión/Recall primera aproximación	35
3.7. Ejemplo filtrado por reconstrucción explicativo del proceso a seguir.	36
3.8. F1 Score de frente y de fondo obteniendo de su punto de intersección el umbral óptimo.	37
3.9. Reconstrucción máscara Persona-Fondo usando la información de la segmentación Frente-Fondo.	37
3.10. Reconstrucción máscara Frente usando la información de la segmentación Persona-Fondo.	38
3.11. Esquema General del Algoritmo	39
4.1. Puntuaciones utilizadas en la evaluación con referencia al <i>ground truth</i>	45

4.2. Resultados Segmentador Persona-Fondo para diferentes umbrales	48
4.3. Gráfica PF con distintos factores de penalización	48
4.4. Gráfica RFR con distintos factores de penalización	49
4.5. Resultados Segmentadores (1) (Vídeo Pedestrians frame 345).	51
4.6. Resultados Segmentadores (2) (Vídeo Cubicle frame 2500).	52
4.7. Gráficas Precisión/Recall AVSS_S1 para los métodos SimpleGaussian	54
4.8. Gráficas Precisión/Recall AVSS_S1 para los métodos FuzzyGaussian .	54
4.9. Gráfica PF con distintos factores de penalización primera aproximación	55
4.10. Gráfica RFR con distintos factores de penalización primera aproximación	55
4.11. Imágenes (1) Algoritmo Persona-Fondo Desarrollado	61
4.12. Imágenes (2) Algoritmo Persona-Fondo Desarrollado	61
4.13. Gráfica puntuación PF de la segunda aproximación para distintos factores de penalización	62
4.14. Gráfica puntuación RFR de la segunda aproximación para distintos factores de penalización	63
4.15. Imágenes (1) Algoritmo Frente-Fondo Desarrollado	66
4.16. Imágenes (2) Algoritmo Frente-Fondo Desarrollado	66
4.17. Gráfica puntuación PF de la tercera aproximación para distintos factores de penalización	67
4.18. Gráfica puntuación RFR de la tercera aproximación para distintos factores de penalización	68

Índice de tablas

4.1. Resumen de secuencias de vídeo utilizadas en la evaluación así como el número de frames totales y evaluados	42
4.2. Vídeos Dataset del Proyectos (1) mostrando el frame y correspondiente <i>ground truth</i>	43
4.3. Vídeos Dataset del Proyectos (2) mostrando el frame y correspondiente <i>ground truth</i>	44
4.4. Puntuaciones Promedio Mapas de Confianza Segmentador Persona-Fondo	47
4.5. Puntuaciones Promedio Segmentadores Frente-Fondo	50
4.6. Puntuación Mapas de Confianza Vídeo AVSS_S1 segmentados	53
4.7. Puntuación Limite AND	57
4.8. Puntuación Limite OR	58
4.9. Puntuación Reconstrucción Mejora Segmentador Persona-Fondo	60
4.10. Puntuación Reconstrucción Mejora Segmentador Frente-Fondo	65
C.1. Parámetros originales Segmentador SimpleGaussian	96
C.2. Puntuaciones Segmentador Original	96
C.3. Puntuaciones Parámetro LearningRate modificado	96
C.4. Puntuaciones Parámetro NoiseVariance modificado	97
C.5. Puntuaciones Parámetro Sensitivity modificado	97
C.6. Puntuaciones Parámetro configuración final	97
D.1. Precisión Fondo Segmentadores para cada Vídeo	101
D.2. Precisión Frente Segmentadores para cada Vídeo	102
D.3. Recall Fondo Segmentadores para cada Vídeo	103
D.4. Recall Frente Segmentadores para cada Vídeo	104
D.5. F1Score Fondo Segmentadores para cada Vídeo	105
D.6. F1Score Frente Segmentadores para cada Vídeo	106
D.7. Precisión Fondo Segunda Aproximación para cada Vídeo	108

D.8. Precisión Frente Segunda Aproximación para cada Vídeo	109
D.9. Recall Fondo Segunda Aproximación para cada Vídeo	110
D.10. Recall Frente Segunda Aproximación para cada Vídeo	111
D.11. F1Score Fondo Segunda Aproximación para cada Vídeo	112
D.12. F1Score Frente Segunda Aproximación para cada Vídeo	113
D.13. Precisión Fondo Tercera Aproximación para cada Vídeo	115
D.14. Precisión Frente Tercera Aproximación para cada Vídeo	116
D.15. Recall Fondo Tercera Aproximación para cada Vídeo	117
D.16. Recall Frente Tercera Aproximación para cada Vídeo	118
D.17. F1Score Fondo Tercera Aproximación para cada Vídeo	119
D.18. F1Score Frente Tercera Aproximación para cada Vídeo	120

Capítulo 1

Introducción

1.1. Motivación

En la actualidad el uso de sistemas de visión artificial ha adquirido una gran importancia en múltiples ámbitos. Este crecimiento se debe principalmente a dos factores, el avance que ha experimentado el procesamiento digital de imágenes y vídeo en los últimos años, y el abaratamiento de precios de las herramientas de captura de imágenes y vídeo.

La gran implantación de las cámaras de vídeo en la sociedad en la que vivimos hace que sea inviable tener personal suficiente para controlar y gestionar las grabaciones realizadas. Por este motivo la algoritmia referente a la detección automática de objetos, seguimiento, reconocimiento de acciones y demás tecnologías aplicadas al análisis y comprensión de la imagen digital han adquirido en nuestros días un gran peso e importancia.

La segmentación Frente-Fondo [1], consiste principalmente en la identificación de aquellas zonas de la escena que se desean aislar del fondo de la escena. Las técnicas más comunes utilizan la evolución en el tiempo de alguna característica de bajo nivel de la escena (nivel de gris, flujo óptico, etc) con el fin de modelar que es el fondo y todo aquello que no cumpla el modelo es clasificado como frente. Por otro lado, la segmentación Persona-Fondo [2] consiste en identificar aquellas zonas de la imagen que con mayor probabilidad no son personas y por lo tanto pertenecen al fondo de la escena.

Ambos tipos de segmentación están diseñados para un objetivo distinto. La segmentación Frente-Fondo permite la obtención de los objetos de fondo de manera precisa, por contra tiene diversos problemas difíciles de resolver como pueden ser sombras, movimiento de la cámara (*jitter*), ruido de la cámara, etc. Por el otro lado,

la segmentación Persona-Fondo busca clasificar por completo la persona pese a cometer fallos clasificando fondo como frente, su principal ventaja es la eliminación de otros objetos de frente dado que solo busca personas. Su principal desventaja es la sobreestimación o detección de posibles personas para lograr su objetivo de asegurar que zonas de la escena no presentan personas.

El principal objetivo del Proyecto Final de Carrera es mejorar la segmentación Persona-Fondo utilizando información de segmentación Frente-Fondo. En este proyecto nos hemos centrado en la unión de las cualidades de ambos tipos de segmentación, para suplir las deficiencias de cada una de forma independiente puesto que, la segmentación Frente-Fondo está orientada a la correcta clasificación de los píxeles de frente manteniendo el equilibrio con los píxeles de fondo. Por el contrario la segmentación Persona-Fondo [3] está orientada principalmente a la correcta clasificación de los píxeles de fondo, es decir, es preferible clasificar un píxel de fondo como persona antes que clasificar un píxel de persona como fondo. Por consiguiente, el uso de la segmentación Persona-Fondo añadiéndole la información de la segmentación Frente-Fondo, nos puede generar una segmentación más fiel y exacta eliminando a su vez los problemas generados por cada tipo de segmentación de forma independiente.

1.2. Objetivos

El principal objetivo de este proyecto es el desarrollo de un segmentador Persona-Fondo usando la información de la segmentación Frente-Fondo. Para ello, primero se realizará un análisis del Estado del Arte de la segmentación Persona-Fondo, es decir, la segmentación de la escena en dos clases de diferente valor semántico, fondo y persona, dando una mayor importancia a la correcta detección del fondo en detrimento del frente. Se prefiere la correcta clasificación de los píxeles de fondo, es decir, es preferible clasificar un píxel de fondo como persona antes que clasificar un píxel de persona como fondo.

Posteriormente el siguiente paso será, el estudio del Estado del Arte de la segmentación del Frente-Fondo de la escena, con el objetivo de aplicar las mismas técnicas y esquemas sobre la segmentación Persona-Fondo.

Por lo tanto, tras el estudio de ambos tipos de segmentación y observando sus ventajas e inconvenientes se tiene como objetivo, la mejora de la segmentación Persona-Fondo aprovechando la información obtenida en la segmentación Frente-Fondo. Se estudiarán diferentes formas de combinar las fuentes de información.

Para aclarar de una manera más precisa los objetivos se presenta un esquema de la evolución a seguir.

1. Estudio del Estado del Arte actual de los segmentadores Persona-Fondo y Frente-Fondo.
2. Selección de algoritmos de segmentación Persona-Fondo y Frente-Fondo.
3. Identificación de ventajas e inconvenientes de segmentación Persona-Fondo y Frente-Fondo.
4. Uso de la segmentación Frente-Fondo para mejorar la segmentación Persona-Fondo.
5. Uso de la segmentación Persona-Fondo para mejorar la segmentación Frente-Fondo.
6. Evaluación de ambas segmentaciones.
7. Evaluación del algoritmo propuesto.
8. Conclusiones y trabajo futuro.
9. Elaboración de la memoria del proyecto.

1.3. Medios

Los medios necesarios para la realización de este proyecto han sido facilitados por el grupo de investigación *Video Processing and Understanding Lab* (VPULab) del departamento de Tecnología Electrónica y de las Comunicaciones de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid (UAM). Los principales elementos utilizados para la realización de este proyecto han sido:

1. Parque de PCs (Windows/Linux) interconectados a través de la red de área local y con acceso a Internet y a los servidores VPULab.
2. Software para el desarrollo del proyecto, Matlab.
3. Base de datos de secuencias de vídeo.
4. Biblioteca de libros y revistas de Tratamiento Digital de Imágenes.

1.4. Estructura de la memoria

La memoria seguirá la siguiente distribución:

1. Introducción.

2. Estado del Arte.
3. Algoritmo.
4. Evaluación.
5. Conclusiones y trabajo futuro.

Capítulo 2

Estado del Arte

2.1. Introducción

En este capítulo de la memoria se procederá al estudio detallado del Estado del Arte relacionado con este trabajo. En la sección 2.2 se describe la segmentación Persona-Fondo. A continuación, en la sección 2.3 se han consultado y estudiado diversos surveys de las técnicas de segmentación Frente-Fondo. Para finalizar este capítulo, en la sección 2.4, se extraen conclusiones del mismo y se exponen las ideas principales que utilizaremos para el desarrollo del algoritmo que proponemos en este trabajo.

2.2. Segmentación Persona-Fondo

La segmentación Persona-Fondo consiste en identificar aquellas zonas de la imagen que con mayor probabilidad no son personas y por lo tanto pertenecen al fondo de la escena. [2]. De esta forma se consigue una segmentación orientada a determinar qué zonas de la escena no son persona mientras que los segmentadores Frente-Fondo están orientados a la correcta clasificación de los objetos de frente, por lo que son orientaciones completamente dispares.

Este tipo de segmentación puede ser de mucha utilidad en muchas aplicaciones de procesamiento automático de vídeo. Puede ser usado tanto como un pre-procesado como post-procesado en algoritmos de *tracking*, estimación de densidad de personas, etc. El objetivo de este tipo de segmentador como ya se ha comentado antes, es la determinación de las áreas donde no hay personas penalizando los píxeles clasificados incorrectamente como fondo cuando son persona con mayor penalización. Este proceso nos da una máscara final con sesgo en el fondo y no en la persona como suele ser en el caso de los algoritmos de segmentación y detección tradicionales.

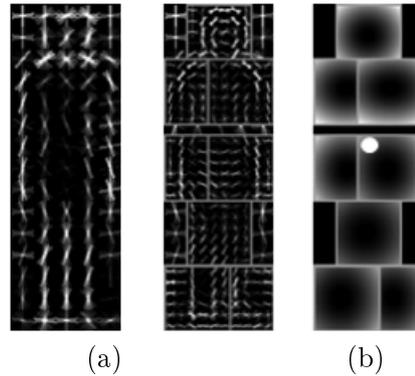


Figura 2.1: Modelo Holístico o persona entera (a), Modelo por partes al doble de resolución (b) [2]

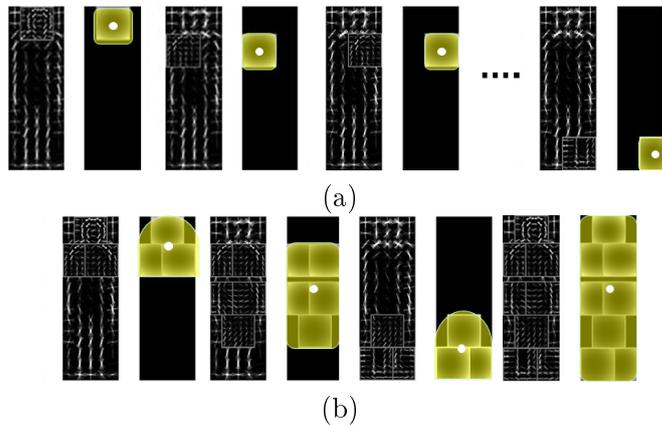


Figura 2.2: Modelo Independiente (a), Modelo Dependiente (b) [2]

El segmentador Persona-Fondo trata de garantizar que ningún píxel de Persona sea clasificado como fondo, pese a clasificar erróneamente píxeles cercanos a la persona que son fondo como Persona. Por ello los autores proponen detectar cualquier posible parte del cuerpo o combinaciones de estas, garantizando que cualquier parte de la escena que pueda contener una persona es clasificada como frente.

En este caso el detector utilizado en este segmentador [4], está constituido por un modelo de persona entera u holístico y un modelo basado en partes al doble de resolución que permite refinar la detección inicial. El modelo final permite flexibilidad en la localización de cada una de las partes del cuerpo mediante modelos “deformables”. La Figura 2.1-(a) representa el modelo de cuerpo u holístico, mientras que 2.1-(b) representa el modelo por partes al doble de resolución y el modelo de deformación de cada parte.

En el algoritmo de segmentación Persona-Fondo [2], se describen cinco métodos

o combinaciones de partes del cuerpo distintos, IBP (*Independet Body Part*) Fig 2.2(a)[2], DBP (*Dependet Body Part*) Fig 2.2:(b)[2], las versiones extendidas IEBP (*Independet Extend Body Part*) y DEBP (*Dependet Extend Body Part*) y la última DEBP-P que es la versión con pos-procesamiento.

Si consideramos el detector por partes multiescala, ver Figura 2.1, por cada píxel, escala y parte del cuerpo tenemos la confianza $P_n(x, y, s)$ de dicho píxel. Cada parte del cuerpo se modela por la respuesta obtenida del filtrado HOG F_n para la parte del cuerpo n y un vector de dos dimensiones v_n , que define la posición relativa de la parte del cuerpo (donde está anclada), respecto a la posición del cuerpo entero. Además, para permitir flexibilidad o “deformación” de las partes del cuerpo, se modela la posición de cada parte del cuerpo con un vector de cuatro dimensiones d_n con los coeficientes de una función cuadrática, que define el coste para cada posible emplazamiento de la parte del cuerpo con respecto al cuerpo total (donde está anclado). El valor de la confianza final se obtiene con la resta del valor de F_n menos su deformación definida por d_n .

El método IBP se define para ocho partes independientes y se define la posición de anclaje de cada parte respecto de ella misma, ver Figura 2.2-(a). Para mejorar la robustez de la detección se usa el método DBP, usando D_m con $m = 1 \dots M$ partes dependientes como combinación de las partes independientes. Cada parte del cuerpo generada por varias partes independientes genera a su vez una parte dependiente, la cual tendrá los mismos datos que una parte independiente simple. Se escogen cuatro combinaciones de partes independientes $M = 4$ para mantener la correlación con el cuerpo humano, cabeza, tronco, piernas y el cuerpo entero. Para mejorar la confianza y recuperar detecciones fallidas se decide extrapolar las detecciones reutilizando la información de otras partes dependientes. Cada parte dependiente D'_m es dada por el máximo entre la parte dependiente original D_m y la media de las otras partes dependientes, todas ellas relativas a la parte dependiente D_m .

En el caso de que solo podamos observar dos partes dependientes del cuerpo, se pueden recuperar y posteriormente normalizar las partes restantes de cada persona. Una vez se tienen los métodos IBP y DBP, se pueden extender los mapas de confianza de cada parte del cuerpo para obtener los métodos IEBP y DEBP. En el caso del primero se extienden las partes independientes y en el segundo caso se extiende la combinación de partes independientes. Una vez hemos obtenido la confianza del cuerpo entero, se escoge el mayor valor de confianza de todas las escalas $C(x, y)$. Se presenta ahora en la Figura 2.3, el resultado visual del método DEBP en escala de grises Figura 2.3-(b) y en color Figura 2.3-(c) y la máscara final post-procesada Figura 2.3-(d).

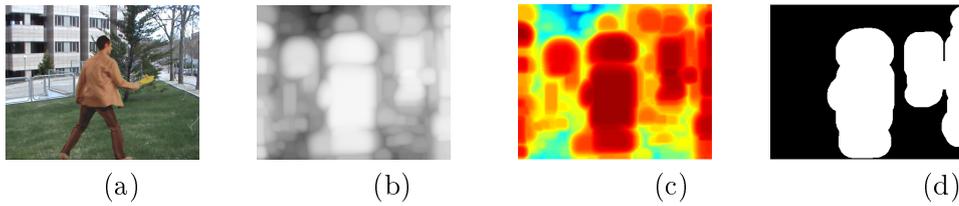


Figura 2.3: Imagen Original (a) Mapa de Confianza escala de grises (b) Mapa de Confianza en color (c) Máscara Final Método DEBP-P (Imagen obtenida dataset VPU UAM).

Una vez se ha obtenido el método DEBP, se propone una versión final pos-procesada de DEBP que consiste en una máscara final obtenida binarizando $C(x, y)$, siendo posteriormente post-procesada con una erosión con un kernel con forma de disco, con el tamaño de la parte del cuerpo menor dentro de cada escala (múltiples inter escalas). Tras esto se analizan las componentes conexas para quitar las zonas que sean más pequeñas que la menor parte del cuerpo detectable. Después se dilata con el mismo disco pero del tamaño de la mayor parte del cuerpo dentro de cada escala, para generar la máscara final. Este es el método DEBP-P.

Finalmente, se obtiene un segmentador Persona-Fondo el cual busca clasificar por completo la persona pese a cometer fallos clasificando fondo como frente, su principal ventaja es la eliminación de otros objetos de frente que no sean personas dado que solo busca personas. Una desventaja que tiene este segmentador es la sobre estimación o detección de posibles personas para lograr su objetivo de asegurar que zonas de la escena no presentan personas.

2.3. Segmentación Frente-Fondo

En esta sección se va a realizar el estudio del Estado del Arte de la segmentación Frente-Fondo, centrándonos en el tipo *Background Subtraction* (BGS). En primer lugar, se hará una introducción y definición de BGS, sección 2.3.1. A continuación, se identifican los principales problemas que nos podemos encontrar en la segmentación Frente-Fondo en sección 2.3.2. Posteriormente se realizará una clasificación de los métodos encontrados en el Estado del Arte en función de los procesamientos realizados y metodologías utilizadas, sección 2.3.3. Para finalizar la sección se expondrán conclusiones obtenidas en el estudio del Estado del Arte, sección 2.3.4.

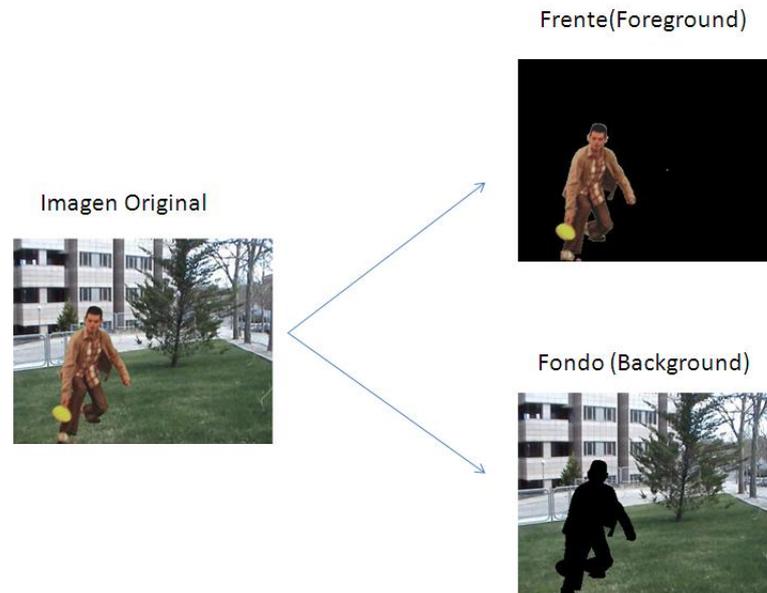


Figura 2.4: Esquema resultado de una segmentación en las dos clases Fondo y Frente (Imagen obtenida dataset VPU UAM)

2.3.1. Introducción

La segmentación Frente-Fondo tiene como objetivo la discriminación entre objetos del primer plano (frente o *foreground*) de una imagen, del resto de los objetos o fondo (*background*). Existen muchas técnicas distintas de segmentación de secuencias de vídeo dependiendo del tipo de secuencias y de los objetivos en cuestión. Los segmentadores que se buscan para este proyecto están basados en modelado de fondo, siendo una serie de algoritmos que mantienen un modelo del fondo a través del cual, para cada imagen se pueden extraer (por comparación u otros métodos) los objetos en movimiento (primer plano o frente) y el fondo. La Figura 2.4 muestra un ejemplo básico de la separación entre frente o *foreground* del fondo o *background*.

En este proyecto se ha optado por el uso de los segmentadores de tipo *Background Subtraction* (BGS), siendo también conocido como *Foreground Detection*. BGS genera un modelo de fondo que posteriormente se usará para la obtención del frente de la imagen original. Puesto que este tipo de algoritmia lleva mucho tiempo siendo estudiada y se han producido múltiples avances y mejoras, es necesario clasificar los distintos algoritmos además de hacer referencia a las dificultades con las que nos podemos encontrar en su uso [5, 6, 7, 8, 9].

Previamente a dar una clasificación comenzaremos con una simple explicación del proceso BGS, [10] que normalmente está formado por tres fases:

1. Inicialización del modelo de Fondo (*Background Initialization*): Su función es crear el modelo de fondo, si bien es cierto, esta fase suele estudiarse de manera independiente dado que normalmente se asume que se puede generar con las primeras frames del vídeo sin ningún objeto. Claramente esto es el caso ideal puesto que en los escenarios reales no tiene por qué iniciar el vídeo sin objetos de frente al principio. Normalmente se inicia el modelo con la primera frame o con un modelo de fondo obtenido a través de un conjunto de frames, las cuales pueden o no tener objetos de frente.

2. Actualización del modelo de Fondo (*Background Maintenance*): Su función es la actualización del modelo de fondo, puesto que a lo largo de la secuencia del vídeo hay elementos y objetos de frente que aparecen y desaparecen. Además esta actualización ayuda a la adaptación a los problemas comunes de esta temática que más tarde serán desarrollados. Es muy común encontrar en este tipo de algoritmos algunos parámetros que pueden ser configurados dependiendo del tipo de secuencia que se quiera procesar, como por ejemplo la tasa de aprendizaje o *Learning rate*. A menudo encontramos este parámetro representado por ' α ', lo que nos dice cuanto porcentaje de información del modelo de fondo y cuanta del frame actual usamos para generar el nuevo modelo de fondo. La ecuación 2.1 es un ejemplo de fórmula básica de actualización del modelo de fondo siendo β_{t+1} el nuevo modelo de fondo, I_t la imagen actual y β_t el modelo de fondo a actualizar.

$$\beta_{t+1}(x, y) = (1 - \alpha)\beta_t + \alpha I_t(x, y) \quad (2.1)$$

Otro posible parámetro es la tasa de actualización indicándonos cada cuanto modificamos el modelo de fondo, siendo muy subjetivo además de dependiente de la secuencia a procesar.

3. Detección de Frente (*Foreground Detection*): Su función es el etiquetado de los píxeles, en función de un umbral u otro posible factor, en píxel de fondo o de frente. La Figura 2.5 muestra un ejemplo esquemático de BGS, en el cual se puede apreciar el proceso de segmentación de una secuencia.

BGS tiene múltiples aplicaciones como para la detección de abandono de objetos, computo de estadísticas de tráfico de las carreteras [11], en aeropuertos[12]. Otro posible ámbito es la vigilancia de animales [13]. Otros posibles ámbitos pueden ser aplicaciones en las que interactúan las personas y las cámaras (Microsoft's Kinect), codificación de vídeo [14] y muchas otras.

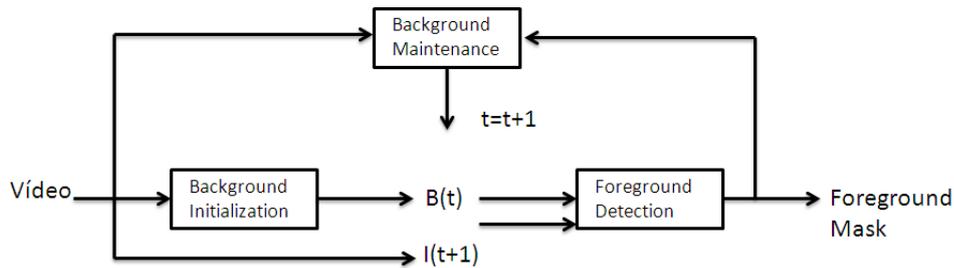


Figura 2.5: Esquema BGS básico de una secuencia de vídeo, incluyendo las etapas de Inicialización y Mantenimiento del Fondo al igual que la fase de Detección de Frente

2.3.2. Problemática en Background Subtraction

En el Estado del Arte de segmentación BGS, se han identificado típicos problemas asociados con entornos reales o no controlados. Puesto que obtener las condiciones ideales para la segmentación es muy difícil, siendo algunas de estas una iluminación constante, un fondo estático, etc. Por tanto se identifican los posibles problemas que podemos encontrar [15]:

1. Ruido: Muy dependiente del dispositivo de captura usado, así como la codificación usada.
2. *Jitter*: Movimiento de la cámara durante la grabación ya sea por viento, movimiento de la plataforma donde permanece la cámara, etc.
3. Auto configuración de las cámaras: El modelo automático de una cámara produce un descenso en los niveles de color u otros parámetros que afectan a la calidad de la grabación o captura.
4. Iluminación: La iluminación no suele ser constante, por lo que nuestro algoritmo ha de adaptarse a dichos cambios no siendo siempre una tarea fácil, sino que es uno de los principales problemas para estos algoritmos.
5. *Bootstrapping*: Durante el tiempo de entrenamiento el fondo no está disponible. Por lo que no se puede representar una imagen de fondo durante este periodo..
6. Camuflaje: Algunas características de los píxeles de frente pueden ser asumidas por el modelo de fondo y por tanto no se puede discriminar si es frente o fondo.
7. Foreground aperture: Cuando algún objeto de frente tiene regiones uniformes de color, por lo que en esas regiones podemos no reconocer cambios y por tanto, no clasificarlo como frente cuando si lo es.

8. Movimiento de objetos de fondo: Algunos objetos de fondo pueden sufrir una modificación en su posición, pese a esto siguen siendo de fondo pudiendo ser erróneamente clasificados como frente.
9. Inserción de objetos en el fondo: Aparición de nuevos objetos que debemos clasificar como fondo, pese a que no estuvieran desde el principio en la secuencia.
10. Fondos dinámicos: Un claro ejemplo son las hojas de un árbol que se mueven, pese a que lo hagan no son elementos de frente sino de fondo.
11. Fantasmas: Cuando el aprendizaje del algoritmo no es suficientemente rápido y se siguen clasificando como frente elementos que en un momento anterior pertenecían a este, pero en el actual no.
12. Frente estático: Cuando un elemento de frente se queda parado y es absorbido por el fondo, por lo que se clasifica erróneamente como fondo.
13. Sombras: Las sombras que generan los elementos de frente se denotan erróneamente como frente.

2.3.3. Clasificación del Estado del Arte de Background Subtraction

Una vez se conoce el funcionamiento a grandes rasgos de BGS y las posibles problemáticas que se pueden encontrar, se procede a la realización de una clasificación de los algoritmos. Una de las más comunes es la diferenciación entre recursivos y no recursivos [7, 9], otra posible la distinción entre modelos básicos, los cuales no tienen ningún parámetro modificable, y los paramétricos o no paramétricos [8]. Pero dado que el desarrollo en esta temática ha generado diversas ramas, se necesita una clasificación más detallada. Por lo que siguiendo [5], se ha decidido clasificar los algoritmos de segmentación Persona-Fondo en dos grandes grupos principalmente. El primero de estos sería Modelos Tradicionales, en los cuales están englobados los primeros segmentadores desarrollados, y los cuales tienen una operativa de mayor simplicidad que los otros. El segundo son los modelos Recientes o Actuales, los cuales se basan en los anteriores aplicando nuevas metodologías y procesamientos desarrollados. Uno de los motivos de esta clasificación, es que los métodos recientes han sido desarrollados para mejorar las deficiencias que los modelos tradicionales poseen, por lo que muchos de los métodos recientes son versiones mejoradas de los tradicionales. La Figura 2.6 nos muestra un esquema de la clasificación que se ha realizado de la segmentación Frente-Fondo.

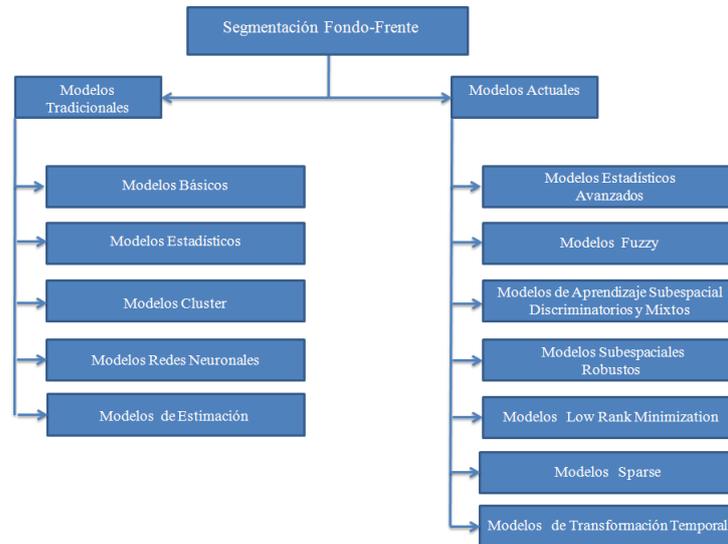


Figura 2.6: Esquema Clasificación Métodos de Segmentación Frente-Fondo

- Modelos Tradicionales: Dentro estos se pueden dividir en categorías según el tipo de procesamiento que realice cada modelo.
 - ✧ Modelos Básicos: Son los primeros desarrollados, con una complejidad mínima. Generan el modelo de fondo con la media, la mediana, o histogramas [8] y con una umbralización de la diferencia del modelo y la imagen actual se obtiene la salida.
 - ✧ Modelos Estadísticos: Se denominan así puesto que, utilizan funciones estadísticas para la generación del modelo de fondo. Ofrecen robustez frente a cambios de iluminación, fondo dinámico y al *jitter* de las cámaras. Este tipo puede clasificarse en:
 - Modelos Gaussianos: Este tipo de modelos asumen que la variación en la intensidad de los píxeles puede ser modelada por una Gaussiana siendo la primera aproximación realizada [16, 17]. Como ya se ha hecho referencia antes, la problemática encontrada en esta temática genero el desarrollo de nuevos modelos, trayendo esto el uso de mezclas de gaussianas, para la obtención de una mejor representación del modelo de fondo [18]. Otra aproximación englobada en estos modelos son los KDE (*Kernel Density Estimation*) [19], siendo un modelo no paramétrico, que estima las probabilidades de fondo a nivel de píxel.
 - Modelos basados en *Support Vector Models (SVM)*: Son modelos estadísticos mucho más sofisticados como puedan ser SVM [20], SVR [20],

SVDD [21]. En estos métodos cada nuevo píxel que aparece es evaluado para decidir si es fondo o no, la actualización del modelo de fondo se realiza mientras haya nuevos píxeles que incorporar al fondo. Para la caracterización y posterior clasificación de los píxeles como fondo o frente, utiliza las características del flujo óptico y la diferencia entre imágenes.

- Modelos basados en *Subspace Learning Models*: Estos modelos utilizan PCA (*Principal Component Analysis*), generan un modelo de fondo a partir de N frames, representado con la imagen media y una matriz de proyecciones que tiene los p primeros auto-vectores del PCA [22].

Los modelos Estadísticos ofrecen una mayor robustez a cambios de iluminación y a los fondos dinámicos. Los modelos Gaussianos y *Support Vector Models* son buenos para fondos dinámicos, mientras que los *Subspace Learning Models* son buenos para cambios de iluminación. En general los modelos estadísticos son los más usados por su compromiso entre su coste computacional y su rendimiento. En el caso de MOG (*Mixture of Gaussians*) son robustos para el auto configurado de parámetros de la cámara, *foreground aperture*, fondo dinámico, inserción de objetos en el fondo, fantasmas y frente estático.

- ✧ Modelos clúster: Suponen que cada píxel de la imagen, puede ser representado temporalmente por un clúster, siendo esto un conjunto. Se separan en tres categorías:
 - Modelos K-means: Asigna un grupo de clústers a cada píxel en el frame. Los clústers se ordenan de acuerdo con la probabilidad de que modelan el fondo y están adaptados para hacer frente a variaciones de fondo e iluminación [23].
 - Modelos Codebooks: Las muestras de cada píxel se agrupan en un conjunto de codewords en base a una métrica de distorsión de color junto con límites de brillo. El número de codewords es distinto en función de los cambios sufridos por el píxel [24].
 - clústering Secuencial Básico: Las intensidades de los píxeles se clasifican según un modelo clúster, posteriormente son calculados los centros y probabilidades de estos. Una vez es realizado este proceso los clústers con una probabilidad superior a un umbral son asignados como fondo [25].

Los Modelos clúster destacan por su buena adaptación a fondos dinámicos

- Modelos Híbridos: Estos modelos unen las ventajas de los modelos KDE y GMM (Modelo Mezcla de Gaussianas), siendo ambos modelos tradicionales gaussianos [36]. Este método usa la información de color de la secuencia, además de tener en cuenta la información de color de los vecinos de cada píxel.
- Modelos No Paramétricos: Se tienen dos aproximaciones para este tipo de modelos, ViBe y PBAS.
 - ▷ ViBe (*Visual Background Extractor*): En este método se genera el fondo agregando valores antes observados para cada píxel. Este método introduce diversas innovaciones, la primera es una política de selección aleatoria que otorga una vida útil de manera exponencial en descomposición sin problemas de los valores de las muestras que constituyen los modelos de píxeles. Otra innovación es la consistencia espacial que genera con los píxeles vecinos y por último da una mejora en la inicialización del modelo de fondo que en el segundo frame ya se dispondría de un modelo de fondo para realizar la segmentación [36]. Este método tiene problemas en escenarios con fondos oscuros, sombras y cambios frecuentes en el fondo.
 - ▷ PBAS (*píxel-Based Adaptive Segmenter*): Genera el fondo a partir de las últimas observaciones de los píxeles. La principal diferencia con el modelo anterior, es que la decisión en cuanto al modelo de fondo se realiza a nivel de bloque, y no de píxel. Una de las innovaciones introducidas con este método es que cada píxel tiene un parámetro de aprendizaje, por lo que con el umbral a nivel de bloque y a nivel de píxel cambia la estimación del fondo dinámicamente [37]. La gran diferencia con los modelos paramétricos, es que al tener los dos umbrales la actualización del modelo de fondo es dinámica y no se necesitan parámetros.
- Modelos Multi-Kernel: Este método emplea múltiples kernels, además de distintas representaciones del color para poder tener un espacio que represente mejor las características WGKVS (*Weighted Gaussian Models*) [38].

Los Modelos No Paramétricos ofrecen una buena robustez ante la problemática del *jitter* de las cámaras.

- ✧ Modelos *Fuzzy*: Estos métodos reciben su nombre del hecho de que, toman en consideración las imprecisiones y las incertidumbres generadas en el

proceso de segmentación. Introducen conceptos *fuzzy* (difusos) en todas las etapas del modelado de fondo.

- Modelado de Fondo *Fuzzy*: De este tipo de métodos surgen dos aproximaciones T2-FMOG-UM y T2-FMOG-UV, para la incertidumbre de la media y varianza respectivamente. Ambos métodos parten de la aproximación MOG de los métodos tradicionales [39], nos permiten atenuar la variación de color que se generan por el movimiento de los objetos de frente [40]. Su robustez es mayor que la de las aproximaciones de las que parten, trabajando bien con fondos dinámicos y el jitter de las cámaras.
- Detección de Frente *Fuzzy*: Se puede usar una función lineal de saturación para evitar una decisión tajante en la clasificación de los píxeles de fondo o frente. Las dos aproximaciones divergen en la integral que realizan de los resultados, pero las características usadas en ambos casos son las mismas textura y color. La primera aproximación de las dos es la que aplica la integral Sugeno [41], y la otra la integral de Choquet [42]. Siendo ambos procesos matemáticos de la Teoría de la Medida Difusa. Tienen buena robustez cuando trabajan con escenarios de fondos dinámicos y sombras.
- Actualización de Fondo *Fuzzy*: La idea es que la actualización de fondo use la pertenencia del píxel a la clase de fondo o frente, obteniendo esta pertenencia del modelo anterior. Tenemos dos realizaciones de estos modelos, la primera adaptando la tasa de aprendizaje a la clasificación del píxel antes mencionada [43] y la segunda es la unión de dos normas de la actualización del fondo [44]. Este método otorga una mayor robustez en escenarios con cambios de iluminación y sombras.
- Características *Fuzzy*: La primera aproximación encontrada es la creación de una nueva característica *fuzzy correlogram*, el algoritmo se computa sobre el fondo y el frame actual y se obtienen los valores. Posteriormente se aplica el cálculo de *K-L divergence distance*, si las distancias fuesen menores que un umbral se consideran fondo. Las otras dos aproximaciones parten de la primera siendo los autores de la misma los que han desarrollado las siguientes [45, 46, 47].
- post-procesado *Fuzzy*: Puede ser aplicado a cualquiera de los resultados de los métodos anteriores modelos Fuzzy [48].

Estos métodos estacan por su buena adaptación a cambios de iluminación y fondos dinámicos en el Estado del Arte.

- ✧ Modelos de Aprendizaje Subespacial Discriminatorios y Mixtos: Reciben su nombre dado que, generan subespacios para la representación del modelo de fondo y la obtención del frente. Estos métodos parten del método tradicional antes mencionado, *Subspace Learning Models*, pudiéndose clasificar en Reconstructivos o Discriminatorios [49]. El objetivo principal de los métodos reconstructivos es incluir la variabilidad de los datos de entrenamiento obtenidos. Por contra los discriminatorios dan una reconstrucción supervisada de los datos, siendo este método es dependiente de los datos mientras que el primero no lo es.
 - Modelo Subespacial Discriminatorio: La aproximación IMMC (*Incremental Maximum Margin Criterion*), basado en un algoritmo discriminatorio de aprendizaje de subespacios incremental. La desventaja que tiene este método es que necesita el ground truth de los frames [50].
 - Modelo Subespacial Mixto: Este método une un modelo reconstructivo y otro discriminatorio para conformar un fondo más robusto, además de una clasificación más fina de los píxeles [51].
- ✧ Modelos Subespaciales Robustos: Reciben su nombre dado que, generan subespacios para la representación del modelo de fondo y la obtención del frente. En la literatura encontramos diversos métodos pero nos centramos en el RPCA (*Robust Principal Components Analysis*) puesto que los otros por el momento tienen menos desarrollo que este.
 - En este tipo de métodos nos encontramos diversas aproximaciones, como por ejemplo, el uso de PCP (*Principal Component Pursuit*) [52], existiendo otras variantes de RCPA aplicando otros conceptos.
- ✧ Modelos Low Rank Minimization: Estos métodos usan funciones de minimización de matrices entre la matriz de datos o imagen actual y el modelo de fondo obtenido. Son muy útiles en tareas con muchos datos sin embargo, sus actuaciones son a menudo degradadas por valores atípicos.
 - *Contiguous Outliers Detection*: Propone una formulación para detecciones atípicas en la representación *low-rank*, en el que el soporte y la matriz de valores atípicos de *low-rank* se estiman, este método es llamado DECOLOR (*Detecting Contiguous Outlier detection in the Low-rank Representation*) [53].
 - *Direct Robust Matrix Factorization*: Propone una matriz de factorización Directa y robusta (DRMF) asumiendo que una parte de la

matriz de datos esta corrupta por valores atípicos. Esto se realiza para obtener una estimación fiable de la matriz *low-rank* e identificar los valores atípicos [54]. En base a esta aproximación existen otras como *Direct Robust Matrix Factorization-Row* [54], *Probabilistic Robust Matrix Factorization* [55] y *Bayesian Robust Matrix Factorization*[56].

- ✧ Modelos *Sparse*: Recibe su nombre dado que genera la información de grandes datos en modelos con valores escasos. La suposición de escasa descomposición se hace sobre los datos observados. Se pueden dividir en las siguientes categorías los métodos:
 - Modelos *Compressive Sensing*: Se obtienen las mediciones de detección de compresión para una imagen mediante su representación K-sparse, pudiéndose aplicar estas representaciones por cada frame de la secuencia. El fondo se aprende y actualiza en una pequeña representación de dimensión compresada [57].
 - *Structured Sparsity*: Proponer lograr una separación entre fondo y frente usando una estructura *Sparsity*, que es una extensión estándar del concepto en aprendizaje estadístico y compresivo de detección [58].
 - *Dynamic Group Sparsity*: Usa una fórmula de aprendizaje DGS (*Dynamic Group Sparsity*), la idea es que en pocos datos distintos de cero normalmente no son aleatorios, pero tienden a ser un clúster como en una detección de frente. El algoritmo DGS de recuperación asume que los clúster formados son una unión de subespacios, obteniendo de manera eficiente una recuperación estable de datos con un número menor de datos que otros algoritmos [59].
 - *Dictionary Learning*: La suposición que se realiza es que un frame puede ser representado como una combinación lineal de vectores formando un diccionario. Por tanto la inicialización comienza con una imagen de fondo conocida o una estimación de ésta. La valía de este método depende de la robustez del entrenamiento del diccionario además de encontrar valores adecuados para estimar el fondo [60].
 - *Sparse Error Estimation*: En este caso tomamos *Background Subtraction* como un problema de estimación de la señal, donde la escasez de error se aplica después de la minimización de la norma de la diferencia entre el frame actual y el subespacio estimado de fondo, como una aproximación a la estructura de la norma subyacente. Por lo que el obtener el fondo se toma como un problema de recuperación del error escaso [61].

- ✧ Modelos de Transformación Temporal: Llamados así para trabajar las imágenes en el dominio frecuencial.
 - Transformada Rápida de Fourier (FFT): Se estima el fondo con marcas espectrales a partir de fondos multimodales. Esas marcas son luego utilizadas para detectar cambios en la escena [62]. En este caso se obtiene una robustez frente objetos de frente de poco contraste y escenas dinámicas.
 - Transformada Discreta de Coseno (DCT): Genera un fondo a través de la descomposición de frecuencias del historial del píxel. Se calculan las componentes de la DCT para la frame actual para compararlos y obtener un mapa de distancias, el mapa es umbralizado y se determina el fondo [63].
 - *Walsh Transform (WT)*: Obtiene el fondo utilizando aplicando MOG sobre bloques de múltiples tamaños aplicando esta transformación, para obtener las características. Esto provoca tener una alta correlación espacial entre ellas [64].
 - *Wavelet Transform (WT)*: Puede ser usado en modelado de fondo o detección de frente.
 - ▷ *Marr wavelet*: Para modelar el fondo utiliza *Marr wavelet kernel* y características obtenidas de una transformación binaria discreta para obtener el fondo. El modelo de fondo mantiene una muestra de la intensidad de cada píxel en la imagen, utilizándolo para estimar la función de densidad de probabilidad de la intensidad [65].
 - ▷ *Dyadic Wavelet*: Propone el uso de *Dyadic Wavelet (DW)* para detectar objetos del frente. La diferencia entre el fondo y el frame actual es descompuesto en componentes multiescala *wavelet*, siendo las características HSV. Estas son utilizadas para detectar el frente y la componente de saturación se usa para eliminar las sombras [66].
 - ▷ *Orthogonal non-separable Wavelet*: Se usan transformaciones ortogonales no separables de los frames de entrenamiento y se extraen la información apropiada para reconstruir el fondo. Para la actualización del fondo se usa un esquema de media móvil [67].
 - ▷ *Daubechies complex wavelet*: Utiliza la transformación wavelet compleja Daubechies y el modelo de fondo es la mediana y se actualiza

[68].

- *Hadamard Transform (HT)*: Propone una rápida inicialización del fondo a nivel de bloque sin recursividad, para la obtención del mejor modelo de fondo con el menor número de frames [69].

2.3.4. Conclusiones

Tras el estudio de esta sección se han obtenido diversas conclusiones acerca de esta temática, como pueden ser que claramente no existe un método de Background Subtraction que sea robusto a toda la problemática comentada anteriormente. Además, a medida que avanza el estudio en esta temática se obtienen métodos más precisos, al igual que costosos computacionalmente y conceptualmente que generan mejores resultados que los primeros.

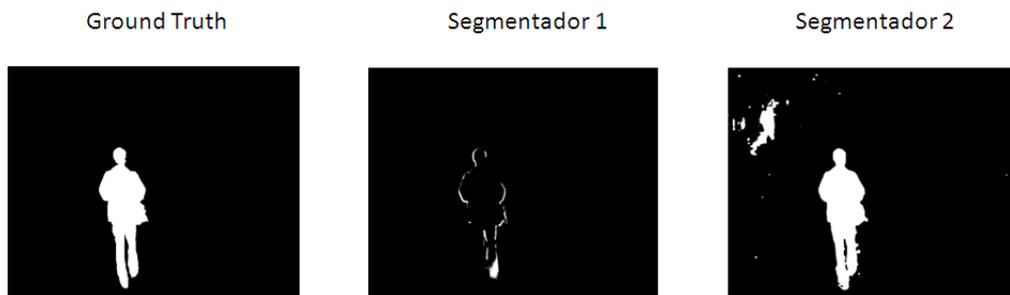


Figura 2.7: Máscara Salida Segmentador Frente-Fondo para dos segmentadores de distinta metodología, siendo el segundo de mayor calidad

En la Figura 2.7, se pueden apreciar dos problemas que nos podemos encontrar en los segmentadores Frente-Fondo, por ejemplo, el segmentador 1 no rellena la persona como se muestra en la imagen de *ground truth*, el segmentador 2 mantiene una zona clasificada como frente cuando ya no lo es.

Los algoritmos que parecen capaces de solventar a la vez diversas problemáticas son ViBe [36], PBAS [37] y SOB-SC [31]. Por tanto tras analizar las clasificaciones e informaciones relativas a cada algoritmo, podemos concluir que cada algoritmo está preparado para hacer frente a algunas que no todas las problemáticas antes descritas. Cada algoritmo tiene deficiencias respecto al tipo de secuencia, inclusive en su funcionamiento. Como se puede apreciar en la Fig 2.7, algunos segmentadores no rellenan la persona u objeto de frente detectado en contraposición otros actualizan muy tarde y quedan partes marcadas como frente que no deberían de ser o se añade ruido.

2.4. Conclusiones

En la sección 2.2 se ha estudiado el segmentador Persona-Fondo, el cual es el utilizado en este proyecto. En la sección 2.3 se ha estudiado los segmentadores Frente-Fondo, se han identificado los diferentes problemas que presentan estos en la sección 2.3.2. A continuación, en la sección 2.3.3 se ha realizado una clasificación de los mismos, en dos grandes grupos los tradicionales y los actuales, basándose muchos de los segundos en los tradicionales. En esta clasificación se han remarcado las principales ventajas o mejoras que introducen cada uno de los segmentadores mencionados.

Se ha hablado en este capítulo de la segmentación Persona-Fondo y la sección Frente-Fondo, siendo la primera un tipo de segmentación orientado a la clasificación correcta de las zonas en las que no hay personas. Si bien es cierto, no es un tipo de segmentación con un alto nivel de estudio, por lo que se dispone de la implementación utilizada en este proyecto ya explicada en secciones anteriores del capítulo.

En cuanto al segundo tipo de segmentación estudiada, es un tipo de segmentación orientado a la correcta clasificación de los objetos de frente, buscando una gran precisión. En este caso, nos encontramos con múltiples aproximaciones con todo tipo de metodologías y fundamentos. Por contra, se tienen múltiples problemáticas a resolver por los algoritmos como ya se ha explicado en una sección anterior del capítulo.

El objetivo principal de este PFC es el desarrollo de, un algoritmo de segmentación Persona-Fondo usando la información de la segmentación Frente-Fondo. Para ello en el siguiente capítulo se propone estudiar diferentes formas de combinar ambas fuentes de información para mejorar los resultados.

Capítulo 3

Algoritmo

3.1. Introducción

Una vez conocidos los dos tipos de segmentaciones explicadas previamente en el Estado del Arte, vamos a proceder a la explicación del sistema desarrollado en este proyecto final de carrera. La principal motivación de este desarrollo surge de suplir las deficiencias que se han podido reflejar en el apartado anterior para ambos tipos de segmentación. Debido a que la segmentación Persona-Fondo [2], está orientada a garantizar que ningún píxel de persona sea clasificado como fondo, se generan muchos falsos positivos, es decir, muchos píxeles de fondo son clasificados incorrectamente como persona. Por tanto, se generan muchos errores en este tipo de segmentación que es lo que se busca mejorar.

Cierto es que este tipo de segmentación no busca que sea lo más fina posible, entendiendo como segmentación fina aquella que delimita con muy alta precisión la silueta de los objetos de frente, sino que todos los píxeles pertenecientes a persona este cubiertos pese a tomar parte del fondo como persona, por lo que introducen errores en la segmentación. En cuanto a la segmentación de Frente-Fondo la problemática es más amplia que en el caso anterior ya estudiados en la sección 2.3.2, ya sean por sombras, fantasmas, cambios de iluminación, etc. A parte de estos problemas que nos podemos encontrar en las secuencias a analizar, existen otros factores intrínsecos a cada algoritmo. Esto es así dado que algunos de estos algoritmos son desarrollados para escenarios específicos, pudiendo no dar resultados correctos o aceptables en otras circunstancias.

Ejemplos: Dentro de las implementaciones de los algoritmos, muchos de estos algoritmos necesitan que las secuencias a analizar no tenga un comienzo en caliente, lo cual, significa que las secuencias empiecen sin objetos de frente en las primeras frames

y poder aprender el modelo de fondo sin interferencias no deseadas. Otro problema y para ser coherentes el principal para este proyecto, es el hecho de que se busca la segmentación de personas y este tipo de segmentación extrae todo tipo de objeto de frente ya sea persona o no.

Por tanto el motivo de este proyecto es unir a la información obtenida por la segmentación Persona-Fondo, la información obtenida por la segmentación Frente-Fondo tratando de obtener resultados mejores que ambos casos por separado. El resto del capítulo se distribuye en dos secciones. En primer lugar en la sección 3.2, describiremos la segmentación Persona-Fondo, mientras que en la sección 3.3, describimos la segmentación Frente-Fondo. Teniendo en cuenta el estudio realizado en el capítulo anterior hemos seleccionado un abanico que cubra en gran medida las posibilidades mencionadas en la clasificación realizada en el mismo.

Por lo que, en este capítulo describiremos el algoritmo propuesto para combinar las informaciones de ambas segmentaciones con el fin de mejorar la segmentación Persona-Fondo, en la sección 3.4. Para ello en primer lugar debemos describir cada uno de los segmentadores utilizados en el trabajo. En la sección 3.2 se explica el funcionamiento del algoritmo utilizado, en la sección 3.3 se exponen los segmentadores Frente-Fondo utilizados de manera escueta, para más información se dispone del apéndice B. Por último en la sección 3.4 describiremos el algoritmo propuesto que combinará la información de uno de los segmentadores Frente-Fondo con la información del segmentador Persona-Fondo, con el fin de mejorar la segmentación este último.

3.2. Algoritmo Persona-Fondo

El algoritmo de segmentación Persona-Fondo utilizado para este proyecto ha sido implementado por el VPULab [2], como ya se comentó en el capítulo anterior disponiendo de 5 versiones o implementaciones del algoritmo:

- IBP (*Independet Body Part*): En esta implementación se analizan las partes del cuerpo por separado.
- DBP (*Dependet Body Part*): Se unen partes independientes para generar 4 grandes grupos dependientes.
- IEBP (*Independet Extend Body Part*): En este caso se extiende el área que ocupa cada parte del cuerpo detectada en el método IBP.
- DEBP (*Dependet Extend Body Part*): Ocurre lo mismo que en el caso anterior pero para DBP.

- DEBP-P (*Dependet Extend Body Part Postprocess*): Aplica técnicas de post-procesado a DEBP para mejorar las máscaras de salida.

De entre las 5 versiones, en este proyecto se usa la DEBP, siendo uno de los motivos que al ser el mapa de confianza y no directamente una máscara binarizada aporta mayor información, el otro motivo es porque los resultados del algoritmo original post-procesados no siempre son mejores que sin post-procesar. Lo que se busca en este caso con el post-procesado, es en referencia a aspectos de tamaño, mientras que nosotros buscamos un procesado más “inteligente” utilizando la información de Frente-Fondo. Ya que en este tipo de segmentación solo usamos este algoritmo, precedemos a realizar una explicación del funcionamiento del algoritmo:

Antes de la ejecución del algoritmo se necesita configurar tres parámetros, dos de estos importantes para el resultado final y el otro para tener más o menos salidas disponibles:

- Tamaño del Modelo (*model.sbin*): Este parámetro hace referencia al tamaño de las personas que aparecen en la secuencia, si las personas fueran grandes el valor de este debería estar en 8, mientras que si las personas tienen un tamaño medio o pequeño el valor ha de ser 4.
- Escala (*scale_levels*): Dado que en las secuencias no tienen por qué aparecer todas las personas de igual tamaño, este parámetro nos permite buscar personas de diferentes tamaño en la secuencia.
- Umbrales (*thresholds*): Este parámetro nos indica el número de umbrales a evaluar.

En la Figura 3.1, se muestra un ejemplo de los posibles valores para los parámetros del algoritmo de segmentación Persona-Fondo. Se observa que la diferencia en el tamaño de las personas es un detalle importante en este algoritmo, al igual que la escala si tuviéramos diversos tamaños de personas en la escena.

Tras la elección de parámetros ejecutaría el algoritmo, las secuencias no son en vídeo, sino que son los frames.

- Se genera una máscara de procesamiento en función de los parámetros antes seleccionados y se genera un vector de características (HOG, *Histogram of Oriented Gradients*) sobre el frame actual y el modelo.
- Tras esto se empiezan a obtener las puntuaciones del tronco enviando como parámetros la imagen, las características antes obtenidas, y los parámetros a los siguientes detectores. El detector usado en este algoritmo de tipo HOG [70], el



Figura 3.1: Ejemplo visual parámetros algoritmo segmentación Persona-Fondo.

cual genera un vector de dos dimensiones con la posición relativa de la parte a buscar con la de anclaje (donde está el cuerpo completo). Las puntuaciones de confianza obtenidas por los detectores se obtienen:

$$P_n(x, y, s) = F_n(x, y, s) - \langle d_n, \phi(dx_n, dy_n) \rangle \quad (3.1)$$

con

$$(dx_n, dy_n) = (x_n, y_n) - (2(x_0, y_0) + v_{n,0}) \quad (3.2)$$

dándonos el desplazamiento relativo

$$\phi(dx, dy) = (dx, dy, dx^2, dy^2) \quad (3.3)$$

Por lo que el proceso a seguir es la detección de una de las partes dependientes en función de las otras, dándonos cuatro combinaciones.

- ✧ Detector de Medio Cuerpo o Tronco.
 - ✧ Detector de Medio Cuerpo o Tronco basado en cabeza y hombros.
 - ✧ Detector de Medio Cuerpo o Tronco basado en piernas.
 - ✧ Detector de Medio Cuerpo o Tronco basado en cuerpo entero.
- Una vez hemos obtenido las puntuaciones de los tres últimos detectores para obtener la confianza media de estos, mirando todas las escalas S y todos los

píxeles N . Se obtienen los mapas finales de confianza de la parte del cuerpo:

$$C(x, y) = \max_{1 \dots N} \max_{1 \dots S} P_n(x, y, s)$$

- Tras la obtención de las confianzas medias de los detectores basados en las otras partes de cuerpo, esto hace referencia a *Extend* del nombre del algoritmo, puesto que la información de otras partes si se tiene en cuenta. Se combinan estas las puntuaciones medias con las puntuaciones obtenidas por el detector de la parte correspondiente que se está buscando, en este caso el tronco.
- Se realiza la misma operación para la cabeza y hombros, las piernas obteniendo la media de estos dos con el primero ejecutado. Por último se realiza para el cuerpo entero.
- Por último para cada umbral de los antes establecidos, se generan la máscara de salida y la máscara de salida post-procesada obtenidas de aplicar la umbralización al mapa de confianza.

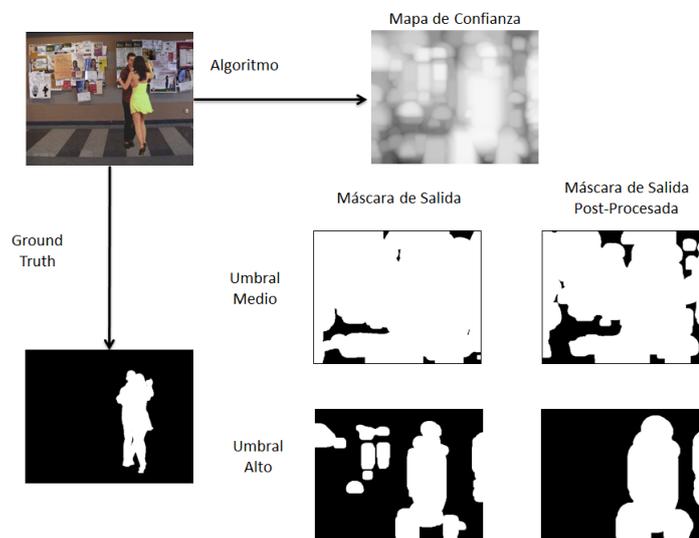


Figura 3.2: Desarrollo del algoritmo Persona-Fondo utilizado para este proyecto

En la Figura 3.2, podemos observar el funcionamiento del segmentador Persona-Fondo a través de la secuencia original. Se obtienen por tanto, los mapas de confianza del mapa original con el procedimiento antes contado, después los mapas son umbralizados para los umbrales decididos en el segmentador.

3.3. Algoritmo Frente-Fondo

3.3.1. Introducción

En esta sección 3.3, se indican los algoritmos seleccionados para el desarrollo de este proyecto. La mayor parte de los algoritmos, se han encontrado como código libre disponible en Internet BGSLibrary¹, siendo explicado el funcionamiento de la misma, los algoritmos disponibles y los escogidos en la sección 3.3.2. Además de estos algoritmos, se dispone de una algoritmo facilitado por la TUB (*Technical University of Berlin*), sección 3.3.3.

3.3.2. Librería BGSLibrary

Esta librería incluye 33 algoritmos de BGS desarrollados cada uno por su autor y unificados en esta aplicación. La clasificación que realiza el autor de los algoritmos, coincide con la realizada en el Estado del Arte en la sección 2.3.3 del capítulo 2.3. Dado que el objetivo de este capítulo no es la explicación de cada uno de los algoritmos utilizados, se puede acudir al apéndice B en el cual se presenta un resumen de cada uno. Por tanto, los algoritmos disponibles se muestran clasificados en la “rama” de las estudiadas en el Estado del Arte en el capítulo 2.3 al que pertenecen.

- Métodos Básicos: StaticFrameDifferenceBGS, FrameDifferenceBGS, Weighted-MovingMeanBGS, WeightedMovingVarianceBGS, AdaptiveBackgroundLearning, DPMeanBGS, DPAdaptiveMedianBGS, DPPratiMediodBGS.
- Métodos *Fuzzy*: FuzzySugenoIntegral, FuzzyChoquetIntegral, LBFuzzyGaussian.
- Métodos Estadísticos de una Gaussiana: DPWrenGABGS, LBSimpleGaussian.
- Métodos Estadísticos de múltiples Gaussiana: DPGrimsonGMMBGS, MixtureOfGaussianV1BGS, MixtureOfGaussianV2BGS, DPZivkovicAGMMBGS, LB-MixtureOfGaussians.
- Métodos *Type-2 Fuzzy*: T2FGMM (UM y UV), T2FMRF (UM y UV).
- Métodos Estadísticos con Características de Color y Textura: DPTextureBGS, LbpMrf, MultiLayerBGS, SJN_MultiCueBGS.
- Métodos No Paramétricos: GMG, VuMeter, KDE, IMBS.
- Métodos *EigenSpace*: DPEigenbackgroundBGS.

¹(<https://code.google.com/p/bgslibrary/> o <https://github.com/andrewssobral/bgslibrary>)

- Métodos Neuronales y Neuronales *Fuzzy*: LBAdaptiveSOM, LBFuzzyAdaptiveSOM.

De todos los que se muestran previamente, se han escogido 25 entre los cuales se han seleccionado al menos uno de cada rama. Se muestran los algoritmos seleccionados para el desarrollo de este proyecto, siendo estos seleccionados por diversos motivos como son la falta de generación de modelo de fondo, el uso de metodologías que no generan imágenes de fondo, la imposibilidad de configurar el segmentador para tener un mejor resultado del original, etc.

- Métodos Básicos: AdaptiveBackgroundLearning, DPMeanBGS, DPAdaptiveMedianBGS, DPPratiMediodBGS.
- Métodos *Fuzzy*: FuzzySugenoIntegral, FuzzyChoquetIntegral, LBFuzzyGaussian.
- Métodos Estadísticos de una Gaussiana: DPWrenGABGS, LBSimpleGaussian.
- Métodos Estadísticos de múltiples Gaussiana: DPGrimsonGMMBGS, MixtureOfGaussianV1BGS, MixtureOfGaussianV2BGS, DPZivkovicAGMMBGS, LBMixtureOfGaussians.
- Métodos *Type-2 Fuzzy*: T2FGMM (UM y UV), T2FMRF (UM y UV).
- Métodos Estadísticos con Características de Color y Textura: MultiLayerBGS.
- Métodos No Paramétricos: GMG, VuMeter, KDE.
- Métodos *EigenSpace*: DPEigenbackgroundBGS.
- Métodos Neuronales y Neuronales *Fuzzy*: LBAdaptiveSOM, LBFuzzyAdaptiveSOM.

Se expondrá el uso e instalación de la aplicación BGSLibrary por encima haciendo referencia a sus beneficios y desventajas. La plataforma está disponible para las plataformas Windows y Linux, en este caso a elección del proyectando se decidió usar la plataforma Linux puesto que, tras los conocimientos obtenidos en la carrera encontraba más fácil el uso de esta plataforma. Previa descarga de la biblioteca es necesario comprobar si se tiene instalado la biblioteca OpenCV puesto que, será necesaria para la posterior ejecución y modificación de los segmentadores. Si no estuviese instalado en la plataforma no hay ningún problema en la obtención puesto que es un software libre.

Una vez terminado el proceso anterior se procede a la instalación de la BGS Library ejecutando los comandos básicos de Linux `cmake`, `make`. Se pueden dar diversos

problemas en la instalación puesto que esta biblioteca necesita librerías que puede que la plataforma Linux no tenga instaladas. Tras la finalización de la instalación se tiene una carpeta contenedora de todos los algoritmos antes mencionados, además del código del programa principal para su ejecución. En Linux la ejecución se realiza a través de la Shell puesto que son scripts que previamente se han de compilar, no teniendo que implementar ningún Makefile puesto que ya se dispone de ello.

Para cada algoritmo se dispone del código fuente, aunque dado que OpenCV trabaja con clases como en el lenguaje de programación JAVA, no se puede acceder a todas las fuentes de información. Además del código fuente se tiene un archivo .XML que contiene los parámetros de ejecución, dándonos pie a configurar cada algoritmo según las necesidades de la secuencia con la que se trabaja.

Para este proyecto ha sido necesaria la comprensión a grandes rasgos de los algoritmos dado que, se ha necesitado modificar el código de cada algoritmo para la obtención de las imágenes de los modelos de fondo y de las máscaras finales de salida. A colación de lo anterior, también se analizó el código para corroborar si los algoritmos generaban un modelo de fondo y se almacenaba, sino se debía modificar el código para que se consiguiese.

3.3.3. Algoritmo SGMM-SOD

Tanto el proyecto como el ejecutable de este algoritmo [71], ha sido facilitado por Rubén de las Heras y por la TUB (*Technical University of Berlin*). El uso de este algoritmo es debido a la buena calificación que ha obtenido en el portal de evaluación *Change Detection*², en la cual se evalúan las implementaciones de esta temática.

- El algoritmo SGMM-SOD [71]: Genera dos modelos de fondo con dos GMM, uno de ellos más preciso para segmentar los elementos en movimiento, corto plazo. Y otro, largo plazo, que se encarga de reconstruir la escena vacía. Ambos modelos tienen los mismos parámetros de configuración, exceptuando la tasa de aprendizaje. El modelo a corto plazo es el más rápido en cuanto a la adaptación. Cuando se detecta un nuevo píxel por el modelo de largo plazo y no por el de corto indica que un nuevo modo para las GMM se ha de generar, así se realiza la actualización del modelo.

Como ya se ha comentado con anterioridad, el proyecto y ejecutable está disponible para Windows y fue facilitado por la TUB.

²<http://www.changedetection.net/>

3.4. Algoritmo Propuesto.

En este apartado se procede a la explicación del algoritmo desarrollado en este proyecto. La principal motivación de este proyecto es el desarrollo de un algoritmo de segmentación Persona-Fondo [2] usando la información de la segmentación Persona-Fondo [5], es unir las ventajas que tienen cada una de las segmentaciones por separado e intentar suplir las deficiencias de ambas. Ya se ha hecho referencia a las desventajas independientes de cada una, como puede ser la generación de un gran número de falsos positivos o detecciones falsas por parte de la segmentación Persona-Fondo. Así mismo, la segmentación Frente-Fondo tiene un gran número de problemáticas a las que hacer frente ya antes mencionadas en el capítulo 2.3.

Con el objetivo mencionado, en primer lugar se planteó utilizar las técnicas de modelado de fondo de los segmentadores Frente-Fondo de tipo BGS, para modelar la confianza generada por el segmentador Persona-Fondo, sección 4.4.3. La segunda aproximación propuesta consiste en combinar directamente a nivel de máscaras binarias los resultados de ambas segmentaciones para mejorar la segmentación Persona-Fondo, sección 4.4.5. Por último, aprovechando el esquema generado en la segunda aproximación propuesta, se propuso la combinación “simétrica” con el objetivo opuesto, es decir, mejorar la segmentación Frente-Fondo con la segmentación Persona-Fondo, sección 4.4.6.

3.4.1. Primera Aproximación

La primera aproximación realizada se basa en aprovechar la evolución temporal generada en la segmentación Frente-Fondo, puesto que los segmentadores generan modelos de fondo los cuales son actualizados en función de la implementación. Con esta aproximación se busca reducir los falsos positivos generados por la segmentación Persona-Fondo, siendo estos idealmente errores aislados temporalmente o al menos, menos estables que la detección correcta de personas, puesto que, el objetivo de la misma es clasificar de manera correcta los píxeles de fondo, prefiriendo clasificar como persona algún píxel de fondo.

En la Figura 3.3 se muestra la evolución temporal de los mapas de confianza del segmentador Persona-Fondo original -(b) correspondientes a las frames del vídeo-(a), la ejecución de la primera aproximación para un segmentador-(c), otra ejecución de la primera aproximación con otro segmentador-(d) y la máscara que posteriormente deberíamos obtener o *ground truth* -(e).

En la Figura 3.4 se puede apreciar como los mapas de confianza originales tienen muchos falsos positivos y como los segmentadores Frente-Fondo suavizan a medida

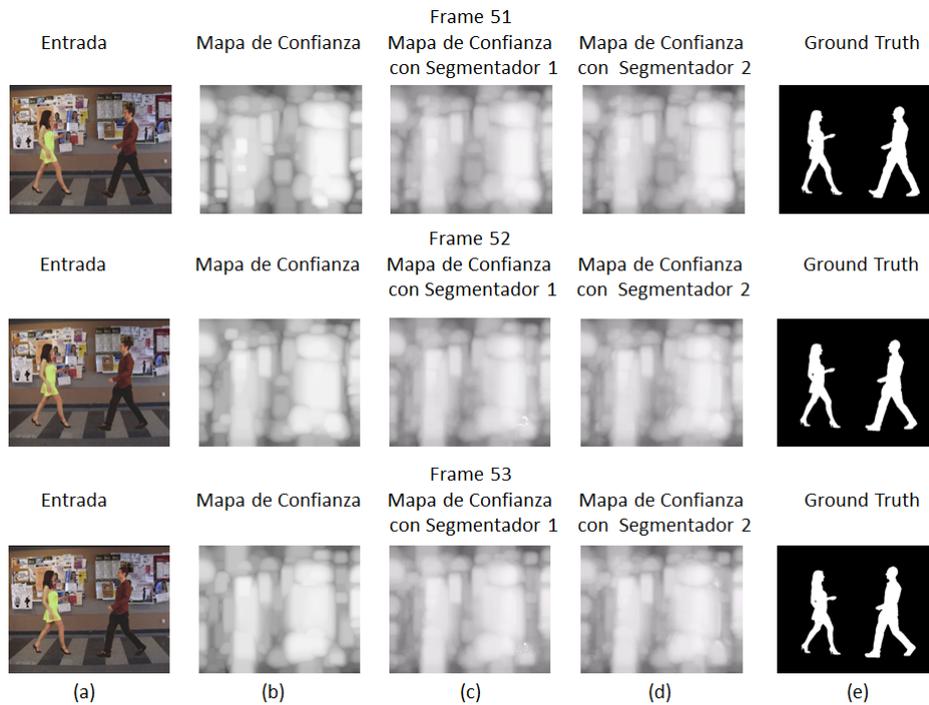


Figura 3.3: Ejemplo de ejecución de la primera aproximación. Frame entrada (a) Salida segmentador Persona-Fondo (b) Salida segmentador 1 Frente-Fondo (c) Salida segmentador 2 Frente-Fondo (d) *Ground Truth* (e)

que va avanzando la secuencia los falsos positivos marcado en amarillo. También se aprecia que, dado que los segmentadores Frente-Fondo generan el modelo de fondo y lo actualizan con el aprendizaje, se observa que los primeros falsos positivos generados se mantienen estables a lo largo de la secuencia marcados en rojo. Se indican en azul, personas que aparecen en la secuencia observando que son cubiertas por completo de manera burda.

En la Figura 3.5 se pueden observar las máscaras obtenidas en la ejecución de la primera aproximación, en las que cabe destacar que si se observan (c) y (d) en referencia a (b) se eliminan algunos falsos positivos encontrados en los mapas de confianza originales. Se aprecia, en cualquiera de las columnas (b) (c) y (d), que las personas son tapadas completamente por el algoritmo Persona-Fondo y se mantiene al usar los segmentadores de Frente-Fondo.

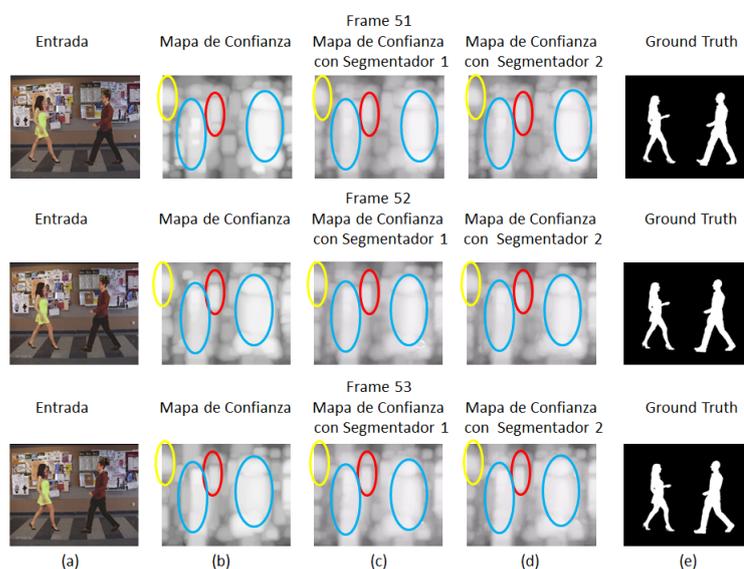


Figura 3.4: Mejoras visuales de los segmentadores Frente-Fondo de la primera aproximación. Frame entrada (a) Salida segmentador Persona-Fondo (b) Salida segmentador 1 Frente-Fondo (c) Salida segmentador 2 Frente-Fondo (d) *Ground Truth* (e)



Figura 3.5: Máscaras de Salida de la ejecución de la primera aproximación. Frame entrada (a) Máscara segmentador Persona-Fondo (b) Máscara segmentador 1 Frente-Fondo (c) Máscara segmentador 2 Frente-Fondo (d) *Ground Truth* (e)

Para tener una idea más clara de la aproximación se presenta ahora el desarrollo necesario para la ejecución de la misma:

1. Se lanza el algoritmo Persona-Fondo [2], sobre el vídeo deseado. Se obtienen por tanto, los mapas de confianza del vídeo, al igual que las máscaras de salida originales y las post-procesadas. Usando solo las máscaras sin post-procesado en el algoritmo.
2. Se lanza el algoritmo Frente-Fondo [5] que se quiera, sobre la secuencia de mapas de confianza previamente obtenida. Se configuran los parámetros del segmentador para obtener una mejora sobre los mapas de confianza anteriormente obtenidos, siendo este el caso en que se desean obtener los modelos de fondo generados por el segmentador.
3. Una vez tenemos toda esta información, disponiendo también del *ground truth* (máscara realizada a mano en la que se segmenta la escena perfectamente y se utiliza para la evaluación de los resultados del algoritmo).
4. Tras esto se evalúan los mapas de confianza originales y los modelos de fondo del algoritmo, utilizando el *ground truth* antes explicado. La evaluación será explicada en el siguiente capítulo de la memoria. También se evalúan de la misma manera las máscaras de salida obtenidas sobre el vídeo original.
5. Se presentan las curvas Precisión/Recall (dos puntuaciones de evaluación) en la Figura 3.6, sobre los píxeles de fondo y de frente, obtenidas de la evaluación de los mapas de confianza. La configuración realizada al algoritmo Frente-Fondo escogido debe generar que, la curva de los mapas de confianza obtenidos del modelo de fondo del segmentador sea mayor que la de los mapas de confianza originales.

3.4.2. Segunda Aproximación

A la vista de los fallos cometidos por el segmentador Persona-Fondo en la máscara binaria DEBP, se propone aprovechar las ventajas del segmentador Frente-Fondo directamente sobre la máscara binaria final. Para ello se propone filtrar o eliminar ruido o falsas detecciones utilizando como filtro selectivo sólo aquellas zonas de la imagen que han sido clasificadas como frente por el algoritmo de segmentación Frente-Fondo. En particular, se propone el uso de un filtrado morfológico por reconstrucción sobre la máscara DEBP y utilizando como marcador la máscara Frente-Fondo de los segmentadores Frente-Fondo. El filtrado por reconstrucción es una operación morfológica realizada con un operador local, el cual preserva los contornos de la imagen.

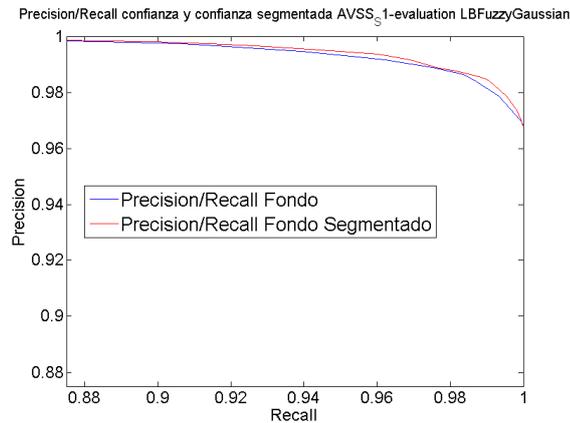


Figura 3.6: Gráfica Precisión/Recall primera aproximación

- Se tienen dos imágenes, en nuestro caso la máscara obtenida por el mapa de confianza y la del segmentador Frente-Fondo del mismo frame. Denotaremos las imágenes como señal X y señal Y indistintamente.
- El proceso de reconstrucción preserva las componentes conexas de X marcadas con Y. Existen dos posibles operaciones, Apertura por Reconstrucción y Cierre por Reconstrucción. En nuestro caso se ha utilizado la primera de las dos, Apertura por Reconstrucción. En la Figura 3.7 se muestra un ejemplo del funcionamiento del filtrado por Reconstrucción utilizado.
- Se utiliza conectividad 8 y se utiliza como marcador la máscara binaria generada por el segmentador Frente-Fondo.
- Se realiza después un proceso iterativo:
 - ✧ Se calcula el mínimo o se realiza la operación AND entre el marcador y la imagen.
 - ✧ Se dilata la imagen resultante del paso anterior.
 - ✧ Se calcula el mínimo entre la imagen dilatada y la imagen original, siendo la operación mínimo para imágenes lógicas la AND.
 - ✧ Repetir hasta que no exista variación significativa entre iteraciones consecutivas.

Se muestra ahora el proceso que se realiza para ejecutar esta aproximación.

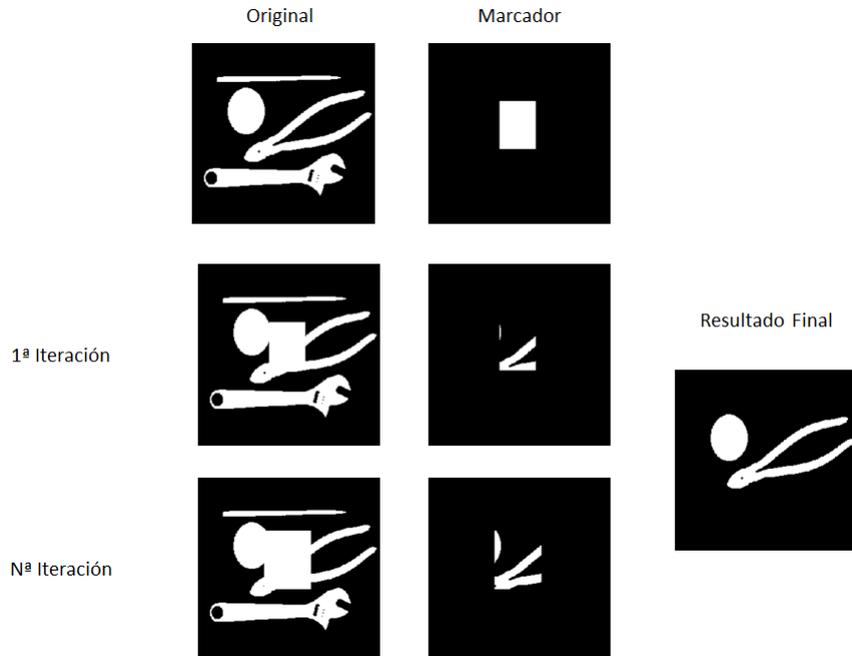


Figura 3.7: Ejemplo filtrado por reconstrucción explicativo del proceso a seguir.

1. Se lanza el algoritmo Persona-Fondo [2], sobre el vídeo deseado. Se obtienen por tanto, los mapas de confianza del vídeo, al igual que las máscaras de salida originales y las post-procesadas. Usando solo las máscaras sin post-procesado en el algoritmo.
2. Se lanza un algoritmo Frente-Fondo sobre el vídeo original, configurando los parámetros para obtener la mejora máscara de salida.
3. Una vez tenemos toda esta información, disponiendo también del *ground truth*.
4. Para la reconstrucción nos interesa obtener la máscara que se encuentre en un punto equitativo, siendo este bueno para fondo como para frente. Por lo que se usan las curvas F1Score (puntuaciones de la evaluación) para frente como para fondo como se aprecia en la Figura 3.8, y el punto donde intersequen es el que se busca indicándonos así el umbral a escoger.
5. Una vez obtenido el umbral (específico para los mapas de confianza originales) se pasa a leer dichas máscaras obtenidas del mapa de confianza original.
6. Usando como marcador la máscara del algoritmo Frente-Fondo [5], queriendo reconstruir la máscara obtenida por el segmentador Persona-Fondo[2]. Por tanto

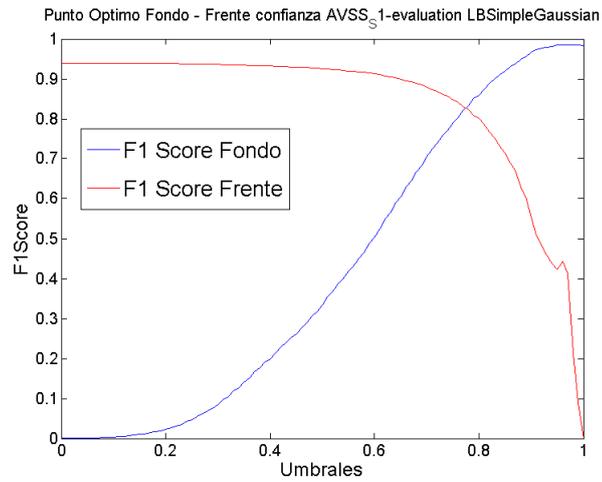


Figura 3.8: F1 Score de frente y de fondo obteniendo de su punto de intersección el umbral óptimo.

se busca obtener una máscara que cubra por completo a las personas, además de eliminar posibles errores en la máscara.



Figura 3.9: Reconstrucción máscara Persona-Fondo usando la información de la segmentación Frente-Fondo.

En la Figura 3.9, se puede apreciar una mejora de la máscara Persona-Fondo tras aplicar la información generada por la segmentación Frente-Fondo, ya que se eliminan muchos falsos positivos de la máscara original.

3.4.3. Tercera Aproximación

Tras la realización de la segunda aproximación y darnos cuenta que el mismo proceso podría aplicarse también para mejorar la segmentación Frente-Fondo, surge la tercera aproximación del proyecto. La única diferencia que se encuentra con la sección

4.4.5, es que en este la imagen a filtrar es la máscara generada por el segmentador Frente-Fondo y el marcador es la máscara generada por el segmentador Persona-Fondo. Por lo que, el proceso antes explicado es el mismo para esta aproximación. También en este apartado se va a mostrar el proceso de ejecución de esta aproximación.



Figura 3.10: Reconstrucción máscara Frente usando la información de la segmentación Persona-Fondo.

En la Figura 3.9, se puede apreciar una mejora de la máscara Frente-Fondo tras aplicar la información generada por la segmentación Persona-Fondo, ya que se eliminan ciertas clasificaciones de frente erróneas y se elimina ruido generado en la segmentación.

3.4.4. Conclusiones

En este capítulo, se han estudiado diversas posibilidades de unificar la información de ambas segmentaciones para los resultados de ambas. Cabe destacar, que la primera aproximación busca el uso de la información generada por los segmentadores Frente-Fondo para mejorar la información generada por los mapas de confianza. En el caso de la aproximación segunda y tercera, lo que se busca es la combinación de la información de ambas segmentaciones a través de un proceso independiente a las segmentaciones. Por lo que para concluir con este capítulo se muestra en la Figura 3.11, un esquema del funcionamiento general de las aproximaciones.

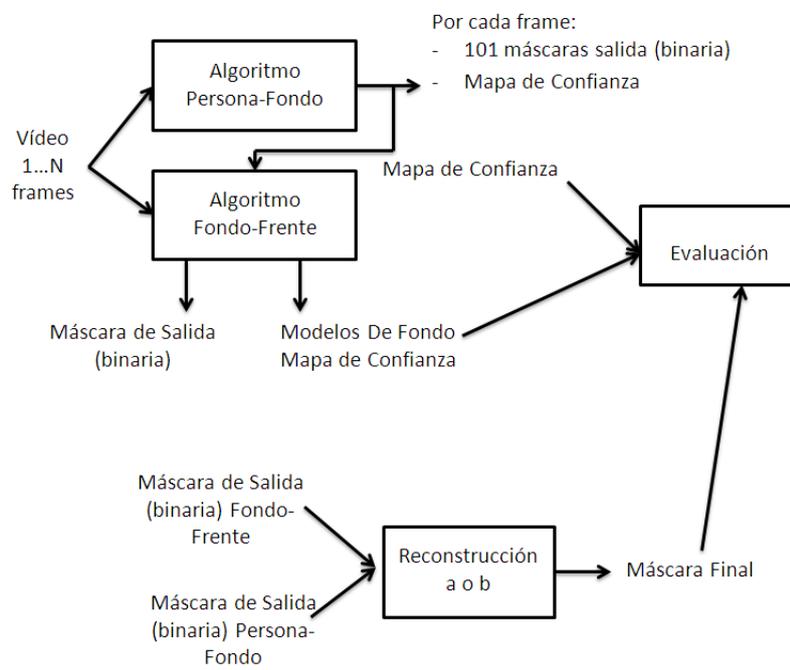


Figura 3.11: Esquema General del Algoritmo

Capítulo 4

Evaluación

4.1. Introducción

Como es común a todas las áreas de investigación, cada nuevo desarrollo realizado necesita ser evaluado ya no solo para ver sus resultados, sino para obtener una realimentación de a qué nivel se encuentra en comparación a las demás implementaciones y desarrollos del Estado de Arte. La segmentación de objetos en secuencias de vídeo ha generado un gran interés en los últimos años. Por lo que, se dispone de diversas estrategias para evaluar la calidad de nuestro algoritmo de segmentación. Las técnicas disponibles con este fin se pueden clasificar en: subjetivas [72] y objetivas [73]. Las técnicas subjetivas son en base al criterio humano de evaluación, por lo tanto, para aplicarse se necesita evaluar con cierto número de personas. Las objetivas, ‘*ground-truth*’, que es una segmentación anotada, normalmente hecha a mano, usada 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 [73]. 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 vídeo de características diferentes que constituyen el grupo de vídeos de prueba (*‘dataset’*) utilizados en el proyecto, grupo que explicaremos en el siguiente apartado.

Este capítulo está estructurado en tres secciones, la sección 4.2 contendrá la información del dataset utilizado. En la sección 4.3 se hablará de la métrica utilizada para la evaluación del sistema y por último en la sección 4.4 se presentan los resultados y

conclusiones de los mismos.

4.2. Dataset

Un *dataset* es el conjunto de elementos de entrada que se usan para la evaluación de una implementación, en nuestro caso está formado por distintos vídeos usados en las evaluaciones del Estado del Arte. Además de estos, en nuestro proyecto se han escogido otros vídeos del grupo VPULab (*Video Processing and Understanding Lab*) de la Escuela Politécnica Superior de la Universidad Autónoma de Madrid. En la Figura 4.1, se muestra un resumen de los vídeos que conforman el dataset del proyecto, indicando su procedencia frames totales y los que se evalúan.

Vídeo	Frames Total	Frames Evaluación	Biblioteca
Avss_S1	983	6	AVSS
BaileMulti	601	601	VPU
BaileUni	601	601	VPU
BusStation	1250	1250	CD
Corridor	5400	5400	CD
Cubicle	7400	7400	CD
DaSide0811-seq6	71	7	TUD-campus
DaSide0811-seq7	201	20	TUD-crossing
Office	2050	2050	CD
Park	600	600	CD
Pedestrians	1099	1099	CD
Pets_2009	795	5	PETS2009
Pets_S1	532	6	CD
Tenis1	307	307	VPU
Train56	103	5	TrecVid

Tabla 4.1: Resumen de secuencias de vídeo utilizadas en la evaluación así como el número de frames totales y evaluados

Se disponen de 15 vídeos para la evaluación, cabe destacar que no todos los vídeos se evalúan sobre todas las frames del vídeo, puesto que en algunos no se tenía el *ground truth* para todas las frames. Los vídeos Avss_S1, DaSide0811 seq6 y seq7, Pets_2009, Pets_S1 y Train56 se han seleccionado puesto que, el algoritmo Persona-Fondo del Estado del Arte se evalúa sobre estos. Además, BusStation, Corridor, Cubicle, Office, Park y Pedestrians han sido obtenidos de la *Change Detection*¹, página de evaluación de segmentadores con alto renombre. Siendo seleccionados aquellos en los que aparecen personas completas y por último, los vídeos del grupo VPULab. En las Figuras 4.2 y

¹(<http://www.changedetection.net/>)

4.3, se presentan los vídeos utilizados para la evaluación de este proyecto, mostrando un frame y su correspondiente *ground truth*.

Vídeo	Frame	Ground Truth
AVSS_S1		
BaileMulti		
BaileUni		
BusStation		
Corridor		
Cubicle		
DaSide0811-seq6		
DaSide0811-seq7		

Tabla 4.2: Vídeos Dataset del Proyectos (1) mostrando el frame y correspondiente *ground truth*

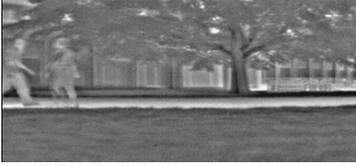
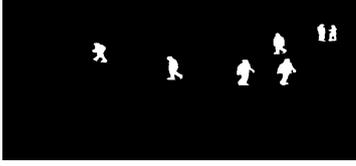
Vídeo	Frame	Ground Truth
Office		
Park		
Pedestrians		
Pets_2009		
Pets_S1		
Tenis1		
Train56		

Tabla 4.3: Vídeos Dataset del Proyectos (2) mostrando el frame y correspondiente *ground truth*

4.3. Métrica

Como ya se ha comentado en la introducción, las métricas utilizadas son objetivas basadas en la comparación de la máscara de salida obtenida con el *ground truth*. Además siguiendo el criterio de evaluación de [2], se ha definido el valor de fondo a

		Ground Truth		
		Positivos (ceros)	Negativos (unos)	
Algoritmo	Positivos (ceros)	TP	FP	TP+FP
	Negativos (unos)	FN	TN	FN+TN
		TP+FN	FP+TN	

Figura 4.1: Puntuaciones utilizadas en la evaluación con referencia al *ground truth*

ceros y las clasificaciones se consideran de la siguiente manera:

- Verdaderos Positivos (TP): Detección correcta de valor cero (fondo)
- Verdaderos Negativos (TN): Detección correcta de valor uno (frente).
- Falsos Positivos (FP): Detección incorrecta de valor cero.
- Falsos Negativos (FN): Detección incorrecta de valor uno.

En la Figura 4.1, se pueden apreciar a lo que corresponde cada puntuación en referencia al *ground truth*.

Además de estas cuatro medidas, se tienen otras 3 medidas más el producto de la combinación de las cuatro anteriores. Además de esto, para obtener la mejor configuración de los segmentadores se tiene otra medida que pondera por igual el fondo y el frente.

Precisión: Es el número total de píxeles correctos (fondo o frente) dependiendo lo que estemos midiendo, con respecto al número total de píxeles detectados por el algoritmo de este tipo (fondo o frente). Dado que buscamos el equilibrio en cuanto a la detección de fondo y frente, se realizan dos cálculos de la precisión correspondientes.

$$\text{Precisión de Fondo} = \frac{TP}{TP+FP}$$

$$\text{Precisión de Frente} = \frac{TN}{TN+FN}$$

En las tablas de resultados la precisión de fondo será representada por PF y la de frente por PFR.

Recall: Es el número total de píxeles correctos (fondo o frente) dependiendo lo que estemos midiendo, con respecto al total real (obtenido de *ground truth*) del tipo que estemos midiendo.

$$\text{Recall de Fondo} = \frac{TP}{TP+FN}$$

$$\text{Recall de Frente} = \frac{TN}{TN+FP}$$

En las tablas de resultados el recall de fondo será representado por RF y la de frente por RFR.

F1Score: Es la relación de las medidas anteriores, combinando la correcta detección del algoritmo con respecto a la Precisión y el Recall concediendo el mismo peso a las dos medidas.

$$\mathbf{F1Score\ de\ Fondo} = 2 \times \frac{PF \times RF}{PF + RF}$$

$$\mathbf{F1Score\ de\ Frente} = 2 \times \frac{PFR \times RFR}{PFR + RFR}$$

En las tablas de resultados el recall de fondo será representada por F1F y la de frente por F1FR.

La medida que se va a utilizar para escoger la mejor configuración de cada segmentador se ha tomado del proyecto final de carrera [74]. Se podrían haber aplicado muchas otras funciones pero se decidió tomar esta medida simple. Siendo la configuración escogida la que mayor valor tenga en esta medida.

$$\mathbf{Mejor\ Configuración} = \mathbf{F1Score\ (fondo)} + \mathbf{F1Score\ (frente)}$$

4.4. Resultados

Una vez se ha desarrollado el objetivo de este proyecto, el desarrollo de un algoritmo de segmentación Persona-Fondo, se ha de realizar una evaluación del mismo. La evaluación no sólo se realiza para obtener los resultados del mismo, sino que para observar la mejora generada por la nueva implementación. Por lo que, se tendrá una evaluación similar al desarrollo seguido para la implementación del segmentador. Por tanto, el capítulo estará estructurado en cinco subsecciones en las cuales se mostrarán los distintos resultados y aproximación obtenidos en este proyecto. Se mostrarán los resultados obtenidos para los segmentadores Persona-Fondo, sección 4.4.1. A continuación, se presentarán los resultados del segmentador Frente-Fondo, sección 2.3. Puesto que en este proyecto no se ha realizado solo una aproximación o implementación, se muestran los resultados de las mismas con su consiguiente explicación de los resultados. Los resultados de la primera aproximación se muestran en la sección 4.4.3, los de la segunda en la sección 4.4.5 y la tercera aproximación se mostrará en la sección 4.4.6.

4.4.1. Segmentación Persona-Fondo

Para la evaluación de los segmentadores Persona-Fondo, se ha procedido a evaluar las máscaras sin post-procesar, generadas por la implementación DEBP dado que, si

usásemos las máscaras post-procesadas se debería realizar un post-procesado a las máscaras de los segmentadores Frente-Fondo y no es el objetivo de este proyecto. Las máscaras corresponden a los Mapas de confianza para los 101 umbrales (de 0 a 1 en pasos de 0.01) utilizados en el algoritmo. Para esta evaluación se han utilizado los 15 vídeos disponibles en el *dataset* de evaluación. Por tanto, se presenta ahora una tabla con las puntuaciones obtenidas para cada uno de los vídeos.

	PF	PFR	RF	RFR	F1F	F1FR
AVSS_S1	0.999	0.088	0.674	0.998	0.803	0.159
BaileMulti	0.996	0.202	0.467	0.982	0.633	0.331
BaileUni	0.997	0.206	0.475	0.990	0.641	0.337
BusStation	0.998	0.045	0.399	0.970	0.567	0.408
Corridor	0.980	0.038	0.439	0.655	0.603	0.323
Cubicle	0.999	0.021	0.243	0.998	0.389	0.606
DaSide0811-seq6	0.999	0.223	0.451	0.999	0.622	0.363
DaSide0811-seq7	0.999	0.177	0.506	0.999	0.669	0.296
Office	0.999	0.076	0.386	0.998	0.556	0.425
Park	0.999	0.017	0.294	0.998	0.453	0.528
Pedestrians	0.999	0.013	0.375	0.999	0.539	0.448
Pets_2009	0.999	0.079	0.719	0.991	0.836	0.146
Pets_S1	0.999	0.076	0.695	0.968	0.819	0.136
Tennis1	0.997	0.168	0.535	0.984	0.695	0.287
Train56	0.996	0.172	0.525	0.979	0.687	0.292

Tabla 4.4: Puntuaciones Promedio Mapas de Confianza Segmentador Persona-Fondo

La tabla 4.4 muestra como ya se comentó en capítulos anteriores, que este algoritmo tiene como objetivo primordial la clasificación correcta de los píxeles de fondo. Por tanto, si observamos la tabla 4.4, apreciamos que en general la precisión de fondo es muy superior a la de frente como cabía esperar. Además, como ya se ha comentado los vídeos tienen gran influencia sobre los resultados. Si bien es cierto, como en este caso la evaluación se realiza para los 101 umbrales y lo que se muestra en la tabla es el valor medio de las puntuaciones, hay que destacar que para umbrales muy altos y muy bajos se generan peores puntuaciones que para los umbrales medios.

En la Figura 4.2 se muestran resultados visuales del algoritmo Persona-Fondo, no se busca definir la persona como tal, sino que se busca que ningún píxel de la persona sea clasificado erróneamente como fondo. Así mismo, se puede apreciar que la definición de las zonas de fondo es más precisa y por tanto nos da los buenos resultados que se han comentado en la tabla 4.4. Consiguiente de lo anterior las puntuaciones para el frente son malas, como ya se ha remarcado anteriormente.

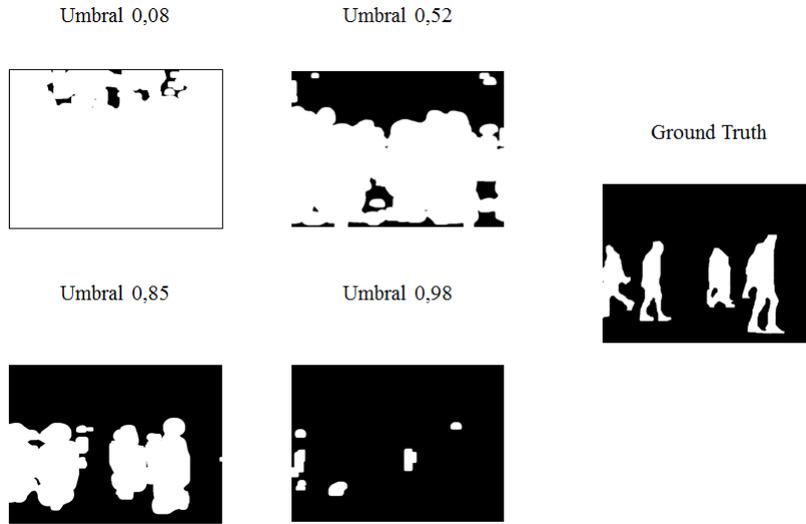


Figura 4.2: Resultados Segmentador Persona-Fondo para diferentes umbrales

Siguiendo la evaluación de [2], se modifica el factor de penalización para los falsos positivos, dado que este algoritmo prefiere clasificar píxeles de fondo como persona y por tanto se aplican estos factores para ver cómo afectan al algoritmo.

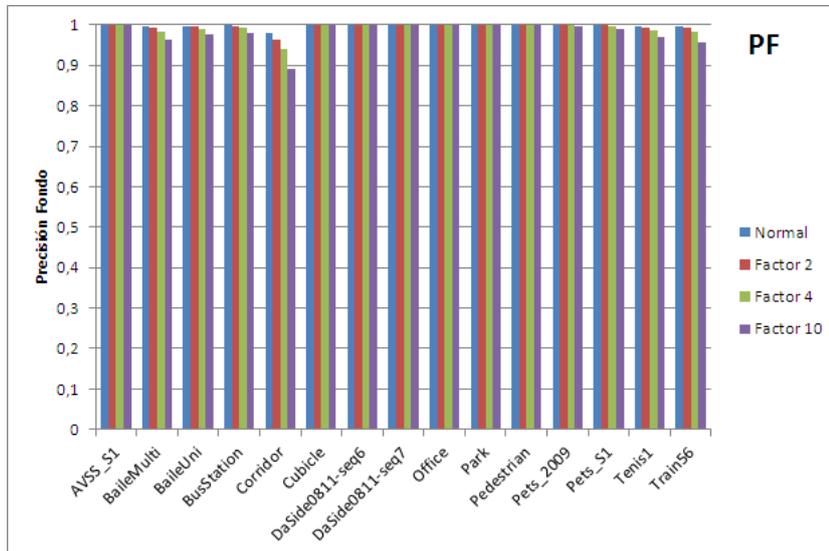


Figura 4.3: Gráfica PF con distintos factores de penalización

En las Figuras 4.3 4.4 se observa la repercusión de modificar el factor de penalización para los falsos positivos, siguiendo la evaluación usada por [2]. La puntuación PF y RFR son las que se ven afectadas por estos factores pero no en gran medida, también influyen los vídeos utilizados, por ejemplo el vídeo Corridor, sufre una gran

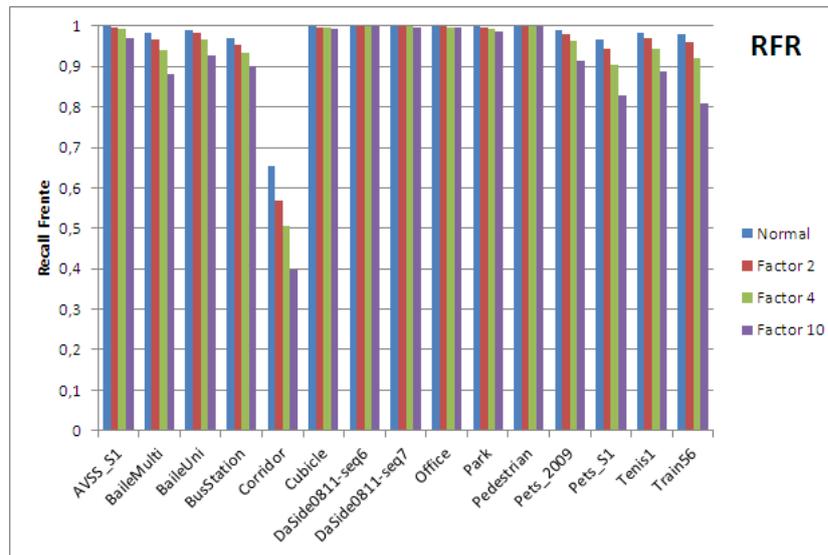


Figura 4.4: Gráfica RFR con distintos factores de penalización

penalización puesto que es un vídeo termal. Además de ser termal, las personas cuando aparecen en el vídeo hasta que no avanzan los frames no se ve por completo el cuerpo, por lo que el detector fallará hasta poder ver la persona entera, por lo que este vídeo es más complicado para los segmentadores sin factor de penalización claramente se verá muy afectada la puntuación para este vídeo.

4.4.2. Segmentación Frente-Fondo

Como ya se comentó en el capítulo anterior, se han usado 27 algoritmos sobre los 35 algoritmos disponibles puesto que, se tenía la posibilidad de modificar su configuración para obtener un mejor resultado en función del vídeo. Los resultados que se van a mostrar, son los obtenidos con la configuración óptima para cada uno de ellos. En el apéndice C, se muestra detalladamente el proceso seguido para un algoritmo y un vídeo.

Dado que en esta parte del proyecto disponemos de múltiples segmentadores y se utiliza el *dataset* completo, se presenta ahora una tabla promedio 4.5 de las puntuaciones obtenidas en las evaluaciones para cada segmentador y para una análisis más exhaustivo se debe acudir al apéndice D sección D.1, en el cual, se muestran seis tablas siendo una para cada puntuación de las obtenidas en la evaluación para cada segmentador y para cada vídeo.

	PF	PFR	RF	RFR	F1F	F1FR
AdaptiveBGLearning	0.970	0.418	0.960	0.607	0.965	0.559
DPAdaptiveMedianBGS	0.979	0.557	0.960	0.715	0.969	0.659
DPEigenbackgroundBGS	0.986	0.325	0.878	0.845	0.925	0.552
DPGrimsonGMMBGS	0.973	0.476	0.959	0.679	0.965	0.589
DPMeanBGS	0.968	0.495	0.966	0.602	0.966	0.566
DPPratiMediodBGS	0.974	0.475	0.957	0.621	0.964	0.577
DPWrenGABGS	0.974	0.503	0.967	0.684	0.970	0.634
DPZivkovicAGMMBGS	0.970	0.510	0.968	0.645	0.968	0.593
FuzzyChoquetIntegral	0.969	0.611	0.973	0.574	0.971	0.637
FuzzySugenoIntegral	0.970	0.605	0.977	0.591	0.974	0.632
GMG	0.977	0.467	0.965	0.700	0.971	0.637
KDE	0.985	0.325	0.901	0.771	0.939	0.550
LBAdaptiveSOM	0.979	0.528	0.957	0.714	0.967	0.660
LBFuzzyAdaptiveSOM	0.982	0.453	0.944	0.769	0.962	0.641
LBFuzzyGaussian	0.968	0.550	0.979	0.617	0.973	0.621
LBMixtureOfGaussians	0.966	0.532	0.974	0.589	0.970	0.586
LBSimpleGaussian	0.969	0.482	0.97	0.634	0.969	0.591
MOGV1BGS	0.959	0.634	0.985	0.498	0.972	0.547
MOGV2BGS	0.972	0.478	0.968	0.632	0.969	0.611
MultiLayerBGS	0.974	0.691	0.981	0.640	0.977	0.680
SGMM-SOD	0.990	0.592	0.946	0.850	0.966	0.730
SigmaDeltaBGS	0.962	0.552	0.984	0.509	0.973	0.558
T2FGMM_UM	0.959	0.503	0.974	0.387	0.966	0.435
T2FGMM_UV	0.978	0.444	0.953	0.728	0.965	0.615
T2FMRF_UM	0.954	0.501	0.978	0.322	0.965	0.479
T2FMRF_UV	0.970	0.499	0.962	0.575	0.965	0.578
VuMeter	0.966	0.624	0.971	0.561	0.968	0.599

Tabla 4.5: Puntuaciones Promedio Segmentadores Frente-Fondo

En la tabla 4.5 se observa que ninguno de los segmentadores de Frente-Fondo alcanza el máximo de la puntuación posible, si bien es cierto, muchos de ellos tras su exhaustiva configuración obtienen una alta puntuación. Hay que reseñar que en cuanto al fondo se obtienen muy buenas puntuaciones, pero en el frente son buenas pero no llegan al nivel del fondo lo que posteriormente se verá reflejado en las puntuaciones de F1 Score que son el resumen. En este caso, que los segmentadores no tengan muy buena puntuación es debido en cierta manera a la problemática desarrollada en el capítulo 2, sección 2.3.2. Los resultados de la tabla 4.5 en rojo, son los segmentadores que al observar su resultado visualmente y a través del apéndice D, sus puntuaciones son los que mejores resultados nos dan. Se muestran ahora unos ejemplos visuales de dos segmentadores que generan un muy buen resultado y a su vez, dos segmentadores que generan no muy buenos resultados.



Figura 4.5: Resultados Segmentadores (1) (Vídeo Pedestrians frame 345).

En la Figura 4.5, se obtiene la persona que en ese momento aparece el vídeo, en ambos casos se obtiene la persona definida, en el segmentador SGMM-SOD se obtiene la persona un poco más grande que en el *ground truth*, al contrario que en el segmentador DPAdaptiveMedianBGS que tiene píxeles sin clasificación como persona.

Como se puede apreciar en la Figura 4.6, la máscara de salida generada por los segmentadores no es para nada buena, mientras que hay que destacar que la figura de la persona que aparece en el *ground truth* si se corresponde a la obtenida por el segmentador. Solo con ver las imágenes que se generan, se aprecia la deficiencia de estos segmentadores ante las problemáticas ya antes referenciadas. Un claro ejemplo

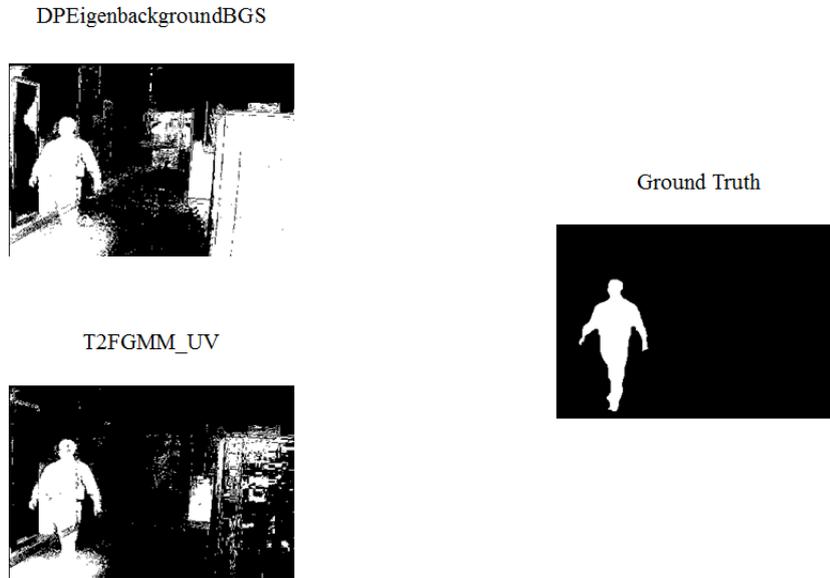


Figura 4.6: Resultados Segmentadores (2) (VÍdeo Cubicle frame 2500).

de una de las problemáticas que se observa, son las sombras debajo de la persona, otorgando puntuaciones malas en la evaluación correspondiente de este.

Para concluir en esta sección, se destaca la diferencia entre segmentadores dependiendo de la metodología que utilizan. Así mismo, si no se hubieran configurado los segmentadores a su punto óptimo se hubieran obtenido deficiencias tales como, clasificar como frente sólo el borde de las personas mientras que el interior se clasificaría como fondo. Otro problema que se puede encontrar, es la lentitud en el aprendizaje del modelo de fondo quedándose fantasmas en la escena y generando peores puntuaciones. También se ha de reseñar que en el resultado influye en gran medida el vídeo a procesar, dado que dependiendo de las sombras que se generen en el escenario, las oclusiones, el tamaño de los objetos de frente son puntos de desventaja para el segmentador.

4.4.3. Primera Aproximación

Tras el estudio de ambas segmentaciones y la visualización de los resultados obtenidos, se propuso una primera combinación de las informaciones de ambas. En este caso se quiso aplicar la información del tiempo que nos dan los segmentadores Frente-Fondo, sobre la información obtenida por el segmentador Persona-Fondo. Por lo que, se decidió generar vídeos de los mapas de confianza, los cuales posteriormente sería pasados por los segmentadores Frente-Fondo.

Los primero que se ha de destacar, es que de los 32 algoritmos con los que se

contaban, en este caso se han reducido en número dado que, no todos los segmentadores generaban un modelo de fondo, ya que muchos de ellos utilizan metodologías en las cuales no necesitan generar la imagen del modelo de fondo, como pueden ser aquellos que utilizan *codebooks*, análisis por capas y otros modelos que no generan imagen de fondo. Además de los segmentadores que no se han podido utilizar por el motivo anterior, se han tenido que descartar otros diversos dado que los resultados obtenidos eran peores a los generados por el segmentador Persona-Fondo por sí solo.

Una vez realizada esta evaluación, nos dimos cuenta que la mejora obtenida era ínfima siendo la modificación en las puntuaciones en el tercer o cuarto decimal en algunos casos. Por lo que, no se realizó la evaluación completa para todos los vídeos del *dataset*, ni para todos los segmentadores. Además de las tablas con resultados se presentan gráficas de la curva ROC, en la cual se representan la precisión y recall de las puntuaciones obtenidas.

En la tabla 4.6 se presentan los resultados obtenidos de la primera aproximación para el vídeo AVSS_S1, teniendo claro que los resultados presentados se han de enfrentar a los presentados en la tabla 4.4, dado que son las puntuaciones de los Mapas de confianza sin aplicar las técnicas de modelado de fondo de los algoritmos de Frente-Fondo para modelar la confianza del segmentador Persona-Fondo.

	PF	PFR	RF	RFR	F1F	F1FR
AdaptiveBGLearning	0.999	0.090	0.681	0.998	0.808	0.162
GrimsonGMMBGS	0.999	0.091	0.684	0.999	0.811	0.164
ZivkovicAGMMBGS	0.999	0.091	0.684	0.999	0.811	0.164
MeanBGS	0.999	0.091	0.681	0.999	0.808	0.163
WrenGABGS	0.999	0.090	0.679	0.999	0.807	0.162
FuzzySugenoIntegral	0.999	0.089	0.676	0.998	0.805	0.160
FuzzyChoquetIntegral	0.999	0.090	0.677	0.998	0.806	0.161
SimpleGaussian	0.999	0.091	0.685	0.999	0.811	0.164
FuzzyGaussian	0.999	0.092	0.686	0.998	0.812	0.164
MixtureOfGaussians	0.999	0.091	0.684	0.998	0.810	0.163

Tabla 4.6: Puntuación Mapas de Confianza Vídeo AVSS_S1 segmentados

Observando los resultados obtenidos, se aprecia una pequeña mejora para todos los casos, siendo las puntuaciones muy similares al segmentador Persona-Fondo con muy buena puntuación para fondo pero muy mala para frente. Pese a esto, no se genera una mejora en las puntuaciones de frente que merezca reseñarse, dado que las puntuaciones de fondo continúan en la línea de las anteriores. Se presentan ahora dos gráficas de la precisión y recall de los segmentadores utilizados.

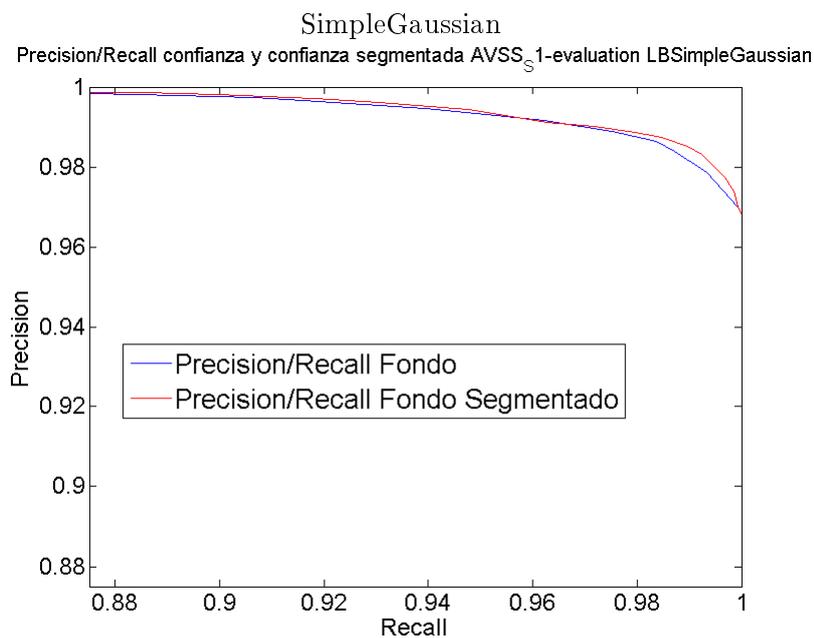


Figura 4.7: Gráficas Precisión/Recall AVSS_S1 para los métodos SimpleGaussian

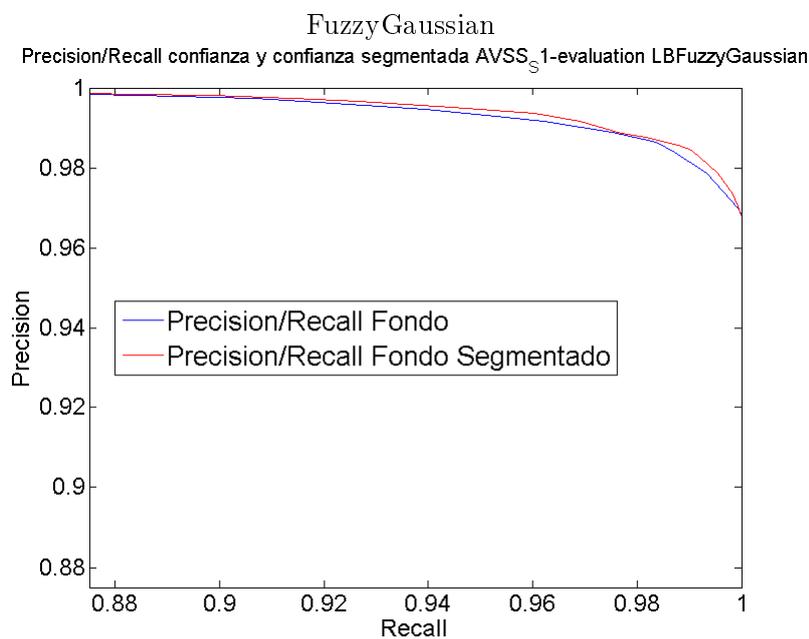


Figura 4.8: Gráficas Precisión/Recall AVSS_S1 para los métodos FuzzyGaussian

Se puede apreciar en las Figuras 4.7 y 4.8 que en las gráficas, ambas curvas se superponen exceptuando en algunas zonas y otras la diferencia es ínfima pero la hay.

Como ya se ha comentado antes la mejora no obtiene grandes resultados.

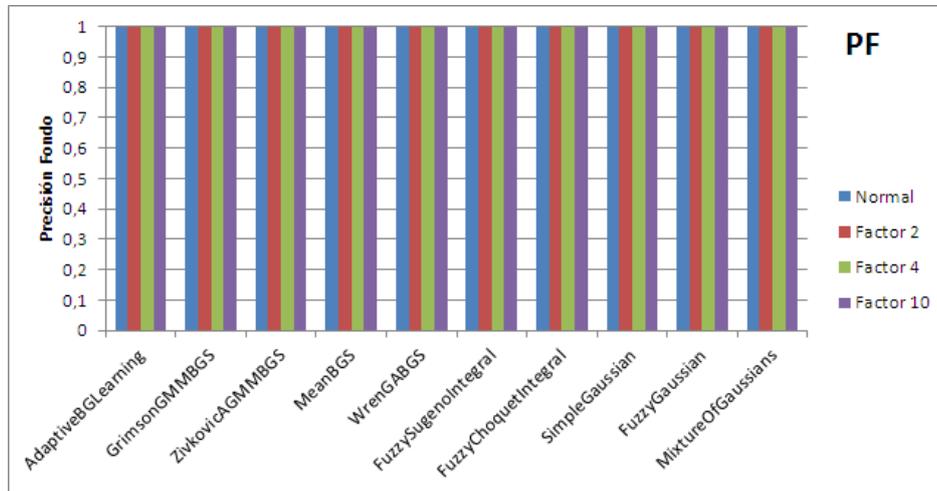


Figura 4.9: Gráfica PF con distintos factores de penalización primera aproximación

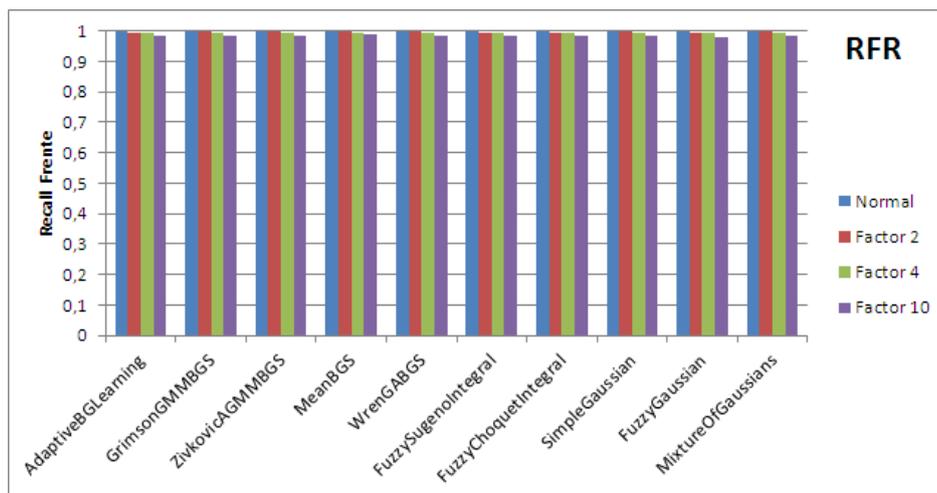


Figura 4.10: Gráfica RFR con distintos factores de penalización primera aproximación

En las Figuras 4.9 4.10, se puede apreciar el efecto de aplicar los factores de penalización usados en la evaluación de [2], dado que esta aproximación se compara con las evaluaciones obtenidas para el segmentador Persona-Fondo. Claramente a medida que aumentamos el factor de penalización se obtienen peores resultados, en este caso solo se muestra esta gráfica puesto que el resto de puntuaciones no sufre casi variación en sus valores. También cabe destacar, que el segmentador Persona-Fondo genera muchas falsas detecciones intermitentes, por lo que usar segmentadores Frente-Fondo con modelado multimodal, podría generar mejores resultados.

En conclusión, no se continuaron realizando más pruebas dado que las mejoras que se obtenían eran prácticamente nulas y en otros muchos casos no se obtenía mejora alguna. Por tanto esta aproximación no ofrece ninguna mejora ni buen resultado en comparación al segmentador Persona-Fondo utilizado, en este caso, no se quiso desechar esta aproximación puesto que tras su desarrollo es interesante dejar constancia para posibles trabajos futuros de los resultados de esta.

4.4.4. Segunda y Tercera Aproximación

De manera breve y previa a la exposición y razonamiento de las aproximaciones realizadas, vamos a razonar la idea de las mismas. En sí, la idea se extrae de los resultados obtenidos en la primera aproximación 4.4.3, y es que si los segmentadores Frente-Fondo generan máscaras en las que el objetivo es la definición de los objetos de frente de manera precisa. Por el contrario, el segmentador Persona-Fondo busca la clasificación correcta de las zonas de fondo, surge la idea de combinar ambas máscaras binarias obtenidas de los segmentadores.

Se pueden dar miles de combinaciones posibles, pero en este caso se pensó dos posibles combinaciones que serían los límites de las aproximaciones, siendo estas la realización de la operación AND entre las dos máscaras y la operación OR. En el caso de la primera, lo que buscamos es quedarnos de ambas máscaras las zonas donde coincidan clasificaciones de píxeles como frente, sin embargo, para la otra operación los que buscamos es quedarnos con las clasificaciones de ambas máscaras como frente. Por tanto, la operación OR es una analogía con el segmentador Persona-Fondo, puesto que garantiza que no dejamos nada de frente sin clasificar, mientras que la operación AND busca recuperar y clasificar los más fielmente posible el frente. A parte de ser los límites, estas combinaciones se utilizan para tener una referencia de que resultado “óptimo” para el frente en el caso de la AND y de fondo para el caso de la OR se puede llegar con la información que disponemos.

Se van a presentar ahora las tablas promedio de ambas operaciones, dado que las mismas han sido realizadas para todos los vídeos del *dataset* y para los 27 segmentadores disponibles.

En la tabla 4.7 se pueden apreciar los resultados obtenidos en la operación AND de ambas máscaras, se obtiene buena puntuación para el fondo en todos los segmentadores. Esto es debido al hecho de que, con esta operación al solo quedarnos con las zonas clasificadas como frente en ambas y que coincidan, por tanto, errores generados en los segmentadores de Frente-Fondo y en el segmentador Persona-Fondo son eliminados por la combinación de máscaras. En relación al frente, dado que esta operación busca clasificar de manera muy fiel el frente, se obtienen buenas puntuaciones.

	PF	PFR	RF	RFR	F1F	F1FR
AdaptiveBGLearning	0.970	0.470	0.970	0.595	0.970	0.586
DPAdaptiveMedianBGS	0.978	0.602	0.971	0.693	0.974	0.681
DPEigenbackgroundBGS	0.985	0.396	0.918	0.814	0.949	0.622
DPGrimsonGMMBGS	0.972	0.513	0.969	0.650	0.970	0.614
DPMeanBGS	0.967	0.547	0.975	0.575	0.971	0.590
DPPratiMediodBGS	0.973	0.519	0.966	0.608	0.969	0.603
DPWrenGABGS	0.974	0.558	0.976	0.659	0.974	0.658
DPZivkovicAGMMBGS	0.970	0.541	0.975	0.617	0.972	0.608
FuzzyChoquetIntegral	0.969	0.667	0.981	0.551	0.975	0.665
FuzzySugenoIntegral	0.970	0.665	0.984	0.568	0.977	0.657
GMG	0.977	0.505	0.972	0.688	0.974	0.653
KDE	0.985	0.438	0.941	0.759	0.961	0.640
LBAdaptiveSOM	0.979	0.593	0.969	0.699	0.973	0.688
LBFuzzyAdaptiveSOM	0.981	0.515	0.960	0.745	0.970	0.680
LBFuzzyGaussian	0.967	0.591	0.984	0.597	0.975	0.636
LBMixtureOfGaussians	0.966	0.562	0.980	0.561	0.972	0.597
LBSimpleGaussian	0.968	0.527	0.978	0.610	0.973	0.609
MOGV1BGS	0.959	0.669	0.989	0.482	0.974	0.555
MOGV2BGS	0.972	0.525	0.978	0.622	0.974	0.641
MultiLayerBGS	0.973	0.721	0.984	0.631	0.979	0.687
SGMM-SOD	0.989	0.651	0.961	0.833	0.974	0.766
SigmaDeltaBGS	0.961	0.590	0.988	0.500	0.974	0.567
T2FGMM_UM	0.959	0.547	0.983	0.373	0.971	0.453
T2FGMM_UV	0.978	0.499	0.966	0.698	0.971	0.647
T2FMRF_UM	0.954	0.575	0.986	0.316	0.970	0.502
T2FMRF_UV	0.971	0.559	0.974	0.567	0.972	0.607
VuMeter	0.967	0.658	0.978	0.552	0.972	0.616

Tabla 4.7: Puntuación Limite AND

En la tabla 4.8, se pueden observar los resultados de la evaluación de la operación OR que se asemejan a los obtenidos en el segmentador Persona-Fondo, dado que esta operación busca cubrir por completo a las personas por lo que se acierta mucho en la clasificación de fondo, pero la de frente falla puesto que hay píxeles en los que se falla su clasificación.

	PF	PFR	RF	RFR	F1F	F1FR
AdaptiveBGLearning	0.999	0.110	0.492	0.976	0.647	0.340
AdaptiveMedianBGS	0.999	0.110	0.491	0.987	0.646	0.336
GrimsonGMMBGS	0.999	0.104	0.462	0.996	0.617	0.328
EigenbackgroundBGS	0.999	0.111	0.492	0.994	0.646	0.338
MeanBGS	0.999	0.110	0.493	0.991	0.647	0.337
PratiMediodBGS	0.999	0.110	0.492	0.978	0.647	0.339
WrenGABGS	0.999	0.110	0.493	0.989	0.647	0.337
ZivkovicAGMMBGS	0.999	0.111	0.495	0.993	0.648	0.338
FuzzyChoquetIntegral	0.999	0.111	0.494	0.987	0.647	0.337
FuzzySugenoIntegral	0.999	0.111	0.495	0.987	0.648	0.338
GMG	0.999	0.111	0.495	0.977	0.649	0.344
KDE	0.999	0.103	0.462	0.976	0.620	0.332
AdaptiveSOM	0.999	0.109	0.490	0.980	0.645	0.340
FuzzyAdaptiveSOM	0.999	0.108	0.485	0.988	0.640	0.334
FuzzyGaussian	0.999	0.111	0.497	0.984	0.650	0.339
MixtureOfGaussians	0.999	0.111	0.496	0.992	0.649	0.338
SimpleGaussian	0.999	0.111	0.494	0.989	0.647	0.338
MOGV1BGS	0.998	0.111	0.498	0.980	0.651	0.338
MOGV2BGS	0.998	0.110	0.492	0.974	0.646	0.341
MultiLayerBGS	0.998	0.111	0.498	0.973	0.652	0.346
SGMM-SOD	0.999	0.109	0.487	0.982	0.642	0.341
SigmaDeltaBGS	0.998	0.111	0.498	0.974	0.652	0.344
T2FGMM_UM	0.999	0.110	0.493	0.978	0.647	0.336
T2FGMM_UV	0.999	0.110	0.489	0.994	0.643	0.336
T2FMRF_UM	0.998	0.109	0.493	0.970	0.648	0.340
T2FMRF_UV	0.998	0.109	0.490	0.972	0.644	0.339
VuMeter	0.997	0.111	0.495	0.974	0.649	0.344

Tabla 4.8: Puntuación Limite OR

4.4.5. Segunda Aproximación

Dado que el objetivo del proyecto es el desarrollo de un segmentador Persona-Fondo usando la información de segmentación Frente-Fondo, en esta aproximación se han combinado las máscaras obtenidas de los segmentadores Persona-Fondo y Frente-Fondo, aplicando un filtrado por reconstrucción ya explicado en el capítulo 3.

En la evaluación se presenta la tabla 4.9, la cual contiene las medidas promedio de las puntuaciones para todos los vídeos y sobre cada segmentador, teniendo seis tablas con mayor detalle en el apéndice D sección D.2dándonos cada una de las tablas la información de una puntuación de cada vídeo para la segunda aproximación.

Además de observar estas tablas, nos debemos fijar en la tabla 4.4 en la sección 4.4.1, dado que es la base de nuestro segmentador Persona-Fondo.

Como se puede apreciar en la tabla 4.9 y para más detalle en las secciones D.2 del apéndice D, se consigue una mejora significativa en las puntuaciones obtenidas tanto para fondo como para frente. Sin duda alguna, la mejora más significativa se produce en referencia al frente, dado que si nos fijamos en la tabla 4.4 de la sección 4.4.1, las puntuaciones de frente sufren una mejora de mayor grado que las de fondo. Esto se debe a que, la información las máscaras de los segmentadores Frente-Fondo que están definidas para definir con mucha precisión la máscara de frente, ayudan a eliminar las clasificaciones erróneas del segmentador Persona-Fondo puesto que las mismas han sido generadas por errores en los niveles de confianza dados a dichos píxeles. Además de las puntuaciones, tenemos el porcentaje de mejora con respecto al algoritmo Persona-Fondo original, calculado sobre los datos de la tabla 4.4, que como cabría esperarse en la mayor parte de las puntuaciones mejora, pero en algunas otras tiene una ínfima pérdida. En muchos casos el porcentaje es muy elevado, pero realmente no es tanto puesto que el valor del algoritmo original es muy bajo por tanto, cuanto se tiene una mejora destacable el porcentaje de mejora se dispara.

	PF(% Δ)	PFR(% Δ)	RF(% Δ)	RFR(% Δ)	F1F(% Δ)	F1FR(% Δ)
AdaptiveBGLearning	0.998(0.1)	0.154(44.3)	0.570(19)	0.962(-0.5)	0.704(11)	0.388(14.5)
DPAdaptiveMedianBGS	0.998(0.1)	0.202(89.3)	0.621(29.7)	0.964(-0.3)	0.744(17.3)	0.394(16.2)
DPEigenbackgroundBGS	0.997(0)	0.127(19)	0.531(10.9)	0.964(-0.3)	0.676(6.6)	0.361(6.5)
DPGrimsonGMMBGS	0.998(0.1)	0.148(39.7)	0.572(19.4)	0.964(-0.3)	0.709(11.8)	0.376(10.9)
DPMeanBGS	0.998(0.1)	0.185(73.3)	0.596(24.5)	0.964(-0.3)	0.724(14.2)	0.391(15.3)
DPPratiMediodBGS	0.996(-0.1)	0.233(118.3)	0.643(34.3)	0.935(-3.3)	0.759(19.7)	0.422(24.5)
DPWrenGABGS	0.998(0.1)	0.173(62.1)	0.597(24.7)	0.964(-0.3)	0.727(14.6)	0.381(12.4)
DPZivkovicAGMMBGS	0.998(0.1)	0.171(60.2)	0.600(25.3)	0.964(-0.3)	0.730(15.1)	0.389(14.7)
FuzzyChoquetIntegral	0.993(-0.4)	0.301(182)	0.673(40.5)	0.874(-9.6)	0.777(22.5)	0.466(37.5)
FuzzySugenoIntegral	0.996(-0.1)	0.253(137)	0.652(36.2)	0.921(-4.8)	0.765(20.6)	0.434(28)
GMG	0.998(0.1)	0.166(55.5)	0.602(25.7)	0.958(-1)	0.731(15.3)	0.411(21.2)
KDE	0.998(0.1)	0.133(24.6)	0.547(14.2)	0.959(-0.8)	0.689(8.7)	0.369(8.8)
LBAdaptiveSOM	0.998(0.1)	0.210(96.8)	0.636(32.8)	0.958(-1)	0.757(19.4)	0.400(18)
LBFuzzyAdaptiveSOM	0.998(0.1)	0.154(44.3)	0.572(19.4)	0.964(-0.3)	0.708(11.6)	0.372(9.7)
LBFuzzyGaussian	0.998(0.1)	0.174(63)	0.600(25.3)	0.964(-0.3)	0.728(14.8)	0.397(17.1)
LBMixtureOfGaussians	0.998(0.1)	0.176(64.9)	0.601(25.5)	0.964(-0.3)	0.729(15)	0.400(18)
LBSimpleGaussian	0.998(0.1)	0.156(46.2)	0.577(20.5)	0.964(-0.3)	0.711(12.1)	0.386(13.9)
MOGV1BGS	0.998(0.1)	0.242(126.7)	0.657(37.2)	0.961(-0.6)	0.771(21.6)	0.414(22.1)
MOGV2BGS	0.997(0)	0.126(18.1)	0.534(11.5)	0.964(-0.3)	0.678(6.9)	0.369(8.8)
MultiLayerBGS	0.997(0)	0.310(190.4)	0.701(46.4)	0.935(-3.3)	0.803(26.6)	0.448(32.2)
SGMM-SOD	0.998(0.1)	0.252(136.1)	0.661(38)	0.953(-1.5)	0.774(22.1)	0.421(24.2)
SigmaDeltaBGS	0.998(0.1)	0.183(71.5)	0.606(26.5)	0.957(-1.1)	0.732(15.4)	0.416(22.7)
T2FGMM_UM	0.997(0)	0.233(118.3)	0.631(31.8)	0.934(-3.4)	0.749(18.1)	0.417(23)
T2FGMM_UV	0.998(0.1)	0.137(28.4)	0.554(15.7)	0.964(-0.3)	0.694(9.4)	0.369(8.8)
T2FMRF_UM	0.993(-0.4)	0.305(185.8)	0.653(36.4)	0.855(-11.6)	0.762(20.2)	0.469(38.3)
T2FMRF_UV	0.998(0.1)	0.185(73.3)	0.594(24)	0.963(-0.4)	0.722(13.9)	0.396(16.8)
VuMeter	0.994(-0.3)	0.304(184.8)	0.691(44.3)	0.908(-6.1)	0.793(25.1)	0.458(35.1)

Tabla 4.9: Puntuación Reconstrucción Mejora Segmentador Persona-Fondo mostrando el porcentaje de incremento con respecto a la tabla 4.4.

En cuanto al fondo, también se mejora puesto que al eliminar los errores de las máscaras del segmentador Persona-Fondo dichas zonas clasificadas erróneamente como frente pasan a ser fondo y por tanto se obtiene una mejora en las puntuaciones del mismo. Las pérdidas que se aprecian en el recall de frente vienen dadas, por el hecho de que al mejorar la precisión de frente se ve reducido de manera ínfima el valor del recall. Si ya nos fijamos en las tablas de la sección D.2 en el apéndice D más detenidamente, se observa que las mejoras tienen mucha dependencia de los vídeos del *dataset* y del segmentador utilizado. Por tanto, si solo observamos el promedio de dichas tablas, observamos una mejora significativa del algoritmo de segmentación Persona-Fondo.

Con las tablas quedan claros los datos, pero sin duda observar la mejora de los resultados nos ayudará a apreciar con más detalle la mejora obtenida.

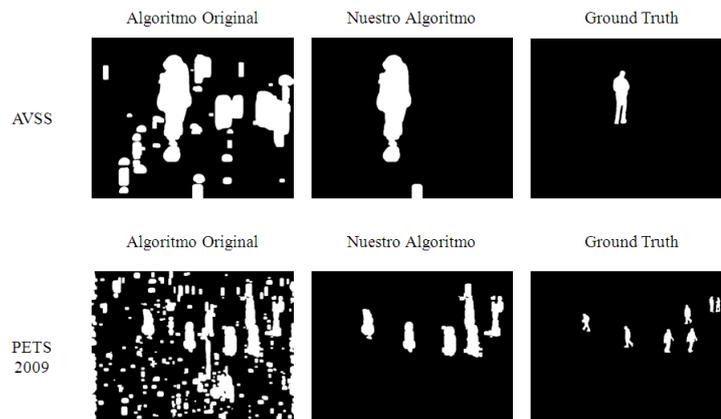


Figura 4.11: Imágenes (1) Algoritmo Persona-Fondo Desarrollado

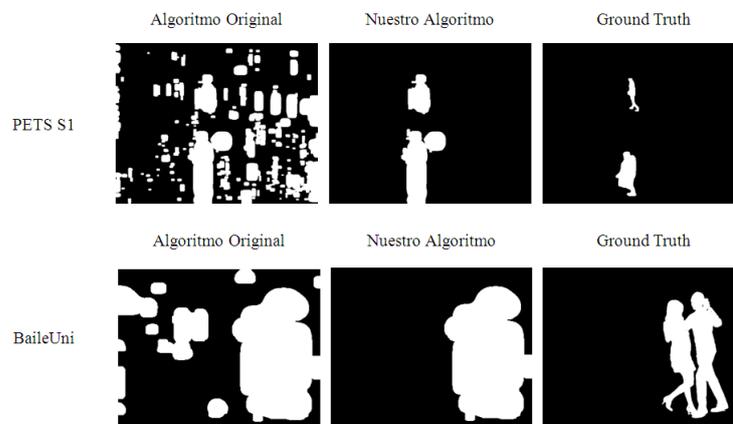


Figura 4.12: Imágenes (2) Algoritmo Persona-Fondo Desarrollado

A través de las Figuras y 4.12, se puede apreciar de manera significativa la mejora obtenida a través del algoritmo desarrollado. Se puede apreciar que, se eliminan prácticamente todas las zonas de errónea clasificación como frente por el algoritmo original. Sí que es cierto, que hay ciertas zonas que no son eliminadas teniendo su motivo en la calidad de la segmentación Frente-Fondo, ya que es la información introducida para mejorar el algoritmo.

Una vez que tenemos ambas evaluaciones, tanto la visual como la de puntuaciones nos comparamos con las tablas 4.7 y 4.8 de la sección 4.4.4 que son los límites que hemos establecido para frente y para fondo respectivamente, y estamos dentro de los mismos par todos los segmentadores en algunos casos se mejora y en otros se empeora, por lo que los límites que se establecieron guardan relación con las puntuaciones obtenidas en esta evaluación.

En esta aproximación buscamos mejorar la segmentación Persona-Fondo, por lo que si se sigue la evaluación de [2], se van a aplicar ahora los factores de penalización para la puntuación falsos positivos como en la evaluación del algoritmo original. En este caso, se van a mostrar gráficas de las puntuaciones a las cuales está modificación afecta en mayor medida.

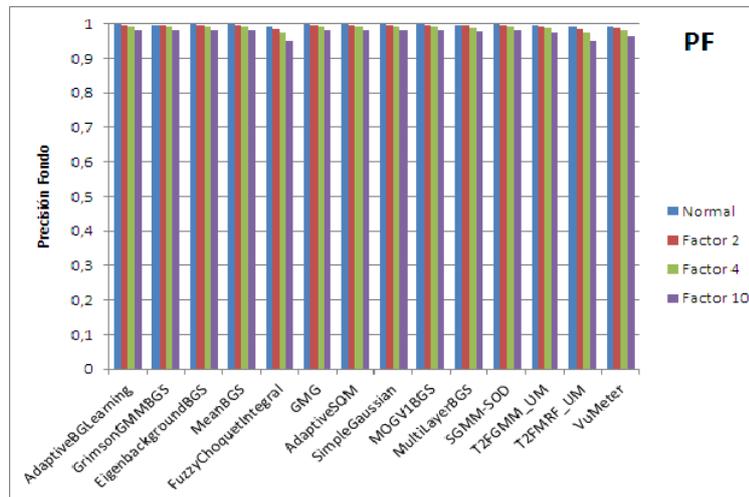


Figura 4.13: Gráfica puntuación PF de la segunda aproximación para distintos factores de penalización

En las Figuras 4.13 y 4.14 se pueden apreciar el efecto que provocan los factores de penalización, el cual cumple con lo que cabría esperar, puesto que el algoritmo original penaliza más los falsos positivos que otros tipos de segmentadores, por tanto, a medida que se aumenta el factor el valor de la puntuación es menor, pero esta pérdida es ínfima por lo que se mantiene estable. No se muestran en las gráficas el

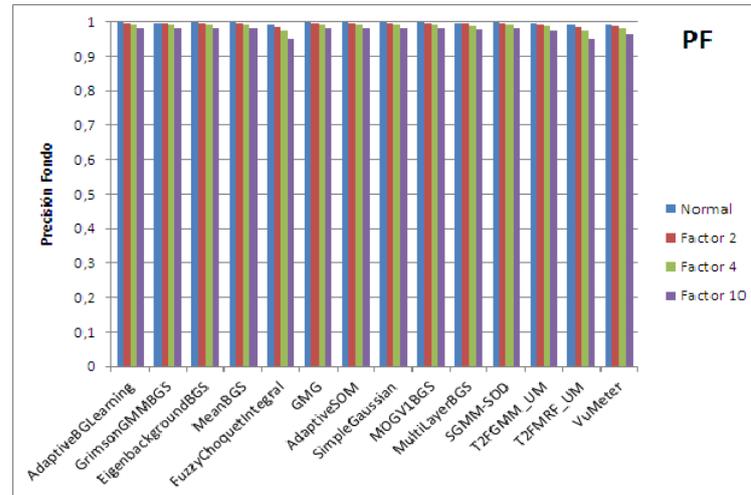


Figura 4.14: Gráfica puntuación RFR de la segunda aproximación para distintos factores de penalización

resultado de todos los segmentadores, puesto que por problemas de visibilidad se han reducido el número de segmentadores a mostrar.

Para concluir, las mejoras obtenidas en esta aproximación son claras tanto visualmente como por la evaluación realizada. Además de esta forma de unificar las informaciones disponibles, existen múltiples más las cuales se podrían haber desarrollado.

4.4.6. Tercera Aproximación

Como ya se comentó anteriormente, tras los resultados obtenidos en la aproximación anterior siendo esta el objetivo de este proyecto, se pensó en realizar la misma aproximación que en la anterior pero buscando mejorar la segmentación Frente-Fondo. Al igual que en el caso anterior, se ha realizado la evaluación sobre todo el *dataset* del proyecto además de todos los segmentadores disponibles. Por tanto, también se dispone en la sección D.3 en el apéndice D, de seis tablas pertenecientes a las seis puntuaciones obtenidas en la evaluación. Se presentan ahora los resultados promedio obtenidos en la evaluación de esta aproximación.

Si se observan los resultados obtenidos en la evaluación en la tabla 4.10, se aprecian buenas puntuaciones para frente como para fondo teniendo en mente que este tipo de algoritmos buscan una clasificación lo más ajustada posible a los objetos de frente. Además, nos hemos de fijar en los resultados obtenidos para los segmentadores Frente-Fondo originales en la tabla 4.5 en la sección 4.4.2 de este capítulo. Se obtienen en este caso mejoras no tan significativas como en la aproximación anterior tanto para frente

como para fondo. Cabe destacar, que ya de por sí los algoritmos Frente-Fondo al haber sido configurados a su punto óptimo, ya generaban buenos resultados por si solos por lo que su mejora es en eliminación de los ruidos generados por los segmentadores y otras posibles problemáticas ya anteriormente nombradas.

	PF(% Δ)	PFR(% Δ)	RF(% Δ)	RFR(% Δ)	F1F(% Δ)	F1FR(% Δ)
AdaptiveBGLearning	0.970(0)	0.444(6.2)	0.963(0.3)	0.604(-0.5)	0.966(0.1)	0.572(2.3)
DPAdaptiveMedianBGS	0.979(0)	0.575(3.2)	0.962(0.2)	0.707(-1.1)	0.970(0.1)	0.671(1.8)
DPEigenbackgroundBGS	0.973(-1.3)	0.497(52.9)	0.963(9.7)	0.671(-20.6)	0.967(4.5)	0.606(9.8)
DPGrimsonGMMBGS	0.986(1.3)	0.349(26.7)	0.888(-7.4)	0.836(23.1)	0.931(3.5)	0.580(1.5)
DPMeanBGS	0.968(0)	0.517(4.4)	0.969(0.3)	0.592(-1.7)	0.968(0.2)	0.580(2.5)
DPPratiMediodBGS	0.975(0.1)	0.493(3.8)	0.958(0.1)	0.626(0.8)	0.965(0.1)	0.594(2.9)
DPWrenGABGS	0.974(0)	0.532(5.8)	0.971(0.4)	0.674(-1.5)	0.972(0.2)	0.652(2.8)
DPZivkovicAGMMBGS	0.970(0)	0.527(3.3)	0.970(0.2)	0.636(-1.4)	0.970(0.2)	0.603(1.7)
FuzzyChoquetIntegral	0.969(0)	0.646(5.7)	0.976(0.3)	0.565(-1.6)	0.972(0.1)	0.659(3.5)
FuzzySugenoIntegral	0.970(0)	0.646(6.8)	0.981(0.4)	0.581(-1.7)	0.976(0.2)	0.656(3.8)
GMG	0.977(0)	0.483(3.4)	0.966(0.1)	0.698(-0.3)	0.971(0)	0.644(1.1)
KDE	0.985(0)	0.373(14.8)	0.917(1.8)	0.769(-0.3)	0.948(1)	0.590(7.3)
LBAdaptiveSOM	0.979(0)	0.557(5.5)	0.959(0.2)	0.710(-0.6)	0.969(0.2)	0.672(1.8)
LBFuzzy AdaptiveSOM	0.982(0)	0.476(5.1)	0.947(0.3)	0.762(-0.9)	0.963(0.1)	0.659(2.8)
LBFuzzyGaussian	0.968(0)	0.575(4.5)	0.981(0.2)	0.610(-1.1)	0.974(0.1)	0.635(2.3)
LBMixtureOfGaussians	0.966(0)	0.545(2.4)	0.976(0.2)	0.579(-1.7)	0.971(0.1)	0.595(1.5)
LBSimpleGaussian	0.969(0)	0.506(5)	0.974(0.4)	0.625(-1.4)	0.970(0.1)	0.604(2.2)
MOGV1BGS	0.959(0)	0.655(3.3)	0.987(0.2)	0.490(-0.6)	0.973(0.1)	0.556(1.6)
MOGV2BGS	0.972(0)	0.505(5.6)	0.973(0.5)	0.630(-0.3)	0.972(0.3)	0.632(3.4)
MultiLayerBGS	0.974(0)	0.706(2.2)	0.981(0)	0.639(-0.2)	0.977(0)	0.684(0.6)
SGMM-SOD	0.990(0)	0.622(5.1)	0.948(0.2)	0.847(-0.4)	0.967(0.1)	0.749(2.6)
SigmaDeltaBGS	0.962(0)	0.573(3.8)	0.986(0.2)	0.507(-0.4)	0.973(0)	0.566(1.4)
T2FGMM_UM	0.960(0.1)	0.540(7.4)	0.980(0.6)	0.382(-1.3)	0.969(0.3)	0.452(3.9)
T2FGMM_UV	0.978(0)	0.476(7.2)	0.957(0.4)	0.719(-1.2)	0.967(0.2)	0.635(3.3)
T2FMRF_UM	0.954(0)	0.570(13.8)	0.985(0.7)	0.319(-0.9)	0.969(0.4)	0.502(4.8)
T2FMRF_UV	0.971(0.1)	0.535(7.2)	0.970(0.8)	0.572(-0.5)	0.970(0.5)	0.599(3.6)
VuMeter	0.966(0)	0.637(2.1)	0.973(0.2)	0.559(-0.4)	0.968(0)	0.607(1.3)

Tabla 4.10: Puntuación Reconstrucción Mejora Segmentador Frente-Fondo mostrando el porcentaje de incremento con respecto a la tabla 4.5.

También se puede apreciar en la tabla 4.10, que el porcentaje de mejora en algunos casos es negativo y en otros positivos lógicamente puesto que al mejorar las máscaras de salida las puntuaciones han de verse modificadas con respecto a las originales, y pese a que en algún caso alguna medida tenga un porcentaje reseñable de pérdida lo obtiene como mejora en otra puntuación. Para el fondo se obtiene una mejora escasa dado que lo que mejoramos son la eliminación del alguna clasificación errónea como frente, por lo que las puntuaciones que si se ven mejoradas son las relativas al frente como era de esperar.

Al igual que en la aproximación anterior se presentan ahora unos resultados visuales, con el motivo de obtener una mejor apreciación de las mejoras obtenidas.

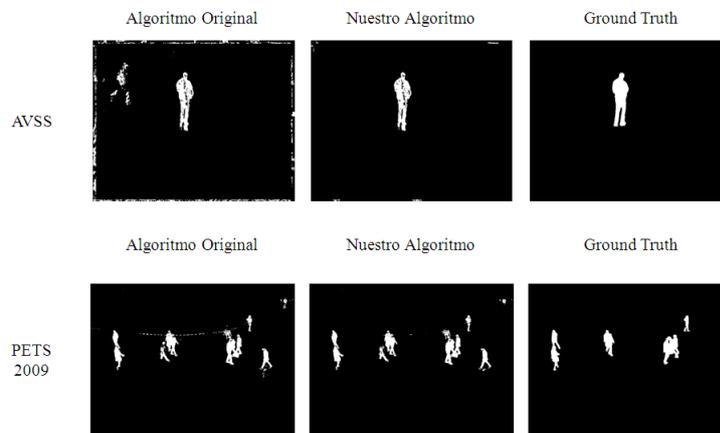


Figura 4.15: Imágenes (1) Algoritmo Frente-Fondo Desarrollado



Figura 4.16: Imágenes (2) Algoritmo Frente-Fondo Desarrollado

A través de las Figuras 4.15 y 4.16, se puede apreciar la mejora obtenida a través

del algoritmo desarrollado. Se puede apreciar que, se eliminan zonas donde la clasificación ha sido errónea de frente por el algoritmo original. Sí que es cierto, que hay ciertas zonas que no son eliminadas teniendo su motivo en la calidad de la segmentación Persona-Fondo, ya que es la información introducida para mejorar el algoritmo.

Una vez que tenemos ambas evaluaciones, tanto la visual como la de puntuaciones nos comparamos con las tablas 4.7 y 4.8 de la sección 4.4.4 que son los límites que hemos establecido de frente y fondo respectivamente, y estamos dentro de los mismos par todos los segmentadores en algunos casos se mejora y en otros se empeora, por lo que los límites que se establecieron guardan relación con las puntuaciones obtenidas en esta evaluación.

En esta aproximación buscamos mejorar la segmentación Frente-Fondo usando la información de la segmentación Persona-Fondo. por lo que si se sigue la evaluación de [2], se van a aplicar ahora los factores de penalización para la puntuación falsos positivos como en la evaluación del algoritmo Persona-Fondo. En este caso, se van a mostrar gráficas de las puntuaciones a las cuales está modificación afecta en mayor medida.

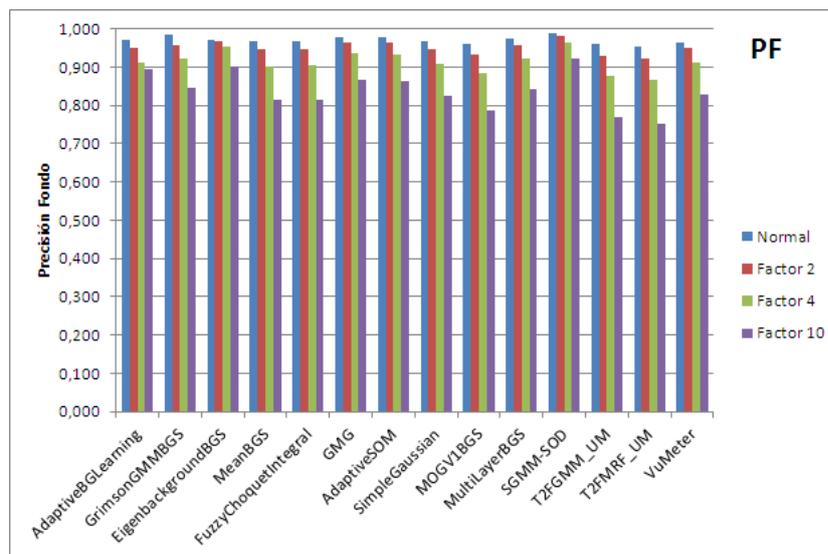


Figura 4.17: Gráfica puntuación PF de la tercera aproximación para distintos factores de penalización

En las Figuras 4.17 4.18 se pueden apreciar el efecto que provocan los factores de penalización, el cual cumple con lo que cabría esperar, puesto que el algoritmo original penaliza más los falsos positivos que otros tipos de segmentadores, por tanto, a medida que se aumenta el factor el valor de la puntuación es menor. En este caso es al contrario que en el resto de las aproximaciones, los factores de penalización afectan

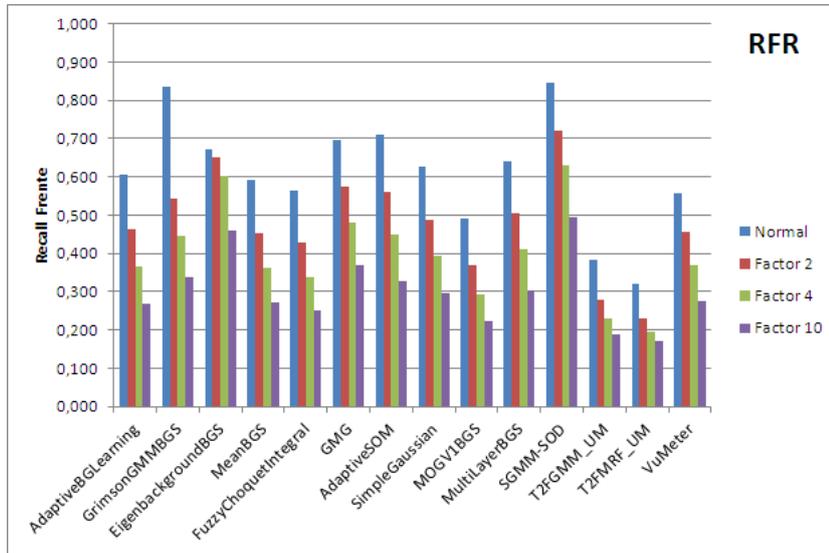


Figura 4.18: Gráfica puntuación RFR de la tercera aproximación para distintos factores de penalización

en mayor medida, puesto que los segmentadores Frente-Fondo no están diseñados para clasificar de manera precisa el fondo, sino que buscan la clasificación fina del frente.

Para finalizar, en esta aproximación las mejoras observadas en las puntuaciones no son tan claras como en la anterior, pero si es cierto que visualmente se aprecian mejoras claras. Además del desarrollo realizado se podrían haber realizado otras implementaciones diversas para mejorar los resultados.

4.4.7. Conclusiones

Para concluir con el capítulo, los algoritmos de segmentación Persona-Fondo como ya suponíamos nos daban puntuaciones altas en referencia al fondo, puesto que su objetivo principal es clasificar de forma correcta las zonas donde no encontramos personas. Por su parte, los segmentadores Frente-Fondo nos dan una correcta y fina de clasificación de los objetos de frente ya que este es su objetivo. En el caso de la primera aproximación, los resultados obtenidos no han sido los esperados debido a la alta variabilidad de los mapas de confianza. Por tanto, al unir ambas informaciones en las dos aproximaciones realizadas, la primera de estas la mejora del algoritmo de segmentación Persona-Fondo y en la segunda que tuvo lugar a la colación de los resultados de la primera, se mejoran los resultados obtenidos de forma independiente obteniendo puntuaciones superiores y un resultado visual con mejoría, como los que se han mostrado en secciones anteriores.

Por tanto, ambas aproximaciones nos dan una mejora significativa tanto para frente como para fondo, además de conseguir prácticamente eliminar los posibles errores generados en cada uno de las dos segmentaciones de forma independiente.

Capítulo 5

Conclusiones y Trabajo Futuro

5.1. Conclusiones

Al principio de este proyecto en la sección 1, se ha expuesto la gran evolución que ha generado la tecnología de tratamiento de vídeo a lo largo del tiempo, las posibles problemáticas a grandes rasgos del análisis de vídeo que se realiza en la actualidad. También se han expuesto de manera escueta las dos ramas de la segmentación que se iban a usar en este proyecto con sus ventajas e inconvenientes. Tras esto, se expuso el objetivo principal del proyecto el cual es la mejora de la segmentación Persona-Fondo utilizando la información de la segmentación Frente-Fondo.

Posteriormente, se ha analizado el Estado del Arte de la segmentación Persona-Fondo y Frente-Fondo, capítulo 2. En él se han analizado algoritmos de segmentación Persona-Fondo, sección 2.2, y algoritmos de segmentación Frente-Fondo, sección 2.3. Tras el estudio de ambas segmentaciones, se extrajeron las metodologías utilizadas en los algoritmos al igual que sus principales ventajas e inconvenientes de los mismos.

Una vez se tenía la información necesaria, en el capítulo 3, se ha explicado en detalle el algoritmo que se ha escogido de la segmentación Persona-Fondo, sección 3.2, se han seleccionado los algoritmos de segmentación Frente-Fondo de una amplia selección disponible, sección 3.3, así mismo cada uno de los seleccionados han sido explicados en el apéndice B.

A continuación se desarrollaron las aproximaciones presentadas en este trabajo 3.4, la primera aproximación 4.4.3 era aplicar la información que generan los segmentadores Frente-Fondo sobre los mapas de confianza generados por el segmentador Persona-Fondo, se observó que la mejora era mínima, esto se debe principalmente al comportamiento “intermitente” de la segmentación Persona-Fondo, por lo que la aplicación directa del modelado de fondo a esta no consigue grandes mejoras. Tras esto,

se desarrolló la segunda aproximación 4.4.5 la cual trataba de unir las informaciones de ambas segmentaciones a través de sus máscaras binarias de salida para mejorar la segmentación Persona-Fondo, obteniendo esta vez sí una mejora sustancial de los resultados del segmentador original. Por último, se realizó una tercera aproximación 4.4.6 siguiendo esta el mismo proceso para mejora de la segmentación Frente-Fondo que en la segunda aproximación, pudiéndose observar mejoras en la segmentación Frente-Fondo.

Tras esto, se procedió a la evaluación, sección 4, para la segmentación Persona-Fondo original, sección 4.4.1, como para los segmentadores Frente-Fondo, sección 4.4.2. Por tanto, ya teniendo una base de los resultados originales, se procedió a la evaluación del algoritmo propuesto, secciones 4.4.3, 4.4.5 y 4.4.6.

Una vez realizada la evaluación, se puede concluir que se ha alcanzado el objetivo principal del proyecto siendo este, la mejora de la segmentación Persona-Fondo gracias a la adición de la información de segmentación Frente-Fondo. Ya sea visualmente como con las puntuaciones, la mejora obtenida es significativa ya sea en Precisión, Recall o F1Score. En el caso de la segmentación Persona-Fondo, obtenemos una mejora clara en las puntuaciones del frente, puesto que se eliminan falsas detecciones y por tanto la precisión y recall aumentan en comparación a la originales. Además, se consiguió otro objetivo que no estaba en este proyecto, siendo el mismo la mejora de la segmentación Frente-Fondo, utilizando la información de segmentación Persona-Fondo. Las mejoras obtenidas no son tan significantes como en el objetivo principal del proyecto, pero si se han de tener en cuenta. En este caso, la mejora que se obtiene es la corrección de clasificaciones erróneas de frente, como puedan ser fantasmas, sombras y otras problemáticas.

5.2. Trabajo Futuro

Una vez finalizado este proyecto, existen diversas opciones de investigación futuras a través de las cuales se podrían mejorar los resultados obtenidos en este proyecto. Por lo que, se van a resumir ahora algunas de las ideas posibles:

1. En primer lugar, sería interesante el uso de otro tipo de segmentadores Frente-Fondo distintos a los pertenecientes a *Background Subtraction* utilizados, dado que otro utilizar segmentadores con otro enfoque puede mejorar en múltiples factores. Además, muchos de ellos pueden contener una algoritmia que favorezca a tener unos mejores resultados de segmentación.
2. Con respecto a las máscaras obtenidas por cada tipo de segmentación, se po-

dría aplicar un post-procesado a las máscaras para en el caso de la segmentación Frente-Fondo con un cierre, por ejemplo, conseguir que las personas estén completamente clasificadas como frente (color blanco en la máscara) si no lo están. A parte de este ejemplo, hay múltiples posibilidades para el post-procesado de las máscaras.

3. En cuanto a la primera aproximación, debido a que la confianza por el segmentador Persona-Fondo es muy variable sería interesante probar a modelarlo en el tiempo con otras técnicas, como el modelado de fondos multimodales para intentar detectar aquellas regiones que se repiten aleatoriamente de las que corresponde a personas y en teoría se repetirán de forma más constante en el tiempo, pudiendo así mejorar con la información temporal la segmentación Persona-Fondo.
4. En cuanto a las aproximaciones, se podrían haber utilizado diversos métodos y procesos para unir las informaciones de ambas segmentaciones. Pudiendo ser uno de los casos, no usar solo las máscaras binarias, sino utilizar los mapas de confianza en escala de grises y otras muchas implementaciones posibles.
5. Se podrían mejorar aspectos de la arquitectura del algoritmo diseñado para una mayor eficiencia y velocidad de proceso. En cuanto al coste computacional, no está contemplado como objetivo de este proyectos, pero la ejecución de las aproximaciones y del sistema final, se ve muy afectado por el número de frames que se van a evaluar.
6. Sería interesante la posibilidad de evaluarnos con otros segmentadores Persona-Fondo.

Bibliografía

- [1] T. B. Moeslund, A. Hilton, and V. Krüger, “A survey of advances in vision-based human motion capture and analysis,” *Computer Vision and Image Understanding*, vol. 104, no. 2-3, pp. 90–126, 2006.
- [2] A. Garcia-Martin, A. Cavallaro, and J. Martinez, “People-background segmentation with unequal error cost,” in *IEEE International Conference on Image Processing*, pp. 157–160, 2012.
- [3] A. Garcia-Martin, A. Hauptmann, and J. Martinez, “People detection based on appearance and motion models,” in *IEEE International Conference on Advanced Video and Signal-Based Surveillance*, pp. 256–260, 2011.
- [4] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, “Object detection with discriminatively trained part-based models,” *IEEE Transactions on Acoustics Speech and Signal Processing/Pattern Analysis and Machine Intelligence*, vol. 32, no. 9, pp. 1627–1645, 2010.
- [5] T. Bouwmans, “Traditional and recent approaches in background modeling for foreground detection: An overview,” *Computer Science Review*, vol. 11-12, pp. 31 – 66, 2014.
- [6] S. Brutzer, B. Hoferlin, and G. Heidemann, “Evaluation of background subtraction techniques for video surveillance,” in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1937–1944, 2011.
- [7] S. Y. Elhabian, K. M. El-Sayed, and S. H. Ahmed, “Moving object detection in spatial domain using background removal techniques – state-of-art,” *COMPUTER. SCI*, vol. 1, pp. 32–54, 2008.
- [8] S. Herrero and J. Bescós, “Background subtraction techniques: Systematic evaluation and comparative analysis,” in *Advanced Concepts for Intelligent Vision Systems*, vol. 5807, pp. 33–42, 2009.
- [9] D. Parks and S. Fels, “Evaluation of background subtraction algorithms with post-processing,” in *IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 192–199, Sept 2008.

- [10] F. El Baf, T. Bouwmans, and B. Vachon, "A fuzzy approach for background subtraction," in *IEEE International Conference on Image Processing*, pp. 2648–2651, 2008.
- [11] L. Unzueta, M. Nieto, A. Cortes, J. Barandiaran, O. Otaegui, and P. Sanchez, "Adaptive multicue background subtraction for robust vehicle counting and classification," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 2, pp. 527–540, 2012.
- [12] P. Blauensteiner and M. Kampel, *Visual surveillance of an airports apron-An overview of the AVITRACK project*. 2004.
- [13] C. Spampinato, Y.-H. Chen-Burger, G. Nadarajan, and R. B. Fisher, "Detecting, tracking and counting fish in low quality unconstrained underwater videos.," *VISAPP*, vol. 2008, pp. 514–519, 2008.
- [14] M. Paul, W. Lin, C. T. Lau, and B.-S. Lee, "Pattern-based video coding with dynamic background modeling," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, pp. 1–15, 2013.
- [15] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers, "Wallflower: principles and practice of background maintenance," in *IEEE International Conference on Computer Vision, The Proceedings of the Seventh*, pp. 255–261, 1999.
- [16] H. Kim, R. Sakamoto, I. Kitahara, T. Toriyama, and K. Kogure, "Robust foreground extraction technique using gaussian family model and multiple thresholds," in *Computer Vision*, vol. 4843, pp. 758–768, 2007.
- [17] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of the human body," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, 1997.
- [18] T. Bouwmans, F. El Baf, and B. Vachon, "Background modeling using mixture of gaussians for foreground detection-a survey," *Recent Patents on Computer Science*, vol. 1, no. 3, pp. 219–237, 2008.
- [19] A. Elgammal, D. Harwood, and L. Davis, "Non-parametric model for background subtraction," in *Computer Vision*, vol. 1843, pp. 751–767, 2000.
- [20] H.-H. Lin, T.-L. Liu, and J.-H. Chuang, "A probabilistic svm approach for background scene initialization," in *International Conference on Image Processing*, vol. 3, pp. 893–896, 2002.
- [21] A. Tavakkoli, M. Nicolescu, and G. Bebis, "A novelty detection approach for foreground region detection in videos with quasi-stationary backgrounds," in *Advances in Visual Computing*, vol. 4291, pp. 40–49, 2006.
- [22] T. Bouwmans, "Subspace learning for background modeling: A survey," *Recent Patents on Computer Science*, vol. 2, no. 3, pp. 223–234, 2009.
- [23] D. E. Butler, V. M. Bove, and S. Sridharan, "Real-time adaptive foreground/background segmentation," *Journal on Advances in Signal Processing*, vol. 2005, no. 14, pp. 2292–2304, 1900.

- [24] K. Kim, T. H. Chalidabhongse, D. Harwood, and L. Davis, "Background modeling and subtraction by codebook construction," in *International Conference on Image Processing*, vol. 5, pp. 3061–3064, 2004.
- [25] M. Xiao, C. Han, and X. Kang, "A background reconstruction for dynamic scenes," in *IEEE International Conference on Information Fusion*, pp. 1–7, 2006.
- [26] D. Culibrk, O. Marques, D. Socek, H. Kalva, and B. Furht, "A neural network approach to bayesian background modeling for video object segmentation," pp. 474–479, 2006.
- [27] R. M. Luque, D. Lopez-Rodriguez, E. Mérida-Casermeyro, and E. J. Palomo, "Video object segmentation with multivalued neural networks," in *IEEE International Conference on Hybrid Intelligent Systems*, pp. 613–618, 2008.
- [28] R. M. Luque, E. Domínguez, E. J. Palomo, and J. Muñoz, "A neural network approach for video object segmentation in traffic surveillance," in *Image Analysis and Recognition*, pp. 151–158, 2008.
- [29] R. M. Luque, D. López-Rodríguez, E. Dominguez, and E. J. Palomo, "A dipolar competitive neural network for video segmentation," in *Advances in Artificial Intelligence*, pp. 103–112, 2008.
- [30] L. Maddalena and A. Petrosino, "A self-organizing approach to background subtraction for visual surveillance applications," *IEEE Transactions on Image Processing*, vol. 17, no. 7, pp. 1168–1177, 2008.
- [31] L. Maddalena and A. Petrosino, "The sobs algorithm: what are the limits?," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 21–26, 2012.
- [32] F. Di, M. Cao, and C. Lv, "An updating method of self-adaptive background for moving objects detection in video," in *IEEE International Conference on Audio, Language and Image Processing*, pp. 1497–1501, 2008.
- [33] R. Chang, T. Gandhi, and M. M. Trivedi, "Vision modules for a multi-sensory bridge monitoring approach," in *IEEE International Conference on Intelligent Transportation Systems*, pp. 971–976, 2004.
- [34] D. Mukherjee and Q. JonathanWu, "Real-timevideosegmentation using student'stmixture model," *Procedia Computer Science*, vol. 10, pp. 153–160, 2012.
- [35] T. S. Haines and T. Xiang, "Background subtraction with dirichlet processes," in *Computer Vision*, pp. 99–113, 2012.
- [36] J. Ding, M. Li, K. Huang, and T. Tan, "Modeling complex scenes for accurate moving objects segmentation," in *Computer Vision*, pp. 82–94, 2011.
- [37] M. Hofmann, P. Tiefenbacher, and G. Rigoll, "Background segmentation with feedback: The pixel-based adaptive segmenter," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 38–43, 2012.

- [38] S. Molina-Giraldo, J. C. González, A. M. Álvarez-Meza, and G. Castellanos-Domínguez, “Video segmentation based on multi-kernel learning and feature relevance analysis for object classification.,” in *ICPRAM*, pp. 396–401, 2013.
- [39] C. Stauffer and W. E. L. Grimson, “Adaptive background mixture models for real-time tracking,” in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2, 1999.
- [40] F. El Baf, T. Bouwmans, and B. Vachon, “Type-2 fuzzy mixture of gaussians model: application to background modeling,” in *Advances in Visual Computing*, pp. 772–781, Springer, 2008.
- [41] H. Zhang and D. Xu, “Fusing color and texture features for background model,” in *Fuzzy Systems and Knowledge Discovery: Third International Conference, FSKD 2006, Xian, China, September 24-28, 2006. Proceedings*, pp. 887–893, Springer, 2006.
- [42] F. El Baf, T. Bouwmans, and B. Vachon, “Foreground detection using the choquet integral,” in *IEEE International Workshop on Image Analysis for Multimedia Interactive Services*, pp. 187–190, 2008.
- [43] L. Maddalena and A. Petrosino, “Multivalued background/foreground separation for moving object detection,” in *Fuzzy Logic and Applications*, pp. 263–270, 2009.
- [44] H. Baradaran Kashani, S. A. Seyedin, and H. Sadoghi Yazdi, “A novel approach in video scene background estimation,” *International Journal of Computer Theory and Engineering*, vol. 2, 2010.
- [45] C. Pojala and S. Sengupta, “Detection of moving objects using fuzzy correlogram based background subtraction,” in *IEEE International Conference on Signal and Image Processing Applications*, pp. 255–259, 2011.
- [46] P. Chiranjeevi and S. Sengupta, “New fuzzy texture features for robust detection of moving objects,” *IEEE Signal Processing Letters*, vol. 19, no. 10, pp. 603–606, 2012.
- [47] P. Chiranjeevi and S. Sengupta, “Robust detection of moving objects in video sequences through rough set theory framework,” *Image and Vision Computing*, vol. 30, no. 11, pp. 829–842, 2012.
- [48] M. Sivabalakrishnan and D. Manjula, “Adaptive background subtraction in dynamic environments using fuzzy logic.,” *International Journal of Video & Image Processing & Network Security*, vol. 10, no. 1, 2010.
- [49] D. Skocaj and A. Leonardis, “Canonical correlation analysis for appearance-based orientation and self-estimation and self-localization,” in *CogVis Meeting*, 2004.
- [50] J. Yan, B. Zhang, S. Yan, Q. Yang, H. Li, Z. Chen, W. Xi, W. Fan, W.-Y. Ma, and Q. Cheng, “Immc: incremental maximum margin criterion,” in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 725–730, 2004.

- [51] C. Marghes, T. Bouwmans, and R. Vasiu, “Background modeling and foreground detection via a reconstructive and discriminative subspace learning approach,” in *International Conference on Image Processing, Computer Vision, and Pattern Recognition*, 2012.
- [52] E. J. Candès, X. Li, Y. Ma, and J. Wright, “Robust principal component analysis?,” *Journal of the ACM*, vol. 58, no. 3, p. 11, 2011.
- [53] X. Zhou, C. Yang, and W. Yu, “Moving object detection by detecting contiguous outliers in the low-rank representation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 3, pp. 597–610, 2013.
- [54] L. Xiong, X. Chen, and J. Schneider, “Direct robust matrix factorization for anomaly detection,” in *In International Conference on Data Mining*, 2010.
- [55] N. Wang, T. Yao, J. Wang, and D.-Y. Yeung, “A probabilistic approach to robust matrix factorization,” in *Computer Vision*, pp. 126–139, 2012.
- [56] N. Wang and D.-Y. Yeung, “Bayesian robust matrix factorization for image and video processing,” in *IEEE International Conference on Computer Vision*, pp. 1785–1792, 2013.
- [57] V. Cevher, A. Sankaranarayanan, M. F. Duarte, D. Reddy, R. G. Baraniuk, and R. Chellappa, “Compressive sensing for background subtraction,” in *Computer Vision*, pp. 155–168, 2008.
- [58] J. Huang, T. Zhang, and D. Metaxas, “Learning with structured sparsity,” *The Journal of Machine Learning Research*, vol. 12, pp. 3371–3412, 2011.
- [59] J. Huang, X. Huang, and D. Metaxas, “Learning with dynamic group sparsity,” in *IEEE International Conference on Computer Vision*, pp. 64–71, 2009.
- [60] C. David, V. Gui, and F. Alexa, “Foreground/background segmentation with learned dictionary,” in *IEEE International Conference on Applied Mathematics, Simulation, Modelling, Circuits, Systems and Signals. Washington: Computer Society*, vol. 197, p. 201, 2009.
- [61] M. Dikmen and T. S. Huang, “Robust estimation of foreground in surveillance videos by sparse error estimation,” in *IEEE International Conference on Pattern Recognition*, pp. 1–4, 2008.
- [62] C. R. Wren and F. Porikli, “Waviz: Spectral similarity for object detection,” in *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance*, pp. 55–61, 2005.
- [63] F. Porikli and C. Wren, “Change detection by frequency decomposition: Wave-back,” in *Proc. of Workshop on Image Analysis for Multimedia Interactive Services*, 2005.
- [64] H. Tezuka and T. Nishitani, “A precise and stable foreground segmentation using fine-to-coarse approach in transform domain,” in *IEEE International Conference on Image Processing*, pp. 2732–2735, 2008.

- [65] T. Gao, Z.-g. Liu, W.-c. Gao, and J. Zhang, "A robust technique for background subtraction in traffic video," in *Advances in Neuro-Information Processing*, pp. 736–744, 2009.
- [66] Y.-P. Guan, "Wavelet multi-scale transform based foreground segmentation and shadow elimination," *The Open Signal Processing Journal*, vol. 1, no. 6, pp. 1–6, 2008.
- [67] D. Gao, Z. Jiang, and M. Ye, "A new approach of dynamic background modeling for surveillance information," in *IEEE International Conference on Computer Science and Software Engineering*, vol. 1, pp. 850–855, 2008.
- [68] A. S. Jalal and V. Singh, "A robust background subtraction approach based on daubechies complex wavelet transform," in *Advances in Computing and Communications*, pp. 516–524, 2011.
- [69] D. Baltieri, R. Vezzani, and R. Cucchiara, "Fast background initialization with recursive hadamard transform," in *IEEE International Conference on Advanced Video and Signal Based Surveillance*, pp. 165–171, 2010.
- [70] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 886–893, 2005.
- [71] R. Evangelio and T. Sikora, "Complementary background models for the detection of static and moving objects in crowded environments," in *IEEE International Conference on Advanced Video and Signal-Based Surveillance*, pp. 71–76, 2011.
- [72] K. McKoen, R. Navarro-Prieto, B. Duc, E. Durucan, F. Ziliani, and T. Ebrahimi, "Evaluation of video segmentation methods for surveillance applications," in *Proc. European Signal Processing Conference*, 2000.
- [73] P. Correia and F. Pereira, "Objective evaluation of relative segmentation quality," in *IEEE International Conference on Image Processing*, vol. 1, pp. 308–311, 2000.
- [74] S. Herrero, "Proyecto fin de carrera de técnicas, análisis comparativo and de segmentación, de secuencias de and el modelado, vídeo basadas en and fondo, del," 2009.
- [75] N. McFarlane and C. Schofield, "Segmentation and tracking of piglets in images," *Machine Vision and Applications*, vol. 8, no. 3, pp. 187–193, 1995.
- [76] S. Calderara, R. Melli, A. Prati, and R. Cucchiara, "Reliable background suppression for complex scenes," in *Proceedings international workshop on Video surveillance and sensor networks*, pp. 211–214, 2006.
- [77] M. H. Sigari, N. Mozayani, and H. R. Pourreza, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 138–143, 2008.
- [78] Y. Benezeth, P.-M. Jodoin, B. Emile, H. Laurent, and C. Rosenberger, "Review and evaluation of commonly-implemented background subtraction algorithms," in *IEEE International Conference on Pattern Recognition*, pp. 1–4, 2008.

- [79] Z. Zivkovic, “Improved adaptive gaussian mixture model for background subtraction,” in *IEEE International Conference on Pattern Recognition*, vol. 2, pp. 28–31, 2004.
- [80] Z. Zivkovic and F. van der Heijden, “Recursive unsupervised learning of finite mixture models,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 651–656, 2004.
- [81] J. Zeng, L. Xie, and Z.-Q. Liu, “Type-2 fuzzy gaussian mixture models,” *Pattern Recognition*, vol. 41, no. 12, pp. 3636–3643, 2008.
- [82] Z. Zhao, T. Bouwmans, X. Zhang, and Y. Fang, “A fuzzy background modeling approach for motion detection in dynamic backgrounds,” in *Multimedia and Signal Processing*, vol. 346, pp. 177–185, 2012.
- [83] M. Heikkila and M. Pietikainen, “A texture-based method for modeling the background and detecting moving objects,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 4, pp. 657–662, 2006.
- [84] C. Kertész, “Texture-based foreground detection,” 2011.
- [85] J. Yao and J.-M. Odobez, “Multi-layer background subtraction based on color and texture,” in *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8, 2007.
- [86] S. Noh and M. Jeon, “A new framework for background subtraction using multiple cues,” in *Computer Vision*, vol. 7726, pp. 493–506, 2013.
- [87] A. B. Godbehere, A. Matsukawa, and K. Goldberg, “Visual tracking of human visitors under variable-lighting conditions for a responsive audio art installation,” in *American Control Conference*, pp. 4305–4312, 2012.
- [88] Y. Goyat, T. Chateau, L. Malaterre, and L. Trassoudaine, “Vehicle trajectories evaluation by static video sensors,” in *IEEE Intelligent Transportation Systems Conference*, pp. 864–869, 2006.
- [89] D. Bloisi and L. Iocchi, “Independent multimodal background subtraction,” in *Proceedings of the Third International Conference on Computational Modeling of Objects Presented in Images: Fundamentals, Methods and Applications*, pp. 39–44, 2012.
- [90] N. M. Oliver, B. Rosario, and A. P. Pentland, “A bayesian computer vision system for modeling human interactions,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 831–843, 2000.
- [91] L. Maddalena and A. Petrosino, “A fuzzy spatial coherence-based approach to background/foreground separation for moving object detection,” *Neural Computing and Applications*, vol. 19, no. 2, pp. 179–186, 2010.

Apéndice A

Acrónimos

- BGS - Background Subtraction.
- BRMF - Matriz de Factorización Bayesiana y Robusta.
- DBP - Dependet Body Part.
- DEBP - Dependet Extend Body Part.
- DEBP-P - Dependent Extend Body Part Postprocess.
- DECOLOR - Detecting Contiguous Outlier detection in the Low-rank Representation.
- DGS - Dynamic Group Sparsity.
- DMM - Modelos Discretos de Mezcla.
- DRMF - Matriz de Factorización Directa y Robusta.
- DW - Dyadic Wavelet.
- GMM - Modelo Mezcla de Gaussianas.
- HOG - Histogram of Gradient.
- IBP - Independet Body Part.
- IEBP - Independet Extend Body Part.
- IMBS - Independet Multimodal Background Subtraction.
- IMMC - Incremental Maximum Margin Criterion.

- IRLS - Iteratively Reweighted Least Squares.
- ISM - Implicit Shape Model.
- KDE - Kernel Density Estimation.
- LBP - Local Binary Pattern.
- MGM - Mixture Gaussian Model.
- MOG - Mezcla de Gaussianas.
- MRP - Markov Random Field.
- PBAS - Pixel-Based Adaptive Segmenter.
- PCA - Principal Component Analysis.
- PCP - Principal Component Pursuit.
- PRMF - Matriz de Factorización Probabilística y Robusta.
- RPCA - Robust Principal Components Analysis.
- SAKBOT - Statical and Knowledge-Based Object Tracker.
- STMM - Modelos de Mezcla Student-t.
- SVM - Support Vector Models.
- TUB - Technical University of Berlin.
- ViBe - Visual Background Extractor.
- VPU - Video Processing and Understanding Lab.
- WGKVS - Weighted Gaussian Models.

Apéndice B

Descripción algoritmos Fondo-Frente

En este apéndice, se describen los segmentadores Frente-Fondo disponibles en la biblioteca BGSLibrary sección 3.3.2 del capítulo 3. Se describen de manera breve, sin entrar en detalles muy específicos, si reseñando las principales mejoras que introducen o alguna operativa específica.

MEDIANA ADAPTATIVA [75]: El algoritmo esta insertado en un *tracker*, del cual no se realizará mención alguna puesto que no es el motivo de este proyecto. Se escoge la mediana adaptativa antes que la media, porque rechaza mejor valores atípicos de la distribución de los píxeles. Por lo que el cambio de valores sobre una zona de la imagen no afectará en demasía comparado con la media. Para obtener un modelo de fondo, se utilizan una secuencia de imágenes para generar con la mediana el modelo. Posteriormente la actualización del modelo se realiza con una máscara indicando las regiones de interés que han sufrido modificaciones.

Posteriormente se aplica un operador Laplaciano, para mejorar el modelo de fondo, mejorando las zonas que tienen curvaturas. Es muy crítica la decisión del tamaño del operador. Después ya se aplica la segmentación aplicando una diferencia de frames.

MEDIANA TEMPORAL [76]: Se presenta un algoritmo robusto basado en la eliminación de fondo, que mejora el algoritmo SAKBOT (*Static and Knowledge-Based Object Tracker*) que mejora la fiabilidad en escenarios complejos. En este caso la inicialización del fondo a de ser rápida y precisa incluso si el inicio de la secuencia no es limpia (Si hay personas nada más iniciar la secuencia). La aproximación realizada divide la imagen en bloques de 16x16 píxeles y la actualización de fondo la realiza selectivamente cuando por bloques cuando el bloque tiene un número suficiente de píxeles sin movimiento. El movimiento se capta con una imagen diferencia obtenida

de frames consecutivos y umbralizada. Si el 95 % de los píxeles en un bloque no sufren cambios (no hay movimiento) ese bloque es actualizado y si además, esto ocurre más de diez veces ese bloque se considera estable y no se vuelve a evaluar. Cuando todos los bloques se consideran estables, el modelo de fondo ya está listo.

Para la actualización del modelo se usa la mediana temporal, previamente se han almacenado para cada píxel un vector con sus k valores de intensidad de nivel de gris en el tiempo. Dentro del vector se ordena los valores por intensidad para luego realizar la mediana de estos y así se estima el modelo de fondo. Para la extracción del frente se aplica la técnica básica de diferencia de frames. Posteriormente la máscara se obtiene usando dos umbrales, uno bajo para eliminar ruido y el otro alto para localizar una variación grande de intensidad, ambos umbrales son locales. El hecho de que estos umbrales sean locales es, para combatir la problemática que se encuentra en esta temática como cambios de iluminación y otros más.

Para la eliminación de sombras se propone el uso de un modelo de detección de sombras basado en el matiz, saturación y el valor de un espacio de color HSV. El método usa dos umbrales, el primero para el valor máximo de efecto de oscurecimiento de las sombras en el fondo y el segundo para prevenir de identificar como sombras los puntos donde el fondo está oscurecido de menos con respecto a las sombras. Además, se aplica una validación de objetos para eliminar los pequeños objetos clasificados como frente siendo estos del fondo. Se usa información de color y gradientes de los objetos.

FUZZY SUGENO INTEGRAL [41]: Se realiza una nueva aproximación de BGS, se utiliza una estimación no paramétrica de KDE general para generar un modelo estadístico de fondo. La función de densidad de probabilidad de la intensidad de píxel se estima directamente de datos. Se aplica un algoritmo de filtrado Kalman, para estimar iterativamente la apariencia intrínseca de las texturas dinámicas al igual que las regiones de los objetos de frente.

Para la obtención de las características del color se usa un espacio de color Ohta. Para comprobar la similitud entre el fondo y el frame actual se aplican las características de color de incertidumbre del espacio Ohta. En el caso de las características de las texturas se utiliza un operador LBP. Dicho operador etiqueta los píxeles de un bloque escogiendo un umbral para los vecinos del píxel central y considerando el resultado dado por el operador.

Para la obtención del frente se usa una integral de tipo Sugeno, para integrar las características de dos píxeles en la misma posición del fondo y el frame actual y decidir cuál si son similares o no. Se decide un umbral para ver la similitud.

Para la actualización del fondo, se usa la característica del movimiento del píxel

para decidir si se actualiza o no, por lo que los píxeles que tengan poco movimiento serán actualizados por considerarse fondo.

FUZZY CHOQUET INTEGRAL [42]: Se realiza una nueva aproximación de BGS, se utiliza una estimación no paramétrica de KDE general para generar un modelo estadístico de fondo. La función de densidad de probabilidad de la intensidad de píxel se estima directamente de datos. Se aplica un algoritmo de filtrado Kalman, para estimar iterativamente la apariencia intrínseca de las texturas dinámicas al igual que las regiones de los objetos de frente.

Para la obtención de las características del color se usa un espacio de color Ohta. Para comprobar la similitud entre el fondo y el frame actual se aplican las características de color de incertidumbre del espacio Ohta. En el caso de las características de las texturas se utiliza un operador LBP. Dicho operador etiqueta los píxeles de un bloque escogiendo un umbral para los vecinos del píxel central y considerando el resultado dado por el operador.

Para la obtención del frente se usa una integral de tipo Choquet, para integrar las características de dos píxeles en la misma posición del fondo y el frame actual y decidir cuál si son similares o no. Se decide un umbral para ver la similitud.

Para la actualización del fondo, se usa la característica del movimiento del píxel para decidir si se actualiza o no, por lo que los píxeles que tengan poco movimiento serán actualizados por considerarse fondo.

FUZZY GAUSSIAN [77]: Uno de los modelos más simples para modelado de fondo es *Running Average*, y la posterior extracción de frente se realiza con una diferencia de frames entre el modelo de fondo y el frame actual, además de la posterior umbralización. Se propone por tanto un método de sustracción de fondo de tipo *fuzzy* usando función de saturación lineal, en lugar de las funciones comunes que se usan para determinar si el píxel tiene movimiento o no. El resultado de la sustracción de fondo no es el común, 1 frente y 0 fondo, sino que son valores entre 0 y 1. Previamente a la binarización la máscara obtenida se pasa por un filtro paso bajo, obteniendo una mayor fiabilidad y robustez para detectar objetos en movimiento.

Para la actualización de fondo aplica una modificación del *Running Average* añadiéndole teoría *fuzzy*, puesto que en vez de tener un parámetro del primero global para el frame lo tiene para cada píxel. Siendo este parámetro la tasa de aprendizaje, usada para ver cuanta información del modelo de fondo y el frame actual se usan para actualizar el modelo de fondo.

GAUSSIANA SIMPLE [78]: La mayoría de técnicas da por hecho que las secuencias son imágenes formadas por un fondo fijo y objetos que se mueven. La gran diferencia entre los métodos es como se modela el fondo, en este caso se utilizan la

distribución Gaussiana. En este caso se modela el fondo a través de generar una distribución Gaussiana para cada píxel, por lo que cada píxel está representado por su media de color y su covarianza. La media de la distancia puede ser con *log likelihood* o con la distancia de Mahalanobis. Cada distribución es actualizada con cada nuevo frame.

GAUSSIANA MEDIA [17]: Este algoritmo viene implementado en un *tracker*, por lo que solo se explicara la parte relativa a la segmentación. La diferencia con otros métodos es que no se utiliza el valor de la intensidad del píxel, sino que es la media de los valores de los planos Y,U y V del frame. Posteriormente se modela el cambio en el valor del píxel con una distribución Gaussiana. El modelo de fondo y segmentación se generan como el método de Gaussiana Simple antes descrito.

MEZCLA DE GAUSSIANAS 1[39]: En este algoritmo se modela cada píxel como una mezcla de Gaussianas, basado en la media y varianza de cada una de las Gaussianas de la mezcla, se determina cuál de las Gaussianas pertenece al fondo. Las Gaussianas utilizadas son adaptativas, cada vez que los parámetros de la Gaussiana son actualizados, se evalúan las distribuciones se usa una simple heurística para realizar la hipótesis de cual pertenece al fondo.

En cuanto al modelo de mezcla, cada píxel sufre un proceso siendo esto el cambio de su valor que se almacena en un vector. El vector de la historia del píxel es modelado con una mezcla de K Gaussianas. No todas las Gaussianas pesan por igual, sino que cada una tiene un peso diferente, además de que el número de Gaussianas de la mezcla se determina según la memoria disponible y el poder de procesamiento del dispositivo donde se ejecute el algoritmo. Un nuevo píxel en general, será representado por una de las componentes más predominante en la mezcla y se usará para actualizar el modelo.

Dado que cada píxel tiene una distribución se utiliza una aproximación k-means. Por lo que la aparición de un valor nuevo de píxel se busca en todas las distribuciones hasta que exista una coincidencia, siendo la coincidencia no exacta pudiendo tener una desviación el valor de 2.5, siendo este valor modificable para cada distribución. Si no hay coincidencia, la distribución menos probable es sustituida por una nueva creada a partir del nuevo valor.

En cuanto al modelado de fondo, nos interesa la distribución Gaussiana con mayor evidencia y menor varianza. Se ordenan las Gaussianas según el valor de su peso entre la varianza, siendo este valor el que incrementa a una distribución con más evidencia como la varianza disminuye. Se vuelven a estimar los parámetros de la mezcla y es suficiente para clasificar las distribuciones para ver cuál sirve para el fondo.

MEZCLA DE GAUSSIANAS 2[79]: El algoritmo es una mejora del algoritmo desarrollado en [80], dado que no solo los parámetros sino que también el número de

componentes se adapta para cada píxel. En este caso utiliza el modelo de mezcla de Gaussianas [39], ya explicado anteriormente. En cuanto a la selección del número de componentes, el peso de cada Gaussianas nos da la información de cuanto se parece un dato a esa distribución, por lo que si observamos todos los pesos de las Gaussianas se genera una distribución multimodal. Aplicando una función buscando obtener el más parecido, podemos estimar el peso de cada Gaussianas por lo que se puede ver cuantas componentes son necesarias para cada píxel.

MEZCLA DE GAUSSIANS 3[18]: Es la misma implementación que [39], antes explicado pero en el cálculo de las distancias utiliza la distancia Mahalanobis .

TYPE-2 FUZZY GMM-UM [40]: Se propone para el modelado de fondo el uso de un modelo de mezcla de Gaussianas con *Type-2 Fuzzy*. Este modelo parte de la mezcla de Gaussianas. El modelo usado es T2 FMGM [81], en cuanto a la aplicación del modelado de fondo, cada píxel se caracteriza por la intensidad del color. MGM se compone de K mezclas de componentes de Gaussianas multivariadas la diferencia con otros modelos es que, el vector de medias no es conocido por lo que puede utilizarse para secuencias en las que no funcionen las distribuciones que necesitan conocer las medias.

La detección de frente consiste en clasificar el píxel como fondo o frente, Se ordenan las Gaussianas como en [81] ordenando las Gaussianas que representan el fondo con los pesos mayores con una varianza pequeña.

TYPE-2 FUZZY GMM-UV [40]: Es idéntico al algoritmo anterior pero la incertidumbre la tiene en la varianza.

TYPE-2 FUZZY GMM-UM con MRF [82]: El algoritmo es una mejora de TYPE-2 FUZZY GMM-UM[40], aplicando proceso Bayesiano aplicando restricciones espacio temporales para mejorar el etiquetado en la detección.

TYPE-2 FUZZY GMM-UM con MRF [82]:El algoritmo es una mejora de TYPE-2 FUZZY GMM-UV[40], aplicando proceso Bayesiano aplicando restricciones espacio temporales para mejorar el etiquetado en la detección.

BGS TEXTURAS [83]: El algoritmo usa operadores LBP, puesto que es un gran para obtener las características de textura. El operador etiqueta el píxel de una imagen realizando una umbralización del vecindario. El modelado de fondo realiza el modelado a nivel de píxel que nos permite tener una mayor velocidad de procesamiento. Se considera el vector de características de un píxel a lo largo del tiempo como un proceso, se calcula el histograma con la información del LBP sobre un radio circular desde el píxel. Cada modelo de histograma tiene un peso para poder saber cuál es el que represente el modelo. Se obtiene el histograma del nuevo frame y se compara con el modelo de histogramas que tenemos buscando una coincidencia, y tras eso

se aplica una medida de la distancia entre los histogramas. Aplicando después una umbralización con la distancia, si hay coincidencia se denota como fondo y si no ocurriese sería frente.

TEXTURE-BASED FOREGROUND DETECTION WITH MRF [84]: Es una mejora del algoritmo TBMOD [83], añade una detección global de cambios en la escena y para ello lo que hace es cambiar el espacio de color de la cámara de captación a Luv. Si la diferencia de frames sumando los valores totales es mayor que un umbral significa que hay un cambio global de la escena, por lo que nos hace robustos a cambios de iluminación. Además añade un operador LBP, que trabaja con las texturas en escala de grises. Se calcula sobre vecindarios de 5×5 , mejorando al de la primera versión. Cada canal del frame es normalizado entre 0 y 255, después el frame se transforma a escala de grises para aplicar LBP y obtener los histogramas de los valores para cada píxel. El proceso de actualización sigue siendo el mismo que el original además de ayudarse para el proceso con MRF. Para la segmentación se aplica calculo de distancias y un umbral para la decisión de frente o fondo.

MULTI-LAYER BGS [85]: En este algoritmo se trabaja con las características de textura y color, usando LBP para textura y medidas fotométricas de la invarianza del color. En cuanto a LBP trabaja etiquetando el píxel como fondo o frente umbralizando el vecindario del píxel con el píxel central. La medida fotométrica de la invarianza del color se usa porque nos da una invarianza a los cambios de iluminación

El modelo de fondo utiliza la información de texturas y color, consiste en una lista de modos aprendidos de los datos observados hasta el frame actual. El menor de todos ellos es el representante del fondo, cada píxel tiene un tamaño diferente de lista dependiendo de la variación de valor que hayan tenido. Cada modo está constituido por 7 componentes, los vectores de color, la información de LBP, los pesos y el número de la capa que tiene el modo de fondo. Para la actualización del modelo, obtienen la información del LBP y el color para buscar si algún modo del fondo lo tiene. Si no se encuentra nada, se genera un nuevo modo con la información actual.

MULTICUE BGS[86]: El método propuesto emplea tres tipos de informaciones, la información de textura para obtener información de la ordenación geométrica de las intensidades de los píxeles, la segunda es el color de píxel para compensar los inconvenientes de las texturas y por último una región de apariencia para eliminar errores en el frente.

El frame que entra es convolucionado con un operador Gaussiano para eliminar el ruido y alta frecuencia, tras esto se cambia el espacio de color de RGB a HSV. En este método para el modelo de fondo tenemos dos niveles, usando codebooks para texturas y para color. Además la sustracción del fondo se realiza en dos niveles, el primero a

nivel de píxel utilizando el codebook de texturas y el de color para corregir errores. Posteriormente se aplican operaciones morfológicas para eliminar falsos negativos.

El segundo nivel es sobre las regiones, se aplica una verificación de los blobs por la apariencia lo que nos da los resultados finales de detección.

Los codebooks utilizados están basados en cajas que no hacen suposiciones sobre cambios de iluminación. Para la sustracción de fondo a nivel de píxel se genera un mapa inicial de frente con los operadores de texturas y color. Se eliminan puntos erróneos generados por los operadores de color y texturas.

A nivel de región se aplica un mapa de componentes conexas para detectar las regiones.

GMG [87]: Aplica un método probabilístico de segmentación de fondo. El sistema para cada píxel genera un histograma de color con las componentes RGB, y para eliminar el ruido de la cámara aplica una cuantificación del espacio de color. Para obtener un modelo de fondo toma los N primeros frames para obtener la función de probabilidad de masa de cada píxel, o lo que es lo mismo el modelo de fondo. Posteriormente genera una operación de elevación para generar un vector único de características. Este vector va almacenando los valores de cada frame y posteriormente se realiza la media de los valores.

Posteriormente se aplica una regla Bayesiana, para obtener la probabilidad de que cada píxel pertenezca a fondo o a frente. Tras esto se genera una máscara binaria para después aplicar un método de componentes conexas.

VUMETER [88]: Se propone un modelo no paramétrico basado en la estimación de una densidad de probabilidad discreta. El modelo probabilístico utiliza el valor de las componentes RGB de cada píxel, toma cada píxel como una variable aleatoria discreta con posibilidad de tener dos estados, fondo o frente. La probabilidad de ser fondo o frente depende del vector de color y de la posición del píxel, usando una función de más para obtener dicha probabilidad.

El modelo de actualización se realiza cada nuevo frame, siendo lo que se actualiza la función de masa. Al ser un algoritmo pensado para coches, se usa una función de observación de verosimilitud. Cada coche se toma como un *blob* a analizar, teniendo en cuenta esto, se asume que el centro del coche debe corresponder con el centro del *blob*. Además utiliza un filtro de partículas, que es un algoritmo de muestreo secuencial para estimar propiedades de variables escondidas en un modelo de Markov.

KDE [19]: El objetivo de este algoritmo y en particular de su modelado de fondo, que busca obtener información del frame actual para actualizar continuamente la información del modelo para adaptarse a los cambios en la escena. Por lo que se debe estimar una función de distribución en cada frame. En este caso la función se estima

con un estimador kernel. Una vez estimada la probabilidad se utiliza un umbral para decidir si es píxel de fondo o de frente, el umbral en este caso es global para todo el frame.

Dado que estamos hablando de una estimación con kernel necesitamos obtener su anchura. Hay dos tipos de variaciones en los valores de los píxeles, el primero saltos muy grandes de valor de intensidad por distintos objetos y el segundo son esos periodos cortos de tiempo que tienen variación continua de valor. Por tanto, para estimar la anchura se calcula la desviación media absoluta sobre los valores distintos de intensidad del píxel, si medimos la desviación entre dos valores consecutivos normalmente pertenecerán al mismo momento de tiempo de la distribución y solo unas pocas parejas serán de distribuciones distintas.

IMBS [89]: Uno de los métodos más populares para la detección de objetos en movimiento con cámaras estáticas es Background Subtraction (BS), pudiendo ser este en tiempo real. BS se basa en la resta del frame actual con un modelo de fondo. La creación del modelo de fondo es un resto dado que se pueden encontrar problemas como cambios de iluminación, sombras, camera *jitter*, movimientos de los elementos de fondo, etc.

En la literatura se pueden encontrar múltiples algoritmos de BS, siendo estos clasificados por análisis a nivel de píxel, región, frame e híbrido, otra posible clasificación posible es recursivos, o no. El primero de estos mantiene un modelo de fondo simple el cual se va actualizando por cada nuevo frame. Por el contrario, los no recursivos mantienen un buffer con un número de frames anteriores almacenadas y las analiza para generar el modelo de fondo. Una tercera clasificación los divide en predictivos y no predictivos, modelando como una serie temporal y generan un modelo dinámico de fondo los primeros, mientras que los no predictivos no tienen en cuenta el orden de observación de los frames y dan modelo de fondo de probabilidades.

Este algoritmo es de análisis a nivel de píxel, sin recursividad ni predicción. El algoritmo fue diseñado para tratar con fondo de agua, pero puede ser aplicado a cualquier escenario. El diseño tiene el propósito de conseguir un BS eficiente y rápido. El modelo de fondo en este trabajo utiliza N frames, siendo este número siempre menor que el número de frames totales para tener una menor carga computacional. El modelo de fondo contiene duplas de los valores de RGB y los píxeles asociados a esta combinación de RGB, con este modelado obtenemos información sobre la relación de los canales RGB.

Este modelo produce una discretización de una distribución de probabilidad no conocida, siendo esto uno de los motivos por lo que el algoritmo es más robusto. Se crea un nuevo modelo de fondo cada cierto tiempo, estimando en el algoritmo, la

independencia generada por que la creación del modelo de fondo no sea continua, hace que sea un método que se puede adaptar a cambios de iluminación y otros diversos problemas.

Un problema que se ha de solucionar es la eliminación de los píxeles seleccionados como frente de manera errónea, por lo que se implementa un módulo de eliminación de sombras. Este módulo es esencial para el buen funcionamiento de este algoritmo. En cuanto a la actualización del modelo de fondo no se realiza sobre la imagen completa, sino que solo se aplica sobre los píxeles clasificados como fondo.

EIGENBACKGROUND/ SL-PCA [90]: Para modelar el fondo se genera un espacio de tipo Eigen que describe el rango de apariencias que se han observado. El modelo se genera obteniendo de N frames la media y la matriz de covarianza, pudiéndose diagonalizar la matriz. Se aplica PCA para reducir la dimensionalidad del espacio. Cuando entra una nueva frame se proyecta sobre el espacio Eigen que contiene el modelo de fondo, aplicando la distancia Euclídea y aplicando un umbral se decide si es fondo o frente y se segmenta. El esto del papper es sobre tracking lo cual no nos compete.

ADAPTIVE SOM [30]: El modelo propuesto es crear un modelo de fondo, que aprenda solo con las variaciones de valores de los píxeles, basado en un mapa de movimiento y patrones estacionarios. Se utiliza un mapa de redes neuronales los cuales permiten entrenar el modelo y cada nodo tiene una función lineal de combinación de pesos de las entradas. Cada nodo puede ser representado por un vector de pesos. Cuando llega un nuevo frame que genera un patrón de pesos se mapea con la red neuronal y posteriormente se actualiza cada nodo para obtener el peso del nuevo patrón. Para cada píxel de color tenemos un mapa neuronal formado por vectores de pesos, el conjunto de vectores de pesos genera el modelo de fondo y usado para la obtención del frente.

La inicialización del modelo se realiza con la primera imagen, y se generan los respectivos vectores de pesos. Una vez se tienen los vectores de pesos se genera el mapa de redes neuronales. Una vez que llegan nuevas frames se generan los vectores de pesos y se comparan con el modelo que tenemos, si hay coincidencia quiere decir que pertenece al fondo el píxel, si no ocurre esto sería píxel de frente.

FUZZY ADAPATIVE SOM [91]: Se propone un SOBS (*self-organizing background subtraction*), donde se implementa una aproximación para la detección de objetos en movimiento basado en una modelo de fondo generado automáticamente. Generado por un método de auto organización con un conocimiento previo sobre los patrones involucrados.

El fin de este método es la mejora y ampliación de un algoritmo SOBS intro-

duciendo una coherencia espacial, además de tener en cuenta la incertidumbre en el modelo de fondo. Es muy útil pensar en coherencia espacial como la diferencia local en la intensidad entre píxeles contiguos. Por tanto los píxeles vecinos que presenten una pequeña diferencia de intensidad serán coherentes, si fuese al contrario serían incoherentes. Con este hecho se consigue una mayor robustez a falsas detecciones.

Se propone un método de aproximación a la variante del algoritmo SOBS, en el cual durante la actualización del modelo de fondo, un mecanismo automático y que depende de los datos para reforzar aún más la contribución de los píxeles que pertenecen al fondo.

Las consideraciones expuestas dan pie a la formulación del algoritmo como un procedimiento basado en unas reglas fuzzy, donde fuzzy funciones son ejecutadas a nivel de píxel en la base de la fase de sustracción del fondo y la combinación de las reglas producidas. Las mejoras producen una mayor precisión y nitidez en la detección de objetos en movimientos.

Apéndice C

Ejemplo de Configuración Segmentador Frente-Fondo

En esta sección se presenta un ejemplo del proceso de configuración de los parámetros de un segmentador Frente-Fondo para un vídeo, solo se muestra este ejemplo por simplicidad, espacio y porque no se considera una parte principal de este trabajo solo se muestra este ejemplo, dado que se ha realizado la configuración para cada uno de los vídeos y algoritmos utilizados en este trabajo. El principal objetivo es mejorar los resultados de ambos segmentadores, por lo que se ha intentado partir de los mejores resultados posibles para que las mejoras introducidas fueran realmente relevantes, pero sin ser totalmente exhaustivos porque no se considera una parte principal de este trabajo.

En este caso el vídeo utilizado del *dataset* es AVSS_S1 y el algoritmo usado es SimpleGaussian. En este caso el algoritmo SimpleGaussian tiene tres parámetros para la configuración del mismo. Hay dos modos de obtener la configuración deseada, el primero es subjetivo, basado en la imagen mostrada por el segmentador durante el proceso. El segundo es en función de las puntuaciones ya desarrolladas en el capítulo 4, por lo tanto, estas son objetivas.

En nuestro caso, se han usado ambas pero si bien es cierto, ha tenido mayor influencia las medidas objetivas. En este caso el segmentador tiene un parámetro *learningRate*, el cual nos hace referencia a cuanto porcentaje de información del modelo de fondo y de la imagen actual del vídeo tomamos. El siguiente parámetro que encontramos disponible para su configuración es *noise Variance* y por último tenemos *sensitivity*, el cual hace referencia al umbral de decisión para cuando determinamos si pertenece o no al fondo un píxel o bloque dependiendo al nivel que se trabaje.

Partimos de los parámetros que vienen por defecto escogidos por el desarrolla-

del algoritmo, la figura C.1 nos muestra los parámetros por defecto que tiene el algoritmo.

Sensitivity	noiseVariance	learningRate
66	162	18

Tabla C.1: Parámetros originales Segmentador SimpleGaussian

En la Figura C.2 se observan los resultados obtenidos en la evaluación del segmentador original.

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.996	0.1941	0.8889	0.8783	0.9392	0.3152	1.2543

Tabla C.2: Puntuaciones Segmentador Original

Primero se procede a la modificación del parámetro *learningRate* con los siguientes valores, en la Figura C.3, se pueden observar los valores de los parámetros escogidos y los consiguientes resultados de la evaluación para cada valor.

learningRate=100;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9933	0.1524	0.8495	0.8442	0.9149	0.2548	1.1697

learningRate=10;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9968	0.1842	0.8819	0.8853	0.9354	0.3019	1.2373

learningRate=26;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9966	0.1883	0.8742	0.8782	0.9301	0.3069	1.2369

Tabla C.3: Puntuaciones Parámetro LearningRate modificado

En este caso, como el parámetro original genera puntuaciones más altas nos quedamos con este valor y comenzamos a la modificación del parámetro *noise Variance*, pudiéndose observar en la Figura C.4, los resultados obtenidos en la evaluación para cada variación del parámetro.

noiseVariance=100;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9971	0.1491	0.8413	0.8974	0.9112	0.2541	1.1653

noiseVariance=250;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9963	0.2220	0.9088	0.8690	0.9504	0.3485	1.2989

Tabla C.4: Puntuaciones Parámetro NoiseVariance modificado

Para este parámetro es mejor un valor superior al original, por lo que la configuración actual del segmentador sería *Sensitivity*=66, *noise Variance*=250 y *learningRate*=18. Por lo tanto, procedemos a modificar el parámetro *Sensitivity*, pero en este caso solo subiremos su valor, puesto que en las anteriores ejecuciones se ha observado que es un valor bajo porque clasifica como frente muchas zonas que en ningún momento lo son, observándose los resultados obtenidos en la Figura C.5.

Sensitivity=120;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9947	0.4098	0.9645	0.8143	0.9793	0.5372	1.5165

Sensitivity=160;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9935	0.4997	0.9765	0.7804	0.9849	0.6013	1.5861

Tabla C.5: Puntuaciones Parámetro Sensitivity modificado

Claramente el segmentador sufre una gran mejora con respecto a la configuración original, pero dado que el parámetro *learningRate* no ha mejorado con ningún cambio se ha decidido aumentar su valor tras la mejora producida por el resto de parámetros, pudiendo apreciar los resultados de la última modificación de los parámetros en la Figura C.6.

learningRate=25;

PF	PFR	RF	RFR	F1F	F1FR	F1F+F1FR
0.9930	0.5276	0.9793	0.7663	0.9861	0.6173	1.6033

Tabla C.6: Puntuaciones Parámetro configuración final

El proceso realizado para este segmentador, se ha realizado para cada uno de los segmentadores y para cada vídeo por independiente, dado que las características del

vídeo influyen mucho en los resultados obtenidos.

Apéndice D

Resultados Adicionales a la Evaluación

En este apéndice se muestran todos los resultados obtenidos en la evaluación de los segmentadores Frente-Fondo y las dos aproximaciones realizadas en este proyecto, que por motivos de espacio no se han incluido en el capítulo 4. Se pueden observar los datos de las puntuaciones específicas para cada vídeo del estudio de los segmentadores Frente-Fondo y de las dos aproximaciones de reconstrucción realizadas en este proyecto. Los vídeos se han decidido nombrar de la siguiente manera, puesto que por cuestión de espacio era necesario.

- AV es AV_S1.
- BM es BaileMulti.
- BS es BaileUni.
- BS es BusStation.
- CO es Corridor.
- CU es Cubicle.
- S6 es DaSide0811-S6.
- S7 es DaSide0811-S7.
- OF es Office.
- PK es Park.
- PD es Pedestrians.

- P9 es PETS2009.
- P1 es PETS_S1.
- TE es Tennis1.
- TR es Train56.

D.1. Segmentadores Frente-Fondo

En esta sección se muestran los resultados de la evaluación de los segmentadores Frente-Fondo más detalladamente, puesto que en la Figura D.1 pueden observar los resultados para la precisión de fondo para cada vídeo y algoritmo. En las Figuras D.2, D.3, D.4 D.5 y D.6, se pueden observar la precisión de frente, el recall de fondo y frente y por último el F1 Score de fondo y frente respectivamente.

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.983	0.939	0.971	0.987	0.985	0.994	0.928	0.928	0.966	0.996	0.999	0.993	0.993	0.954	0.937
DPAdaptiveMedianBGS	0.987	0.954	0.968	0.986	0.991	0.996	0.952	0.979	0.979	0.992	0.999	0.994	0.995	0.970	0.939
DPGrimsonGMMBGS	0.994	0.983	0.982	0.997	0.997	0.998	0.954	0.977	0.998	0.995	0.999	0.995	0.997	0.978	0.945
DPEigenbackgroundBGS	0.981	0.928	0.960	0.995	0.997	0.996	0.963	0.910	0.980	0.995	0.999	0.995	0.994	0.953	0.943
DPMeanBGS	0.972	0.947	0.946	0.989	0.994	0.995	0.910	0.925	0.995	0.993	0.999	0.991	0.993	0.937	0.932
DPPratiMediodBGS	0.985	0.959	0.958	0.974	0.986	0.995	0.975	0.939	0.952	0.994	0.998	0.996	0.992	0.959	0.944
DPWrenGABGS	0.992	0.947	0.968	0.991	0.991	0.994	0.932	0.932	0.993	0.992	0.999	0.992	0.995	0.958	0.935
DPZivkovicAGMMBGS	0.978	0.925	0.959	0.992	0.996	0.996	0.951	0.911	0.976	0.994	0.999	0.995	0.994	0.951	0.936
FuzzyChoquetIntegral	0.985	0.942	0.945	0.973	0.990	0.997	0.956	0.967	0.951	0.991	0.997	0.990	0.993	0.941	0.921
FuzzySugenIntegral	0.987	0.935	0.938	0.973	0.990	0.996	0.955	0.964	0.990	0.990	0.997	0.988	0.992	0.937	0.922
GMG	0.993	0.933	0.933	0.992	0.983	0.996	0.985	0.968	0.970	0.998	0.999	0.994	0.990	0.978	0.946
KDE	0.992	0.984	0.990	0.994	0.980	0.999	0.972	0.974	0.991	0.990	0.999	0.993	0.997	0.987	0.931
LBAdaptiveSOM	0.990	0.963	0.966	0.993	0.988	0.994	0.955	0.948	0.997	0.992	0.999	0.994	0.995	0.975	0.936
LBFuzzyAdaptiveSOM	0.992	0.972	0.975	0.995	0.993	0.995	0.959	0.958	0.998	0.993	0.999	0.995	0.996	0.969	0.939
LBFuzzyGaussian	0.991	0.928	0.943	0.993	0.992	0.996	0.909	0.919	0.994	0.994	0.999	0.991	0.992	0.950	0.929
LBMixtureOfGaussians	0.984	0.932	0.935	0.987	0.995	0.996	0.923	0.929	0.971	0.993	0.999	0.990	0.991	0.940	0.931
LBSimpleGaussian	0.990	0.946	0.951	0.991	0.993	0.996	0.882	0.933	0.998	0.994	0.999	0.992	0.993	0.940	0.931
MOGV1BGS	0.987	0.923	0.908	0.985	0.985	0.992	0.887	0.929	0.968	0.992	0.999	0.991	0.991	0.929	0.925
MOGV2BGS	0.988	0.963	0.962	0.991	0.979	0.990	0.929	0.940	0.971	0.996	0.999	0.991	0.994	0.962	0.927
MultiLayerBGS	0.987	0.948	0.939	0.985	0.975	0.993	0.964	0.973	0.960	0.994	0.999	0.995	0.996	0.967	0.934
SGMM-SOD	0.995	0.987	0.989	0.996	0.992	0.998	0.979	0.990	0.998	0.995	0.999	0.997	0.998	0.987	0.944
SigmaDeltaBGS	0.983	0.936	0.906	0.987	0.983	0.995	0.898	0.929	0.966	0.995	0.999	0.991	0.990	0.943	0.926
T2FGMM_UM	0.970	0.952	0.949	0.975	0.986	0.987	0.886	0.918	0.951	0.993	0.995	0.978	0.978	0.957	0.914
T2FGMM_UV	0.994	0.968	0.974	0.990	0.996	0.996	0.943	0.961	0.998	0.994	0.999	0.985	0.996	0.940	0.940
T2FMRF_UM	0.968	0.931	0.917	0.973	0.978	0.987	0.888	0.920	0.951	0.992	0.995	0.984	0.976	0.927	0.919
T2FMRF_UV	0.990	0.945	0.940	0.988	0.978	0.986	0.970	0.932	0.994	0.993	0.995	0.991	0.991	0.932	0.931
VuMeter	0.995	0.891	0.917	0.976	0.987	0.994	0.955	0.963	0.953	0.990	0.997	0.991	0.992	0.958	0.935

Tabla D.1: Precisión Fondo Segmentadores para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.678	0.447	0.716	0.347	0.378	0.125	0.648	0.553	0.344	0.192	0.135	0.560	0.443	0.312	0.397
DPAdaptiveMedianBGS	0.693	0.546	0.948	0.613	0.673	0.150	0.483	0.341	0.891	0.450	0.399	0.608	0.687	0.506	0.361
DPGrimsonGMMBGS	0.537	0.317	0.915	0.150	0.289	0.047	0.440	0.292	0.180	0.277	0.126	0.424	0.373	0.250	0.254
DPEigenbackgroundBGS	0.730	0.680	0.952	0.297	0.331	0.094	0.631	0.475	0.395	0.333	0.335	0.437	0.694	0.345	0.403
DPMeanBGS	0.582	0.438	0.958	0.388	0.453	0.087	0.445	0.555	0.814	0.435	0.339	0.499	0.735	0.372	0.319
DPPratiMediodBGS	0.590	0.346	0.714	0.703	0.459	0.148	0.562	0.553	0.728	0.260	0.222	0.476	0.461	0.487	0.423
DPWrenGABGS	0.583	0.502	0.923	0.323	0.641	0.128	0.556	0.466	0.536	0.407	0.578	0.524	0.579	0.361	0.434
DPZivkovicAGMMBGS	0.643	0.685	0.952	0.358	0.428	0.096	0.566	0.481	0.431	0.373	0.464	0.619	0.698	0.347	0.516
FuzzyChoquetIntegral	0.408	0.496	0.954	0.999	0.454	0.096	0.701	0.444	0.999	0.325	0.840	0.703	0.604	0.631	0.505
FuzzySugenoIntegral	0.338	0.523	0.961	0.999	0.619	0.109	0.702	0.564	0.597	0.399	0.843	0.723	0.672	0.648	0.370
GMG	0.496	0.408	0.861	0.401	0.263	0.152	0.675	0.448	0.427	0.297	0.178	0.622	0.547	0.731	0.495
KDE	0.247	0.314	0.342	0.346	0.499	0.121	0.421	0.297	0.393	0.254	0.311	0.473	0.134	0.453	0.276
LBAdaptiveSOM	0.710	0.438	0.947	0.391	0.742	0.480	0.442	0.360	0.638	0.442	0.415	0.467	0.674	0.428	0.341
LBFuzzyAdaptiveSOM	0.656	0.371	0.939	0.280	0.646	0.165	0.423	0.355	0.458	0.423	0.350	0.445	0.555	0.415	0.308
LBFuzzyGaussian	0.686	0.605	0.959	0.313	0.580	0.251	0.631	0.553	0.530	0.415	0.394	0.741	0.747	0.353	0.490
LBMixtureOfGaussians	0.710	0.530	0.961	0.390	0.477	0.095	0.689	0.534	0.434	0.419	0.525	0.757	0.765	0.368	0.331
LBSimpleGaussian	0.415	0.496	0.955	0.285	0.575	0.120	0.606	0.465	0.354	0.355	0.281	0.622	0.712	0.499	0.499
MOGV1BGS	0.626	0.539	0.962	0.600	0.810	0.272	0.547	0.331	0.799	0.453	0.690	0.743	0.812	0.653	0.674
MOGV2BGS	0.621	0.443	0.854	0.305	0.189	0.166	0.719	0.402	0.399	0.232	0.216	0.745	0.690	0.590	0.595
MultiLayerBGS	0.739	0.569	0.896	0.691	0.774	0.664	0.772	0.481	0.786	0.404	0.719	0.707	0.720	0.789	0.655
SGMM-SOD	0.553	0.371	0.970	0.585	0.806	0.400	0.498	0.345	0.826	0.432	0.612	0.923	0.665	0.518	0.373
SigmaDeltaBGS	0.736	0.551	0.936	0.427	0.581	0.194	0.607	0.472	0.383	0.305	0.282	0.757	0.844	0.587	0.618
T2FGMM_UM	0.592	0.426	0.958	0.905	0.521	0.220	0.396	0.209	0.586	0.324	0.345	0.591	0.489	0.471	0.520
T2FGMM_UV	0.564	0.415	0.944	0.366	0.409	0.101	0.557	0.344	0.356	0.310	0.209	0.562	0.587	0.554	0.389
T2FMRF_UM	0.667	0.491	0.961	0.828	0.369	0.238	0.367	0.304	0.740	0.425	0.526	0.143	0.586	0.603	0.265
T2FMRF_UV	0.667	0.457	0.959	0.470	0.369	0.151	0.383	0.240	0.855	0.293	0.307	0.498	0.754	0.583	0.496
VuMeter	0.424	0.754	0.921	0.708	0.688	0.474	0.614	0.442	0.863	0.372	0.757	0.671	0.534	0.632	0.510

Tabla D.2: Precisión Frente Segmentadores para cada Vídeo

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.993	0.906	0.945	0.977	0.977	0.973	0.955	0.973	0.978	0.981	0.976	0.987	0.969	0.875	0.939
DPAdaptiveMedianBGS	0.992	0.926	0.995	0.994	0.996	0.967	0.878	0.827	0.995	0.996	0.995	0.988	0.989	0.936	0.928
DPGrimsonGMMBGS	0.974	0.734	0.991	0.893	0.898	0.756	0.848	0.778	0.822	0.990	0.969	0.975	0.968	0.729	0.850
DPEigenbackgroundBGS	0.995	0.962	0.996	0.965	0.975	0.878	0.928	0.988	0.930	0.990	0.990	0.976	0.990	0.897	0.932
DPMeanBGS	0.997	0.891	0.997	0.979	0.989	0.912	0.915	0.977	0.991	0.995	0.993	0.984	0.993	0.945	0.929
DPPratiMediodBGS	0.985	0.812	0.948	0.999	0.978	0.975	0.896	0.964	0.999	0.987	0.987	0.977	0.974	0.935	0.938
DPWrenGABGS	0.985	0.916	0.994	0.976	0.995	0.962	0.929	0.959	0.983	0.995	0.995	0.985	0.982	0.899	0.949
DPZivkovicAGMMBGS	0.994	0.966	0.996	0.980	0.986	0.925	0.905	0.986	0.953	0.992	0.992	0.988	0.990	0.903	0.967
FuzzyChoquetIntegral	0.975	0.918	0.997	0.999	0.988	0.915	0.952	0.908	0.999	0.997	0.999	0.994	0.992	0.980	0.983
FuzzySugenolIntegral	0.964	0.935	0.997	0.999	0.995	0.941	0.953	0.946	0.992	0.997	0.999	0.995	0.994	0.984	0.968
GMG	0.977	0.894	0.989	0.978	0.991	0.975	0.928	0.909	0.981	0.980	0.981	0.989	0.980	0.971	0.951
KDE	0.931	0.723	0.758	0.969	0.998	0.928	0.816	0.804	0.944	0.996	0.994	0.981	0.865	0.897	0.913
LBAdaptiveSOM	0.991	0.867	0.995	0.983	0.998	0.993	0.849	0.903	0.988	0.996	0.995	0.979	0.988	0.903	0.927
LBFuzzyAdaptiveSOM	0.988	0.808	0.994	0.963	0.995	0.976	0.828	0.879	0.965	0.995	0.995	0.977	0.982	0.906	0.905
LBFuzzyGaussian	0.991	0.957	0.997	0.974	0.995	0.987	0.965	0.983	0.984	0.994	0.995	0.995	0.994	0.908	0.969
LBMixtureOfGaussians	0.994	0.939	0.997	0.988	0.990	0.925	0.965	0.971	0.981	0.994	0.995	0.996	0.995	0.937	0.937
LBSimpleGaussian	0.966	0.916	0.996	0.965	0.993	0.954	0.983	0.958	0.921	0.992	0.993	0.990	0.991	0.968	0.971
MOGV1BGS	0.989	0.953	0.999	0.994	0.999	0.988	0.971	0.923	0.995	0.996	0.996	0.995	0.995	0.992	0.991
MOGV2BGS	0.984	0.852	0.986	0.965	0.991	0.992	0.963	0.936	0.958	0.982	0.987	0.995	0.989	0.960	0.983
MultiLayerBGS	0.994	0.928	0.991	0.993	0.997	0.996	0.965	0.917	0.995	0.992	0.994	0.992	0.992	0.984	0.984
SGMM-SOD	0.979	0.778	0.998	0.985	0.994	0.991	0.860	0.810	0.987	0.994	0.993	0.998	0.983	0.919	0.919
SigmaDeltaBGS	0.995	0.945	0.997	0.988	0.992	0.990	0.968	0.970	0.984	0.991	0.993	0.995	0.996	0.975	0.986
T2FGMM_UM	0.999	0.878	0.997	0.999	0.993	0.992	0.939	0.889	0.999	0.994	0.997	0.999	0.997	0.941	0.994
T2FGMM_UV	0.983	0.841	0.994	0.984	0.985	0.925	0.913	0.873	0.940	0.992	0.987	0.993	0.982	0.972	0.929
T2FMRF_UM	0.999	0.932	0.998	0.999	0.997	0.995	0.921	0.940	0.999	0.995	0.998	0.948	0.999	0.989	0.955
T2FMRF_UV	0.990	0.903	0.997	0.988	0.997	0.995	0.788	0.858	0.991	0.991	0.997	0.985	0.994	0.983	0.969
VuMeter	0.966	0.978	0.986	0.996	0.992	0.990	0.904	0.912	0.998	0.988	0.994	0.993	0.969	0.943	0.963

Tabla D.3: Recall Fondo Segmentadores para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.473	0.571	0.791	0.676	0.412	0.810	0.528	0.303	0.544	0.836	0.938	0.705	0.609	0.548	0.367
DPAdaptiveMedianBGS	0.588	0.670	0.759	0.650	0.741	0.859	0.721	0.837	0.713	0.688	0.917	0.749	0.753	0.685	0.397
DPGrimsonGMMBGS	0.800	0.907	0.871	0.927	0.964	0.944	0.743	0.834	0.979	0.801	0.955	0.775	0.857	0.814	0.505
DPEigenbackgrounBGS	0.464	0.465	0.699	0.878	0.967	0.879	0.777	0.099	0.750	0.802	0.960	0.794	0.681	0.534	0.440
DPMeanBGS	0.166	0.634	0.584	0.737	0.917	0.851	0.424	0.263	0.926	0.708	0.899	0.646	0.617	0.329	0.327
DPPratiMediodBGS	0.589	0.732	0.675	0.358	0.445	0.851	0.856	0.417	0.308	0.765	0.842	0.830	0.632	0.572	0.445
DPWrenGABGS	0.729	0.628	0.770	0.780	0.787	0.816	0.569	0.340	0.913	0.675	0.878	0.684	0.767	0.582	0.339
DPZivkovicAGMMBGS	0.399	0.438	0.693	0.819	0.957	0.850	0.704	0.119	0.694	0.759	0.951	0.778	0.671	0.511	0.337
FuzzyChoquetIntegral	0.561	0.588	0.578	0.322	0.826	0.914	0.720	0.721	0.287	0.623	0.740	0.563	0.650	0.351	0.162
FuzzySugenoIntegral	0.604	0.521	0.522	0.322	0.822	0.881	0.718	0.673	0.870	0.603	0.706	0.501	0.624	0.302	0.188
GMG	0.796	0.520	0.486	0.781	0.348	0.889	0.908	0.707	0.595	0.905	0.994	0.754	0.596	0.769	0.456
KDE	0.752	0.916	0.944	0.872	0.294	0.969	0.850	0.799	0.887	0.591	0.897	0.700	0.891	0.873	0.330
LBAdaptiveSOM	0.691	0.760	0.749	0.822	0.492	0.779	0.746	0.521	0.958	0.690	0.901	0.758	0.726	0.747	0.373
LBFuzzyAdaptiveSOM	0.737	0.831	0.814	0.884	0.753	0.843	0.778	0.628	0.973	0.728	0.915	0.774	0.778	0.684	0.415
LBFuzzyGaussian	0.704	0.464	0.560	0.827	0.691	0.851	0.387	0.201	0.926	0.744	0.916	0.632	0.593	0.497	0.256
LBMixtureOfGaussians	0.520	0.507	0.499	0.667	0.933	0.860	0.487	0.311	0.620	0.731	0.912	0.563	0.550	0.367	0.308
LBSimpleGaussian	0.708	0.619	0.630	0.778	0.816	0.854	0.164	0.349	0.971	0.773	0.922	0.658	0.658	0.339	0.277
MOGV1BGS	0.593	0.422	0.252	0.626	0.558	0.734	0.214	0.348	0.572	0.695	0.892	0.639	0.530	0.195	0.194
MOGV2BGS	0.635	0.767	0.729	0.790	0.303	0.712	0.536	0.408	0.612	0.839	0.950	0.637	0.740	0.590	0.230
MultiLayerBGS	0.622	0.615	0.520	0.626	0.226	0.800	0.766	0.740	0.448	0.780	0.930	0.778	0.817	0.636	0.301
SGMM-SOD	0.856	0.928	0.921	0.908	0.530	0.930	0.885	0.924	0.974	0.802	0.982	0.886	0.910	0.865	0.455
SigmaDeltaBGS	0.500	0.519	0.254	0.661	0.366	0.815	0.300	0.276	0.546	0.792	0.947	0.608	0.469	0.379	0.207
T2FGMM_UM	0.130	0.678	0.611	0.365	0.556	0.669	0.238	0.275	0.288	0.663	0.572	0.047	0.087	0.551	0.072
T2FGMM_UV	0.789	0.797	0.810	0.771	0.959	0.864	0.652	0.639	0.971	0.782	0.940	0.392	0.797	0.352	0.408
T2FMRF_UM	0.001	0.494	0.340	0.323	0.295	0.667	0.266	0.249	0.287	0.630	0.550	0.360	0.022	0.179	0.162
T2FMRF_UV	0.665	0.611	0.537	0.707	0.295	0.627	0.841	0.425	0.921	0.716	0.571	0.613	0.565	0.245	0.280
VuMeter	0.816	0.168	0.386	0.432	0.392	0.770	0.715	0.643	0.338	0.737	0.863	0.622	0.583	0.625	0.327

Tabla D.4: Recall Frente Segmentadores para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.988	0.922	0.957	0.982	0.980	0.983	0.941	0.950	0.972	0.988	0.987	0.99	0.981	0.912	0.938
DPAdaptiveMedianBGS	0.990	0.940	0.981	0.990	0.993	0.981	0.913	0.896	0.987	0.994	0.997	0.991	0.992	0.952	0.934
DPGrimsonGMMBGS	0.984	0.840	0.986	0.941	0.942	0.854	0.897	0.866	0.895	0.993	0.984	0.985	0.982	0.831	0.895
DPEigenbackgroundBGS	0.988	0.944	0.977	0.980	0.986	0.930	0.945	0.947	0.948	0.993	0.995	0.985	0.992	0.924	0.937
DPMeanBGS	0.984	0.918	0.971	0.984	0.991	0.950	0.912	0.950	0.993	0.994	0.996	0.988	0.993	0.941	0.931
DPPratiMediodBGS	0.985	0.878	0.951	0.986	0.982	0.985	0.934	0.951	0.975	0.990	0.993	0.986	0.983	0.946	0.941
DPWrenGABGS	0.988	0.931	0.981	0.983	0.993	0.978	0.930	0.945	0.988	0.994	0.997	0.989	0.989	0.927	0.942
DPZivkovicAGMMBGS	0.986	0.944	0.977	0.986	0.991	0.958	0.927	0.947	0.962	0.993	0.995	0.992	0.992	0.926	0.951
FuzzyChoquetIntegral	0.980	0.930	0.970	0.986	0.989	0.954	0.954	0.937	0.975	0.994	0.998	0.992	0.992	0.960	0.951
FuzzySugenIntegral	0.975	0.934	0.967	0.986	0.992	0.968	0.954	0.955	0.991	0.994	0.998	0.992	0.993	0.960	0.945
GMG	0.985	0.911	0.960	0.985	0.987	0.985	0.955	0.938	0.975	0.989	0.990	0.992	0.985	0.975	0.948
KDE	0.960	0.833	0.858	0.981	0.988	0.962	0.887	0.880	0.967	0.993	0.997	0.987	0.926	0.940	0.922
LBAdaptiveSOM	0.991	0.912	0.980	0.988	0.992	0.994	0.898	0.925	0.993	0.994	0.997	0.987	0.991	0.937	0.932
LBFuzzyAdaptiveSOM	0.990	0.882	0.984	0.979	0.994	0.986	0.888	0.917	0.981	0.994	0.997	0.986	0.989	0.936	0.922
LBFuzzyGaussian	0.991	0.942	0.969	0.983	0.993	0.991	0.936	0.950	0.989	0.994	0.997	0.993	0.993	0.928	0.948
LBMixtureOfGaussians	0.989	0.935	0.965	0.987	0.992	0.958	0.943	0.949	0.976	0.994	0.997	0.993	0.993	0.938	0.934
LBSimpleGaussian	0.978	0.931	0.973	0.978	0.993	0.974	0.930	0.945	0.954	0.993	0.996	0.991	0.992	0.954	0.950
MOGV1BGS	0.988	0.938	0.951	0.989	0.992	0.990	0.927	0.926	0.981	0.994	0.997	0.993	0.993	0.959	0.957
MOGV2BGS	0.986	0.901	0.974	0.978	0.985	0.991	0.946	0.938	0.960	0.989	0.993	0.993	0.992	0.961	0.954
MultiLayerBGS	0.990	0.936	0.964	0.989	0.985	0.994	0.964	0.944	0.977	0.993	0.996	0.993	0.994	0.975	0.958
SGMM-SOD	0.987	0.869	0.993	0.991	0.993	0.995	0.916	0.891	0.993	0.994	0.996	0.998	0.991	0.952	0.932
SigmaDeltaBGS	0.989	0.940	0.949	0.987	0.987	0.992	0.931	0.949	0.975	0.993	0.996	0.993	0.993	0.959	0.955
T2FGMM_UM	0.984	0.914	0.972	0.987	0.989	0.989	0.911	0.903	0.975	0.993	0.996	0.988	0.987	0.949	0.953
T2FGMM_UV	0.988	0.899	0.984	0.987	0.991	0.958	0.928	0.915	0.967	0.993	0.993	0.989	0.989	0.956	0.935
T2FMRF_UM	0.984	0.932	0.956	0.986	0.987	0.991	0.903	0.929	0.975	0.994	0.997	0.966	0.988	0.956	0.937
T2FMRF_UV	0.990	0.923	0.968	0.988	0.987	0.991	0.869	0.893	0.993	0.992	0.996	0.988	0.993	0.957	0.950
VuMeter	0.980	0.931	0.948	0.986	0.989	0.992	0.926	0.936	0.974	0.989	0.995	0.992	0.980	0.948	0.948

Tabla D.5: F1Score Fondo Segmentadores para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.547	0.492	0.740	0.629	0.539	0.723	0.580	0.391	0.528	0.734	0.633	0.621	0.470	0.385	0.376
DPAdaptiveMedianBGS	0.629	0.591	0.842	0.700	0.682	0.751	0.574	0.469	0.782	0.739	0.808	0.670	0.704	0.573	0.377
DPGrimsonGMMBGS	0.628	0.462	0.890	0.560	0.393	0.644	0.546	0.418	0.565	0.780	0.625	0.548	0.512	0.366	0.337
DPEigenbackgroundBGS	0.549	0.513	0.805	0.702	0.471	0.691	0.694	0.164	0.652	0.780	0.759	0.564	0.670	0.407	0.416
DPMeanBGS	0.236	0.509	0.724	0.682	0.568	0.690	0.428	0.356	0.900	0.751	0.768	0.563	0.650	0.340	0.323
DPPratiMediodBGS	0.582	0.456	0.650	0.597	0.539	0.744	0.676	0.474	0.504	0.725	0.667	0.598	0.499	0.503	0.433
DPWrenGABGS	0.640	0.546	0.833	0.697	0.696	0.725	0.560	0.391	0.854	0.725	0.796	0.593	0.647	0.437	0.371
DPZivkovicAGMMBGS	0.465	0.498	0.801	0.732	0.564	0.696	0.617	0.190	0.628	0.771	0.773	0.687	0.666	0.403	0.402
FuzzyChoquetIntegral	0.466	0.525	0.719	0.996	0.518	0.701	0.709	0.541	0.998	0.677	0.779	0.625	0.624	0.442	0.241
FuzzySugenoIntegral	0.427	0.509	0.674	0.999	0.660	0.712	0.708	0.612	0.875	0.657	0.756	0.592	0.645	0.404	0.247
GMM	0.605	0.438	0.584	0.698	0.667	0.766	0.771	0.545	0.582	0.785	0.682	0.681	0.530	0.743	0.470
KDE	0.360	0.458	0.496	0.740	0.622	0.736	0.559	0.425	0.752	0.625	0.796	0.560	0.225	0.592	0.300
LBAdaptiveSOM	0.695	0.543	0.835	0.754	0.715	0.794	0.549	0.417	0.900	0.739	0.813	0.578	0.682	0.537	0.355
LBFuzzyAdaptiveSOM	0.687	0.503	0.871	0.694	0.664	0.760	0.541	0.443	0.823	0.765	0.807	0.565	0.637	0.508	0.353
LBFuzzyGaussian	0.689	0.500	0.705	0.705	0.652	0.808	0.478	0.291	0.860	0.775	0.808	0.681	0.637	0.401	0.325
LBmixtureOfGaussians	0.590	0.501	0.654	0.661	0.611	0.697	0.567	0.392	0.613	0.766	0.810	0.645	0.609	0.358	0.319
LBsimpleGaussian	0.517	0.540	0.757	0.660	0.668	0.724	0.257	0.396	0.732	0.785	0.783	0.639	0.663	0.398	0.351
MOGV1BGS	0.603	0.465	0.393	0.677	0.673	0.733	0.304	0.323	0.616	0.743	0.811	0.687	0.594	0.287	0.298
MOGV2BGS	0.599	0.535	0.770	0.678	0.507	0.717	0.605	0.397	0.573	0.764	0.726	0.685	0.697	0.583	0.325
MultiLayerBGS	0.669	0.561	0.626	0.674	0.701	0.804	0.767	0.579	0.628	0.790	0.796	0.741	0.762	0.689	0.408
SGMM-SOD	0.661	0.517	0.939	0.816	0.802	0.863	0.633	0.485	0.903	0.816	0.813	0.903	0.751	0.638	0.405
SigmaDeltaBGS	0.591	0.521	0.383	0.656	0.610	0.777	0.397	0.337	0.550	0.786	0.787	0.674	0.539	0.454	0.301
T2FGMM_UM	0.168	0.514	0.744	0.758	0.508	0.698	0.280	0.230	0.479	0.708	0.594	0.087	0.142	0.499	0.119
T2FGMM_UV	0.651	0.528	0.871	0.720	0.542	0.700	0.594	0.442	0.746	0.781	0.720	0.455	0.664	0.416	0.391
T2FMRF_UM	0.835	0.484	0.498	0.812	0.533	0.696	0.286	0.266	0.781	0.688	0.584	0.204	0.041	0.268	0.201
T2FMRF_UV	0.658	0.513	0.686	0.705	0.533	0.659	0.523	0.300	0.900	0.738	0.590	0.550	0.622	0.334	0.353
VuMeter	0.555	0.219	0.501	0.648	0.702	0.783	0.703	0.518	0.673	0.755	0.814	0.643	0.482	0.603	0.381

Tabla D.6: F1Score Frente Segmentadores para cada Video

D.2. Resultados Segunda Aproximación

En esta sección se muestran los resultados de la evaluación de la segunda aproximación realizada en este proyecto más detalladamente, la cual busca la mejora de la segmentación Persona-Fondo usando la información de la segmentación Frente-Fondo. Puesto que, en la Figura D.1 pueden observar los resultados para la precisión de fondo para cada vídeo y algoritmo. En las Figuras D.2, D.3 D.4, D.5 y D.6, se pueden observar la precisión de frente, el recall de fondo y frente y por último el F1 Score de fondo y frente respectivamente.

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.999	0.995	0.997	0.998	0.984	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.996	0.996
DPAdaptiveMedianBGS	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
DPGrimsonGMMBGS	0.999	0.995	0.997	0.998	0.982	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.996	0.996
DPEigenbackgroundBGS	0.999	0.995	0.997	0.998	0.983	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.996	0.996
DPMeanBGS	0.999	0.995	0.997	0.998	0.983	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
DPPratiMediodBGS	0.999	0.995	0.997	0.992	0.984	0.999	0.999	0.999	0.987	0.999	0.999	0.999	0.999	0.997	0.996
DPWrenGABGS	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
DPZivkovicAGMMBGS	0.999	0.995	0.997	0.998	0.984	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
FuzzyChoquetIntegral	0.999	0.996	0.997	0.973	0.984	0.999	0.999	0.999	0.951	0.999	0.999	0.999	0.999	0.997	0.996
FuzzySugenoIntegral	0.999	0.996	0.997	0.973	0.984	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
GMG	0.999	0.996	0.998	0.998	0.982	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
KDE	0.999	0.995	0.997	0.998	0.983	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
LBAdaptiveSOM	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
LBFuzzyAdaptiveSOM	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
LBFuzzyGaussian	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
LBMixtureOfGaussians	0.999	0.995	0.997	0.998	0.984	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
LBSimpleGaussian	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
MOGV1BGS	0.999	0.995	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
MOGV2BGS	0.999	0.995	0.997	0.998	0.980	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
MultiLayerBGS	0.999	0.996	0.998	0.997	0.985	0.999	0.999	0.999	0.989	0.999	0.999	0.999	0.999	0.997	0.996
SGMM-SOD	0.999	0.996	0.997	0.998	0.985	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
SigmaDeltaBGS	0.999	0.995	0.998	0.998	0.984	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
T2FGMM_UM	0.999	0.995	0.997	0.988	0.985	0.999	0.999	0.999	0.991	0.999	0.999	0.999	0.999	0.996	0.996
T2FGMM_UV	0.999	0.995	0.997	0.998	0.984	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
T2FMRF_UM	0.971	0.995	0.997	0.984	0.983	0.999	0.999	0.999	0.974	0.999	0.999	0.999	0.999	0.997	0.996
T2FMRF_UV	0.999	0.995	0.997	0.998	0.983	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.997	0.996
VuMeter	0.999	0.996	0.998	0.992	0.985	0.999	0.978	0.999	0.977	0.999	0.999	0.999	0.999	0.996	0.996

Tabla D.7: Precisión Fondo Segunda Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.130	0.227	0.221	0.054	0.329	0.021	0.236	0.206	0.079	0.018	0.014	0.198	0.208	0.176	0.199
DPAdaptiveMedianBGS	0.130	0.231	0.250	0.201	0.322	0.022	0.234	0.198	0.421	0.122	0.101	0.185	0.224	0.191	0.196
DPGrimsonGMMBGS	0.117	0.214	0.225	0.048	0.151	0.021	0.233	0.190	0.077	0.018	0.013	0.148	0.103	0.175	0.180
DPEigenbackgroundBGS	0.132	0.234	0.257	0.048	0.149	0.021	0.232	0.209	0.099	0.088	0.111	0.100	0.182	0.179	0.180
DPMeanBGS	0.133	0.225	0.259	0.146	0.167	0.021	0.233	0.211	0.359	0.119	0.082	0.216	0.217	0.187	0.196
DPPratiMediodBGS	0.131	0.228	0.267	0.549	0.388	0.022	0.234	0.208	0.561	0.073	0.066	0.153	0.225	0.194	0.197
DPWrenGABGS	0.117	0.229	0.227	0.049	0.298	0.022	0.230	0.208	0.112	0.099	0.286	0.198	0.152	0.177	0.196
DPZivkovicAGMMBGS	0.121	0.234	0.258	0.049	0.176	0.021	0.233	0.205	0.169	0.099	0.231	0.207	0.190	0.179	0.192
FuzzyChoquetIntegral	0.117	0.236	0.246	0.996	0.190	0.021	0.236	0.198	0.998	0.020	0.445	0.236	0.181	0.204	0.198
FuzzySugenoIntegral	0.117	0.237	0.250	0.999	0.238	0.021	0.236	0.198	0.096	0.095	0.445	0.278	0.179	0.205	0.196
GMG	0.116	0.237	0.280	0.078	0.176	0.027	0.233	0.201	0.108	0.087	0.041	0.289	0.205	0.206	0.199
KDE	0.107	0.214	0.213	0.049	0.250	0.021	0.231	0.188	0.078	0.019	0.025	0.120	0.108	0.180	0.188
LBAdaptiveSOM	0.130	0.226	0.256	0.079	0.440	0.323	0.233	0.205	0.270	0.119	0.105	0.192	0.185	0.190	0.196
LBFuzzyAdaptiveSOM	0.129	0.216	0.230	0.049	0.328	0.022	0.232	0.202	0.084	0.114	0.053	0.168	0.112	0.183	0.193
LBFuzzyGaussian	0.129	0.232	0.259	0.051	0.301	0.047	0.235	0.216	0.084	0.110	0.100	0.244	0.220	0.180	0.198
LBmixtureOfGaussians	0.133	0.231	0.260	0.059	0.173	0.021	0.236	0.206	0.080	0.112	0.227	0.273	0.246	0.185	0.196
LBsimpleGaussian	0.117	0.228	0.258	0.048	0.281	0.021	0.236	0.207	0.078	0.077	0.014	0.168	0.224	0.193	0.196
MOGV1BGS	0.129	0.234	0.268	0.191	0.388	0.114	0.235	0.214	0.421	0.129	0.382	0.273	0.252	0.200	0.201
MOGV2BGS	0.107	0.214	0.220	0.049	0.040	0.022	0.231	0.194	0.084	0.018	0.014	0.203	0.125	0.185	0.183
MultiLayerBGS	0.133	0.238	0.283	0.336	0.696	0.496	0.235	0.205	0.523	0.117	0.443	0.294	0.230	0.211	0.201
SGMM-SOD	0.117	0.237	0.259	0.244	0.568	0.192	0.234	0.196	0.406	0.097	0.339	0.290	0.204	0.205	0.191
SigmaDeltaBGS	0.132	0.233	0.284	0.097	0.455	0.022	0.235	0.214	0.085	0.036	0.017	0.285	0.252	0.196	0.201
T2FGMM_UM	0.133	0.223	0.259	0.692	0.293	0.096	0.230	0.186	0.460	0.046	0.101	0.228	0.168	0.176	0.198
T2FGMM_UV	0.117	0.217	0.243	0.071	0.168	0.021	0.233	0.193	0.080	0.041	0.013	0.122	0.163	0.186	0.184
T2FMRF_UM	0.843	0.229	0.261	0.778	0.191	0.075	0.230	0.186	0.709	0.118	0.240	0.088	0.250	0.193	0.177
T2FMRF_UV	0.132	0.227	0.259	0.159	0.191	0.022	0.229	0.186	0.388	0.043	0.086	0.218	0.248	0.190	0.199
VuMeter	0.119	0.254	0.284	0.446	0.561	0.269	0.340	0.201	0.655	0.109	0.427	0.295	0.198	0.207	0.201

Tabla D.8: Precisión Frente Segunda Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.777	0.542	0.523	0.468	0.746	0.247	0.489	0.594	0.414	0.339	0.432	0.901	0.916	0.560	0.603
DPAdaptiveMedianBGS	0.776	0.551	0.586	0.560	0.775	0.251	0.484	0.571	0.760	0.480	0.502	0.891	0.925	0.601	0.596
DPGrimsonGMMBGS	0.760	0.504	0.532	0.437	0.567	0.246	0.481	0.547	0.405	0.342	0.397	0.857	0.779	0.555	0.551
DPEigenbackgroundBGS	0.779	0.557	0.600	0.441	0.581	0.246	0.477	0.602	0.485	0.452	0.559	0.783	0.895	0.567	0.551
DPMeanBGS	0.778	0.537	0.610	0.511	0.615	0.247	0.483	0.605	0.595	0.475	0.459	0.912	0.922	0.590	0.598
DPPratiMediodBGS	0.777	0.544	0.620	0.820	0.768	0.249	0.485	0.596	0.817	0.442	0.521	0.868	0.927	0.608	0.598
DPWrenGABGS	0.761	0.545	0.536	0.449	0.748	0.249	0.473	0.596	0.532	0.463	0.688	0.898	0.865	0.561	0.595
DPZivkovicAGMMBGS	0.769	0.559	0.602	0.451	0.653	0.247	0.481	0.589	0.591	0.464	0.639	0.904	0.898	0.569	0.584
FuzzyChoquetIntegral	0.761	0.566	0.582	0.998	0.658	0.246	0.489	0.569	0.999	0.391	0.783	0.920	0.901	0.631	0.601
FuzzySugenoIntegral	0.759	0.568	0.589	0.999	0.714	0.246	0.489	0.573	0.534	0.460	0.784	0.936	0.902	0.634	0.597
GMG	0.758	0.566	0.636	0.515	0.661	0.256	0.481	0.581	0.525	0.451	0.515	0.939	0.915	0.635	0.602
KDE	0.738	0.505	0.497	0.448	0.696	0.246	0.476	0.541	0.404	0.352	0.536	0.821	0.802	0.570	0.575
LBAdaptiveSOM	0.776	0.537	0.594	0.476	0.806	0.500	0.481	0.589	0.666	0.479	0.651	0.896	0.897	0.597	0.597
LBFuzzyAdaptiveSOM	0.774	0.510	0.544	0.446	0.772	0.253	0.478	0.580	0.439	0.475	0.461	0.878	0.808	0.579	0.590
LBFuzzyGaussian	0.772	0.554	0.610	0.452	0.749	0.273	0.487	0.620	0.449	0.472	0.540	0.922	0.923	0.570	0.601
LBmixtureOfGaussians	0.780	0.552	0.610	0.475	0.671	0.247	0.489	0.591	0.425	0.474	0.651	0.934	0.929	0.584	0.598
LBsimpleGaussian	0.760	0.544	0.601	0.439	0.734	0.247	0.489	0.594	0.406	0.428	0.417	0.876	0.925	0.604	0.596
MOGV1BGS	0.774	0.561	0.625	0.577	0.806	0.322	0.488	0.614	0.759	0.484	0.751	0.931	0.931	0.620	0.609
MOGV2BGS	0.736	0.506	0.515	0.439	0.486	0.248	0.476	0.562	0.429	0.336	0.413	0.899	0.816	0.584	0.562
MultiLayerBGS	0.781	0.571	0.650	0.746	0.899	0.618	0.487	0.592	0.802	0.474	0.781	0.940	0.927	0.644	0.608
SGMM-SOD	0.764	0.566	0.601	0.689	0.848	0.382	0.483	0.566	0.754	0.459	0.734	0.938	0.910	0.633	0.584
SigmaDeltaBGS	0.780	0.557	0.652	0.536	0.792	0.250	0.488	0.614	0.461	0.408	0.458	0.937	0.931	0.612	0.608
T2FGMM_UM	0.780	0.530	0.608	0.864	0.744	0.306	0.472	0.534	0.782	0.382	0.503	0.918	0.878	0.558	0.602
T2FGMM_UV	0.762	0.514	0.570	0.467	0.625	0.247	0.481	0.558	0.438	0.391	0.405	0.822	0.873	0.587	0.565
T2FMRF_UM	0.945	0.546	0.611	0.912	0.648	0.292	0.472	0.534	0.878	0.472	0.649	0.751	0.931	0.605	0.544
T2FMRF_UV	0.777	0.542	0.610	0.522	0.648	0.250	0.469	0.534	0.670	0.367	0.477	0.913	0.930	0.598	0.603
VuMeter	0.764	0.586	0.635	0.771	0.840	0.440	0.559	0.583	0.855	0.474	0.771	0.940	0.908	0.624	0.609

Tabla D.9: Recall Fondo Segunda Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.997	0.976	0.989	0.967	0.593	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
DPAdaptiveMedianBGS	0.997	0.976	0.989	0.966	0.630	0.998	0.999	0.999	0.998	0.996	0.999	0.988	0.965	0.983	0.976
DPGrimsonGMMBGS	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
DPEigenbackgroundBGS	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.996	0.999	0.988	0.965	0.983	0.976
DPMeanBGS	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.996	0.999	0.988	0.965	0.983	0.976
DPPratiMediodBGS	0.997	0.976	0.989	0.768	0.561	0.998	0.999	0.999	0.824	0.997	0.999	0.988	0.965	0.983	0.976
DPWrenGABGS	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.996	0.999	0.988	0.965	0.983	0.976
DPZivkovicAGMMBGS	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.996	0.999	0.988	0.965	0.983	0.976
FuzzyChoquetIntegral	0.997	0.976	0.989	0.326	0.630	0.998	0.999	0.999	0.289	0.997	0.995	0.988	0.965	0.983	0.976
FuzzySugenoIntegral	0.997	0.976	0.988	0.323	0.630	0.998	0.999	0.999	0.998	0.996	0.995	0.988	0.965	0.983	0.975
GMG	0.997	0.976	0.988	0.966	0.537	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
KDE	0.997	0.976	0.989	0.967	0.548	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
LBAdaptiveSOM	0.997	0.976	0.989	0.967	0.538	0.998	0.999	0.999	0.998	0.996	0.998	0.988	0.965	0.983	0.976
LBFuzzyAdaptiveSOM	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
LBFuzzyGaussian	0.997	0.976	0.989	0.967	0.621	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
LBMixtureOfGaussians	0.997	0.976	0.988	0.966	0.630	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
LBSimpleGaussian	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
MOGV1BGS	0.997	0.976	0.988	0.966	0.594	0.998	0.999	0.999	0.998	0.996	0.998	0.988	0.965	0.983	0.976
MOGV2BGS	0.997	0.976	0.989	0.967	0.627	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
MultiLayerBGS	0.997	0.976	0.989	0.926	0.407	0.994	0.999	0.999	0.845	0.989	0.997	0.988	0.965	0.983	0.976
SGMM-SOD	0.997	0.976	0.989	0.966	0.471	0.996	0.999	0.999	0.998	0.996	0.998	0.988	0.965	0.983	0.976
SigmaDeltaBGS	0.997	0.976	0.988	0.967	0.527	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
T2FGMM_UM	0.997	0.976	0.989	0.641	0.624	0.998	0.999	0.999	0.879	0.997	0.999	0.988	0.965	0.983	0.976
T2FGMM_UV	0.997	0.976	0.989	0.967	0.630	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
T2FMRF_UM	0.167	0.976	0.988	0.554	0.610	0.998	0.999	0.999	0.64	0.996	0.996	0.988	0.957	0.983	0.976
T2FMRF_UV	0.997	0.976	0.988	0.967	0.610	0.998	0.999	0.999	0.998	0.997	0.999	0.988	0.965	0.983	0.976
VuMeter	0.997	0.973	0.985	0.779	0.454	0.996	0.857	0.999	0.678	0.992	0.998	0.988	0.965	0.976	0.976

Tabla D.10: Recall Frente Segunda Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.872	0.700	0.683	0.632	0.831	0.395	0.657	0.743	0.584	0.504	0.596	0.948	0.955	0.716	0.751
DPAdaptiveMedianBGS	0.871	0.707	0.733	0.699	0.853	0.400	0.652	0.725	0.853	0.620	0.649	0.942	0.960	0.749	0.745
DPGrimsonGMMBGS	0.861	0.667	0.692	0.606	0.703	0.393	0.649	0.705	0.575	0.507	0.563	0.923	0.875	0.712	0.709
DPEigenbackgroundBGS	0.873	0.712	0.744	0.609	0.714	0.394	0.646	0.749	0.646	0.599	0.697	0.878	0.944	0.722	0.709
DPMeanBGS	0.872	0.695	0.751	0.662	0.740	0.394	0.651	0.752	0.715	0.616	0.613	0.953	0.959	0.740	0.747
DPPratiMediodBGS	0.871	0.701	0.759	0.880	0.846	0.398	0.653	0.745	0.880	0.591	0.670	0.929	0.961	0.754	0.747
DPWrenGABGS	0.862	0.702	0.695	0.616	0.835	0.397	0.642	0.745	0.683	0.607	0.792	0.946	0.926	0.717	0.744
DPZivkovicAGMMBGS	0.868	0.714	0.745	0.619	0.770	0.395	0.649	0.739	0.728	0.608	0.756	0.949	0.945	0.723	0.736
FuzzyChoquetIntegral	0.862	0.720	0.730	0.985	0.772	0.394	0.657	0.723	0.974	0.552	0.861	0.958	0.947	0.771	0.749
FuzzySugenoIntegral	0.861	0.722	0.735	0.986	0.811	0.394	0.657	0.726	0.687	0.605	0.861	0.967	0.948	0.774	0.747
GMM	0.860	0.720	0.772	0.668	0.773	0.405	0.649	0.733	0.679	0.598	0.663	0.968	0.955	0.775	0.750
KDE	0.847	0.668	0.661	0.615	0.797	0.393	0.644	0.700	0.574	0.516	0.680	0.901	0.889	0.724	0.729
LBAdaptiveSOM	0.871	0.696	0.740	0.637	0.873	0.600	0.649	0.739	0.784	0.620	0.764	0.945	0.945	0.746	0.746
LBFuzzyAdaptiveSOM	0.870	0.672	0.701	0.614	0.851	0.402	0.646	0.732	0.605	0.617	0.618	0.935	0.893	0.732	0.741
LBFuzzyGaussian	0.868	0.709	0.751	0.619	0.835	0.418	0.654	0.762	0.614	0.615	0.680	0.959	0.959	0.724	0.750
LBmixtureOfGaussians	0.873	0.708	0.751	0.638	0.784	0.395	0.656	0.741	0.593	0.616	0.765	0.966	0.962	0.736	0.747
LBsimpleGaussian	0.861	0.701	0.745	0.608	0.824	0.395	0.656	0.743	0.576	0.578	0.582	0.933	0.960	0.751	0.745
MOGV1BGS	0.870	0.715	0.762	0.713	0.874	0.458	0.655	0.758	0.852	0.623	0.838	0.964	0.964	0.763	0.756
MOGV2BGS	0.847	0.668	0.677	0.607	0.646	0.396	0.645	0.717	0.597	0.501	0.577	0.946	0.897	0.735	0.718
MultiLayerBGS	0.874	0.724	0.782	0.836	0.932	0.695	0.655	0.741	0.872	0.616	0.860	0.969	0.961	0.781	0.755
SGMM-SOD	0.865	0.720	0.745	0.796	0.900	0.505	0.651	0.721	0.848	0.605	0.825	0.968	0.951	0.773	0.736
SigmaDeltaBGS	0.873	0.712	0.784	0.685	0.862	0.399	0.655	0.758	0.623	0.565	0.619	0.967	0.964	0.757	0.755
T2FGMM_UM	0.874	0.690	0.750	0.906	0.831	0.444	0.641	0.694	0.861	0.539	0.650	0.957	0.934	0.714	0.750
T2FGMM_UV	0.863	0.675	0.722	0.631	0.748	0.394	0.649	0.713	0.603	0.548	0.571	0.901	0.931	0.737	0.721
T2FMRF_UM	0.952	0.703	0.751	0.936	0.763	0.433	0.641	0.694	0.911	0.613	0.763	0.857	0.963	0.752	0.703
T2FMRF_UV	0.872	0.699	0.751	0.671	0.763	0.398	0.638	0.694	0.780	0.525	0.628	0.954	0.963	0.747	0.751
VuMeter	0.863	0.734	0.771	0.846	0.893	0.552	0.691	0.734	0.897	0.616	0.852	0.969	0.951	0.765	0.755

Tabla D.11: F1Score Fondo Segunda Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.227	0.365	0.357	0.417	0.431	0.606	0.380	0.337	0.429	0.528	0.449	0.328	0.337	0.298	0.330
DPAdaptiveMedianBGS	0.227	0.369	0.393	0.415	0.391	0.606	0.377	0.325	0.513	0.536	0.450	0.311	0.358	0.318	0.326
DPGrimsonGMMBGS	0.206	0.347	0.363	0.411	0.363	0.606	0.376	0.314	0.427	0.528	0.447	0.256	0.177	0.295	0.303
DPEigenbackgroundBGS	0.229	0.373	0.402	0.412	0.367	0.606	0.375	0.341	0.460	0.536	0.455	0.182	0.297	0.301	0.303
DPMeanBGS	0.231	0.362	0.406	0.412	0.371	0.606	0.377	0.344	0.435	0.536	0.448	0.354	0.350	0.313	0.327
DPPratiMediodBGS	0.228	0.366	0.415	0.630	0.474	0.606	0.378	0.338	0.629	0.534	0.453	0.265	0.361	0.323	0.327
DPWrenGABGS	0.206	0.367	0.365	0.412	0.381	0.606	0.373	0.339	0.463	0.536	0.461	0.328	0.253	0.299	0.326
DPZivkovicAGMMBGS	0.212	0.374	0.403	0.413	0.373	0.606	0.376	0.335	0.483	0.536	0.458	0.341	0.306	0.302	0.320
FuzzyChoquetIntegral	0.206	0.376	0.390	0.997	0.373	0.606	0.380	0.325	0.998	0.531	0.467	0.379	0.296	0.337	0.329
FuzzySugenoiIntegral	0.205	0.378	0.394	0.999	0.378	0.606	0.380	0.326	0.451	0.536	0.467	0.432	0.294	0.338	0.326
GMG	0.204	0.377	0.430	0.429	0.518	0.606	0.376	0.330	0.463	0.534	0.455	0.445	0.331	0.340	0.330
KDE	0.189	0.347	0.347	0.413	0.509	0.606	0.374	0.311	0.428	0.529	0.460	0.213	0.186	0.303	0.315
LBAdaptiveSOM	0.227	0.363	0.400	0.414	0.524	0.606	0.376	0.334	0.487	0.536	0.461	0.321	0.301	0.317	0.326
LBFuzzyAdaptiveSOM	0.225	0.350	0.368	0.412	0.390	0.606	0.375	0.331	0.430	0.535	0.449	0.285	0.193	0.307	0.322
LBFuzzyGaussian	0.224	0.370	0.406	0.412	0.406	0.606	0.379	0.352	0.434	0.535	0.455	0.389	0.353	0.303	0.329
LBmixtureOfGaussians	0.231	0.370	0.407	0.419	0.375	0.606	0.380	0.336	0.431	0.535	0.460	0.425	0.388	0.310	0.327
LBsimpleGaussian	0.206	0.366	0.403	0.411	0.379	0.606	0.380	0.338	0.428	0.534	0.448	0.286	0.357	0.321	0.326
MOGV1BGS	0.225	0.374	0.417	0.422	0.440	0.606	0.379	0.348	0.512	0.536	0.463	0.423	0.396	0.330	0.333
MOGV2BGS	0.188	0.348	0.356	0.412	0.361	0.606	0.374	0.321	0.437	0.528	0.448	0.333	0.211	0.310	0.308
MultiLayerBGS	0.232	0.379	0.434	0.501	0.739	0.610	0.379	0.335	0.613	0.543	0.465	0.451	0.367	0.346	0.332
SGMM-SOD	0.206	0.377	0.405	0.462	0.618	0.609	0.377	0.323	0.513	0.536	0.464	0.446	0.330	0.338	0.320
SigmaDeltaBGS	0.229	0.372	0.435	0.433	0.525	0.606	0.379	0.348	0.435	0.533	0.450	0.440	0.396	0.325	0.333
T2FGMM_UM	0.232	0.359	0.405	0.745	0.389	0.606	0.372	0.308	0.591	0.528	0.451	0.370	0.277	0.297	0.329
T2FGMM_UV	0.206	0.351	0.385	0.413	0.371	0.606	0.376	0.319	0.428	0.531	0.448	0.216	0.268	0.311	0.310
T2FMRF_UM	0.851	0.367	0.408	0.802	0.402	0.606	0.372	0.308	0.752	0.535	0.461	0.162	0.393	0.321	0.300
T2FMRF_UV	0.229	0.364	0.406	0.413	0.402	0.606	0.371	0.308	0.467	0.528	0.449	0.357	0.391	0.317	0.330
VuMeter	0.209	0.396	0.433	0.617	0.647	0.608	0.462	0.331	0.726	0.540	0.464	0.452	0.319	0.337	0.333

Tabla D.12: F1Score Frente Segunda Aproximación para cada Video

D.3. Resultados Tercera Aproximación

En esta sección se muestran los resultados de la evaluación de la tercera aproximación realizada en este proyecto más detalladamente, la cual busca la mejora de la segmentación Frente-Fondo usando la información de la segmentación Persona-Fondo. Puesto que, en la Figura D.1 pueden observar los resultados para la precisión de fondo para cada vídeo y algoritmo. En las Figuras D.2 D.3 D.4 D.5 D.6, se pueden observar la precisión de frente, el recall de fondo y frente y por último el F1 Score de fondo y frente respectivamente.

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.983	0.940	0.971	0.987	0.984	0.994	0.928	0.928	0.966	0.996	0.999	0.993	0.993	0.954	0.937
DPAdaptiveMedianBGS	0.987	0.954	0.968	0.986	0.990	0.996	0.952	0.979	0.979	0.992	0.999	0.994	0.995	0.970	0.939
DPGrimsonGMMBGS	0.981	0.928	0.960	0.995	0.997	0.996	0.963	0.910	0.981	0.995	0.999	0.995	0.994	0.954	0.944
DPEigenbackgroundBGS	0.995	0.983	0.982	0.997	0.996	0.998	0.954	0.977	0.998	0.995	0.999	0.995	0.997	0.979	0.945
DPMeanBGS	0.972	0.948	0.946	0.989	0.993	0.995	0.910	0.925	0.995	0.993	0.999	0.992	0.993	0.937	0.932
DPPratiMediodBGS	0.985	0.959	0.958	0.974	0.986	0.995	0.975	0.950	0.952	0.994	0.998	0.996	0.992	0.959	0.944
DPWrenGABGS	0.992	0.948	0.968	0.991	0.991	0.994	0.932	0.932	0.993	0.992	0.999	0.992	0.995	0.959	0.935
DPZivkovicAGMMBGS	0.978	0.925	0.959	0.992	0.996	0.996	0.951	0.912	0.977	0.994	0.999	0.995	0.994	0.952	0.936
FuzzyChoquetIntegral	0.985	0.943	0.945	0.973	0.989	0.997	0.956	0.967	0.951	0.991	0.997	0.990	0.993	0.941	0.922
FuzzySugenIntegral	0.987	0.935	0.938	0.973	0.989	0.996	0.955	0.964	0.990	0.990	0.997	0.988	0.992	0.937	0.923
GMG	0.993	0.933	0.933	0.992	0.982	0.996	0.985	0.968	0.970	0.998	0.999	0.994	0.990	0.978	0.946
KDE	0.992	0.984	0.991	0.994	0.979	0.999	0.973	0.974	0.991	0.990	0.999	0.993	0.997	0.987	0.933
LBAdaptiveSOM	0.990	0.964	0.966	0.993	0.987	0.994	0.955	0.948	0.997	0.992	0.999	0.994	0.995	0.975	0.937
LBFuzzyAdaptiveSOM	0.992	0.972	0.975	0.995	0.992	0.995	0.960	0.958	0.998	0.993	0.999	0.995	0.996	0.969	0.939
LBFuzzyGaussian	0.991	0.928	0.943	0.993	0.991	0.996	0.909	0.919	0.994	0.994	0.999	0.991	0.992	0.951	0.929
LBMixtureOfGaussians	0.984	0.933	0.935	0.987	0.994	0.996	0.923	0.929	0.972	0.993	0.999	0.990	0.991	0.940	0.931
LBSimpleGaussian	0.990	0.947	0.951	0.991	0.992	0.996	0.882	0.933	0.998	0.994	0.999	0.992	0.993	0.940	0.931
MOGV1BGS	0.987	0.924	0.908	0.985	0.984	0.992	0.887	0.929	0.968	0.992	0.999	0.991	0.991	0.929	0.925
MOGV2BGS	0.988	0.964	0.962	0.991	0.979	0.990	0.930	0.941	0.971	0.996	0.999	0.991	0.994	0.962	0.927
MultiLayerBGS	0.987	0.948	0.939	0.985	0.975	0.993	0.964	0.973	0.960	0.994	0.999	0.995	0.996	0.967	0.934
SGMM-SOD	0.995	0.987	0.989	0.996	0.991	0.998	0.979	0.990	0.998	0.995	0.999	0.997	0.998	0.987	0.944
SigmaDeltaBGS	0.983	0.937	0.906	0.987	0.983	0.995	0.898	0.930	0.966	0.995	0.999	0.991	0.990	0.944	0.926
T2FGMM_UM	0.970	0.953	0.949	0.975	0.986	0.987	0.888	0.922	0.951	0.993	0.995	0.978	0.978	0.958	0.914
T2FGMM_UV	0.994	0.968	0.974	0.990	0.996	0.996	0.943	0.961	0.998	0.994	0.999	0.985	0.996	0.940	0.940
T2FMRF_UM	0.968	0.932	0.917	0.973	0.978	0.987	0.891	0.921	0.951	0.992	0.995	0.985	0.976	0.927	0.920
T2FMRF_UV	0.990	0.945	0.940	0.988	0.978	0.986	0.971	0.935	0.994	0.993	0.995	0.991	0.991	0.932	0.931
VuMeter	0.995	0.891	0.917	0.976	0.987	0.994	0.955	0.964	0.953	0.990	0.997	0.991	0.992	0.958	0.935

Tabla D.13: Precisión Fondo Tercera Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.785	0.487	0.722	0.362	0.481	0.128	0.648	0.588	0.355	0.192	0.140	0.573	0.459	0.341	0.404
DPAdaptiveMedianBGS	0.762	0.592	0.948	0.634	0.718	0.153	0.483	0.343	0.893	0.450	0.412	0.650	0.691	0.530	0.370
DPGrimsonGMMBGS	0.740	0.720	0.952	0.333	0.344	0.099	0.641	0.502	0.399	0.347	0.348	0.517	0.702	0.373	0.431
DPEigenbackgroundBGS	0.622	0.326	0.920	0.181	0.296	0.048	0.442	0.301	0.193	0.278	0.141	0.469	0.472	0.278	0.263
DPMeanBGS	0.586	0.470	0.959	0.474	0.455	0.091	0.447	0.588	0.871	0.435	0.363	0.553	0.742	0.402	0.327
DPPratiMediodBGS	0.592	0.355	0.714	0.742	0.549	0.146	0.562	0.541	0.732	0.270	0.234	0.553	0.466	0.511	0.431
DPWrenGABGS	0.665	0.549	0.928	0.369	0.690	0.132	0.567	0.491	0.548	0.416	0.593	0.581	0.617	0.395	0.443
DPZivkovicAGMMBGS	0.647	0.727	0.952	0.387	0.426	0.099	0.571	0.504	0.473	0.383	0.483	0.635	0.707	0.374	0.533
FuzzyChoquetIntegral	0.619	0.529	0.963	0.999	0.461	0.099	0.703	0.457	0.999	0.325	0.841	0.717	0.736	0.660	0.578
FuzzySugenoIntegral	0.673	0.559	0.963	0.999	0.615	0.115	0.702	0.589	0.604	0.399	0.844	0.738	0.771	0.681	0.435
GMG	0.504	0.422	0.864	0.426	0.343	0.154	0.680	0.456	0.452	0.300	0.200	0.629	0.571	0.744	0.501
KDE	0.450	0.320	0.403	0.393	0.551	0.131	0.434	0.313	0.429	0.257	0.329	0.576	0.174	0.484	0.357
LBAdaptiveSOM	0.747	0.456	0.947	0.433	0.805	0.575	0.443	0.372	0.724	0.442	0.426	0.514	0.682	0.445	0.350
LBFuzzyAdaptiveSOM	0.729	0.381	0.940	0.329	0.704	0.170	0.426	0.362	0.480	0.428	0.363	0.483	0.592	0.435	0.315
LBFuzzyGaussian	0.754	0.648	0.959	0.358	0.661	0.259	0.633	0.585	0.545	0.420	0.407	0.754	0.752	0.380	0.504
LBMixtureOfGaussians	0.724	0.580	0.961	0.417	0.466	0.099	0.689	0.562	0.447	0.424	0.537	0.760	0.769	0.397	0.339
LBSimpleGaussian	0.461	0.542	0.955	0.324	0.633	0.122	0.608	0.490	0.370	0.367	0.295	0.649	0.718	0.538	0.513
MOGV1BGS	0.753	0.596	0.962	0.606	0.837	0.279	0.548	0.342	0.802	0.453	0.703	0.750	0.814	0.703	0.678
MOGV2BGS	0.674	0.473	0.872	0.340	0.188	0.172	0.730	0.438	0.407	0.234	0.225	0.762	0.776	0.632	0.648
MultiLayerBGS	0.740	0.585	0.897	0.725	0.813	0.730	0.772	0.487	0.793	0.407	0.733	0.710	0.744	0.793	0.666
SGMM-SOD	0.787	0.374	0.970	0.596	0.839	0.459	0.499	0.346	0.833	0.432	0.630	0.927	0.712	0.528	0.392
SigmaDeltaBGS	0.765	0.604	0.936	0.455	0.666	0.197	0.614	0.510	0.391	0.311	0.291	0.762	0.847	0.626	0.624
T2FGMM_UM	0.595	0.450	0.958	0.905	0.571	0.228	0.478	0.313	0.614	0.330	0.362	0.636	0.582	0.528	0.546
T2FGMM_UV	0.671	0.433	0.944	0.416	0.409	0.106	0.564	0.355	0.390	0.319	0.233	0.681	0.602	0.606	0.409
T2FMRF_UM	0.999	0.544	0.962	0.843	0.425	0.251	0.455	0.391	0.742	0.429	0.543	0.330	0.644	0.656	0.334
T2FMRF_UV	0.749	0.495	0.959	0.527	0.425	0.157	0.436	0.326	0.882	0.305	0.323	0.554	0.758	0.633	0.502
VuMeter	0.436	0.763	0.925	0.737	0.748	0.475	0.616	0.457	0.873	0.372	0.779	0.675	0.539	0.650	0.515

Tabla D.14: Precisión Frente Tercera Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.996	0.920	0.946	0.979	0.978	0.974	0.955	0.976	0.979	0.981	0.977	0.987	0.970	0.892	0.941
DPAdaptiveMedianBGS	0.994	0.939	0.995	0.995	0.997	0.968	0.878	0.829	0.995	0.996	0.995	0.990	0.989	0.941	0.931
DPGrimsonGMMBGS	0.996	0.968	0.996	0.973	0.978	0.883	0.931	0.989	0.934	0.990	0.991	0.982	0.990	0.909	0.939
DPEigenbackgroundBGS	0.983	0.745	0.991	0.915	0.903	0.762	0.849	0.787	0.837	0.990	0.973	0.979	0.976	0.771	0.857
DPMeanBGS	0.997	0.905	0.997	0.985	0.991	0.916	0.916	0.979	0.991	0.995	0.994	0.987	0.993	0.951	0.932
DPPratiMediodBGS	0.985	0.817	0.948	0.999	0.980	0.976	0.896	0.951	0.999	0.987	0.988	0.983	0.974	0.940	0.940
DPWrenGABGS	0.989	0.931	0.994	0.982	0.996	0.964	0.932	0.963	0.984	0.995	0.995	0.988	0.984	0.913	0.951
DPZivkovicAGMMBGS	0.994	0.972	0.996	0.984	0.988	0.929	0.906	0.987	0.956	0.992	0.992	0.989	0.990	0.914	0.969
FuzzyChoquetIntegral	0.988	0.928	0.997	0.999	0.990	0.919	0.953	0.913	0.999	0.997	0.999	0.995	0.995	0.983	0.987
FuzzySugenolIntegral	0.991	0.944	0.998	0.999	0.996	0.947	0.953	0.950	0.992	0.997	0.999	0.996	0.995	0.987	0.974
GMG	0.977	0.898	0.989	0.979	0.992	0.975	0.930	0.912	0.983	0.981	0.982	0.989	0.981	0.973	0.952
KDE	0.973	0.731	0.814	0.978	0.998	0.937	0.826	0.817	0.957	0.996	0.995	0.987	0.988	0.908	0.941
LBAdaptiveSOM	0.993	0.876	0.995	0.985	0.998	0.995	0.850	0.909	0.989	0.996	0.996	0.983	0.988	0.910	0.930
LBFuzzyAdaptiveSOM	0.992	0.816	0.994	0.974	0.996	0.978	0.831	0.884	0.970	0.995	0.995	0.980	0.984	0.913	0.909
LBFuzzyGaussian	0.993	0.965	0.997	0.980	0.996	0.988	0.965	0.984	0.986	0.994	0.995	0.995	0.994	0.918	0.970
LBMixtureOfGaussians	0.994	0.951	0.997	0.989	0.992	0.929	0.965	0.973	0.983	0.994	0.995	0.996	0.995	0.944	0.940
LBSimpleGaussian	0.968	0.930	0.996	0.972	0.994	0.956	0.983	0.961	0.928	0.992	0.994	0.991	0.991	0.973	0.973
MOGV1BGS	0.994	0.964	0.999	0.994	0.999	0.990	0.971	0.928	0.995	0.996	0.996	0.995	0.995	0.994	0.991
MOGV2BGS	0.990	0.867	0.988	0.973	0.993	0.992	0.965	0.945	0.962	0.983	0.988	0.995	0.993	0.967	0.987
MultiLayerBGS	0.994	0.932	0.991	0.994	0.997	0.996	0.965	0.920	0.995	0.992	0.994	0.992	0.992	0.984	0.985
SGMM-SOD	0.992	0.781	0.998	0.986	0.995	0.994	0.860	0.811	0.987	0.994	0.993	0.998	0.987	0.922	0.926
SigmaDeltaBGS	0.996	0.956	0.997	0.989	0.993	0.990	0.968	0.972	0.984	0.991	0.993	0.995	0.996	0.979	0.986
T2FGMM_UM	0.999	0.890	0.997	0.999	0.994	0.993	0.958	0.938	0.999	0.994	0.998	0.999	0.998	0.954	0.995
T2FGMM_UV	0.988	0.851	0.994	0.987	0.987	0.929	0.915	0.880	0.951	0.992	0.989	0.996	0.983	0.977	0.934
T2FMRF_UM	0.999	0.946	0.998	0.999	0.998	0.996	0.948	0.960	0.999	0.995	0.999	0.982	0.999	0.991	0.968
T2FMRF_UV	0.993	0.918	0.997	0.990	0.998	0.996	0.830	0.909	0.992	0.992	0.997	0.988	0.994	0.987	0.970
VuMeter	0.968	0.978	0.986	0.996	0.993	0.990	0.905	0.918	0.998	0.988	0.995	0.993	0.969	0.947	0.964

Tabla D.15: Recall Fondo Tercera Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.473	0.570	0.791	0.672	0.380	0.810	0.528	0.303	0.544	0.836	0.938	0.704	0.607	0.547	0.365
DPAdaptiveMedianBGS	0.588	0.669	0.759	0.646	0.629	0.859	0.721	0.837	0.713	0.688	0.917	0.749	0.750	0.685	0.396
DPGrimsonGMMBGS	0.464	0.464	0.699	0.876	0.848	0.879	0.777	0.099	0.750	0.801	0.960	0.794	0.678	0.533	0.438
DPEigenbackgroundBGS	0.800	0.906	0.871	0.925	0.841	0.944	0.743	0.834	0.980	0.801	0.955	0.775	0.853	0.814	0.504
DPMeanBGS	0.166	0.633	0.584	0.733	0.784	0.851	0.424	0.263	0.926	0.707	0.899	0.646	0.615	0.328	0.326
DPPratiMediodBGS	0.589	0.731	0.675	0.356	0.407	0.851	0.856	0.535	0.308	0.765	0.842	0.830	0.626	0.571	0.445
DPWrenGABGS	0.729	0.627	0.770	0.776	0.657	0.816	0.569	0.340	0.913	0.675	0.878	0.684	0.764	0.581	0.338
DPZivkovicAGMMBGS	0.399	0.437	0.693	0.816	0.832	0.851	0.704	0.119	0.694	0.759	0.951	0.778	0.668	0.510	0.335
FuzzyChoquetIntegral	0.561	0.586	0.578	0.322	0.698	0.913	0.72	0.721	0.287	0.623	0.740	0.563	0.650	0.350	0.161
FuzzySugenoIntegral	0.604	0.519	0.521	0.322	0.689	0.880	0.718	0.673	0.870	0.603	0.706	0.501	0.619	0.302	0.187
GMG	0.796	0.519	0.487	0.780	0.324	0.890	0.908	0.707	0.595	0.905	0.994	0.754	0.590	0.769	0.455
KDE	0.752	0.915	0.944	0.871	0.274	0.968	0.85	0.799	0.887	0.591	0.897	0.700	0.891	0.873	0.329
LBAdaptiveSOM	0.691	0.759	0.749	0.818	0.437	0.778	0.746	0.521	0.958	0.690	0.901	0.758	0.723	0.747	0.372
LBFuzzyAdaptiveSOM	0.737	0.830	0.813	0.882	0.651	0.843	0.778	0.628	0.973	0.728	0.915	0.774	0.775	0.684	0.414
LBFuzzyGaussian	0.704	0.462	0.559	0.824	0.604	0.851	0.387	0.201	0.927	0.744	0.916	0.632	0.591	0.496	0.255
LBmixtureOfGaussians	0.519	0.505	0.499	0.662	0.801	0.860	0.487	0.311	0.620	0.730	0.912	0.563	0.548	0.366	0.307
LBsimpleGaussian	0.708	0.618	0.630	0.775	0.689	0.853	0.164	0.349	0.971	0.773	0.922	0.658	0.656	0.338	0.276
MOGV1BGS	0.593	0.420	0.252	0.622	0.459	0.732	0.214	0.348	0.573	0.695	0.892	0.639	0.528	0.194	0.193
MOGV2BGS	0.635	0.766	0.729	0.788	0.278	0.712	0.536	0.408	0.612	0.839	0.950	0.636	0.737	0.591	0.228
MultiLayerBGS	0.622	0.614	0.520	0.625	0.215	0.799	0.766	0.740	0.449	0.780	0.930	0.778	0.816	0.636	0.300
SGMM-SOD	0.856	0.926	0.921	0.906	0.497	0.930	0.885	0.924	0.974	0.802	0.982	0.886	0.903	0.865	0.455
SigmaDeltaBGS	0.500	0.518	0.254	0.656	0.334	0.814	0.300	0.276	0.546	0.792	0.947	0.608	0.467	0.378	0.206
T2FGMM_UM	0.130	0.677	0.610	0.365	0.488	0.668	0.238	0.275	0.288	0.662	0.572	0.047	0.085	0.550	0.071
T2FGMM_UV	0.789	0.796	0.810	0.768	0.837	0.864	0.652	0.639	0.971	0.782	0.940	0.392	0.794	0.350	0.406
T2FMRF_UM	0.001	0.492	0.340	0.323	0.266	0.666	0.266	0.249	0.288	0.630	0.550	0.359	0.021	0.178	0.158
T2FMRF_UV	0.665	0.610	0.536	0.702	0.266	0.627	0.841	0.425	0.921	0.716	0.571	0.613	0.564	0.244	0.279
VuMeter	0.816	0.166	0.386	0.429	0.363	0.770	0.715	0.643	0.338	0.737	0.863	0.622	0.581	0.624	0.327

Tabla D.16: Recall Frente Tercera Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.989	0.930	0.958	0.983	0.981	0.983	0.941	0.951	0.972	0.988	0.988	0.990	0.981	0.921	0.939
DPAdaptiveMedianBGS	0.991	0.947	0.981	0.990	0.993	0.982	0.914	0.897	0.987	0.994	0.997	0.992	0.992	0.955	0.935
DPGrimsonGMMBGS	0.988	0.947	0.977	0.984	0.987	0.933	0.947	0.947	0.950	0.993	0.995	0.989	0.992	0.931	0.941
DPEigenbackgroundBGS	0.989	0.848	0.987	0.953	0.945	0.859	0.898	0.871	0.905	0.993	0.986	0.987	0.986	0.860	0.899
DPMeanBGS	0.984	0.926	0.971	0.987	0.992	0.953	0.912	0.951	0.993	0.994	0.996	0.989	0.993	0.944	0.932
DPPratiMediodBGS	0.985	0.881	0.952	0.986	0.983	0.985	0.934	0.950	0.975	0.990	0.993	0.989	0.983	0.949	0.942
DPWrenGABGS	0.990	0.939	0.981	0.987	0.993	0.979	0.932	0.947	0.989	0.994	0.997	0.990	0.989	0.935	0.943
DPZivkovicAGMMBGS	0.986	0.947	0.977	0.988	0.992	0.960	0.927	0.947	0.963	0.993	0.996	0.992	0.992	0.932	0.952
FuzzyChoquetIntegral	0.987	0.935	0.970	0.986	0.990	0.956	0.954	0.939	0.975	0.994	0.998	0.992	0.994	0.961	0.953
FuzzySugenIntegral	0.989	0.939	0.967	0.986	0.992	0.971	0.954	0.957	0.991	0.994	0.998	0.992	0.994	0.961	0.948
GMG	0.985	0.914	0.960	0.985	0.987	0.985	0.957	0.939	0.976	0.989	0.991	0.992	0.985	0.976	0.949
KDE	0.982	0.838	0.893	0.986	0.988	0.967	0.893	0.889	0.973	0.993	0.997	0.990	0.944	0.946	0.937
LBAdaptiveSOM	0.992	0.917	0.980	0.989	0.992	0.995	0.899	0.928	0.993	0.994	0.997	0.988	0.991	0.941	0.933
LBFuzzyAdaptiveSOM	0.992	0.887	0.984	0.984	0.994	0.987	0.890	0.919	0.983	0.994	0.997	0.987	0.990	0.940	0.924
LBFuzzyGaussian	0.992	0.946	0.969	0.987	0.993	0.992	0.936	0.951	0.990	0.994	0.997	0.993	0.993	0.934	0.949
LBMixtureOfGaussians	0.989	0.942	0.965	0.988	0.993	0.960	0.943	0.951	0.977	0.994	0.997	0.993	0.993	0.942	0.935
LBSimpleGaussian	0.979	0.938	0.973	0.981	0.993	0.975	0.930	0.947	0.959	0.993	0.996	0.992	0.992	0.956	0.951
MOGV1BGS	0.991	0.943	0.951	0.990	0.991	0.991	0.927	0.928	0.981	0.994	0.997	0.993	0.993	0.960	0.957
MOGV2BGS	0.989	0.910	0.975	0.982	0.985	0.991	0.947	0.943	0.963	0.989	0.994	0.993	0.994	0.964	0.956
MultiLayerBGS	0.990	0.938	0.964	0.989	0.985	0.994	0.964	0.946	0.977	0.993	0.997	0.994	0.994	0.975	0.959
SGMM-SOD	0.993	0.871	0.993	0.991	0.993	0.996	0.916	0.891	0.993	0.994	0.997	0.998	0.992	0.953	0.935
SigmaDeltaBGS	0.990	0.946	0.949	0.988	0.988	0.992	0.932	0.950	0.975	0.993	0.996	0.993	0.993	0.961	0.955
T2FGMM_UM	0.984	0.920	0.972	0.987	0.989	0.989	0.922	0.930	0.975	0.993	0.996	0.988	0.988	0.956	0.953
T2FGMM_UV	0.991	0.905	0.984	0.988	0.991	0.960	0.929	0.919	0.973	0.993	0.994	0.991	0.989	0.958	0.937
T2FMRF_UM	0.984	0.939	0.956	0.986	0.987	0.991	0.918	0.940	0.975	0.994	0.997	0.984	0.988	0.958	0.943
T2FMRF_UV	0.992	0.931	0.968	0.989	0.987	0.991	0.895	0.922	0.993	0.993	0.996	0.989	0.993	0.958	0.950
VuMeter	0.981	0.931	0.948	0.986	0.990	0.992	0.927	0.940	0.974	0.989	0.996	0.992	0.980	0.950	0.949

Tabla D.17: F1Score Fondo Tercera Aproximación para cada Video

	AV	BM	BU	BS	CO	CU	S6	S7	OF	PK	PD	P9	P1	TE	TR
AdaptiveBGLearning	0.578	0.515	0.744	0.630	0.623	0.725	0.580	0.399	0.530	0.735	0.640	0.628	0.475	0.408	0.377
DPAdaptiveMedianBGS	0.658	0.617	0.842	0.703	0.737	0.751	0.574	0.472	0.783	0.739	0.816	0.695	0.704	0.588	0.381
DPGrimsonGMMBGS	0.554	0.525	0.805	0.730	0.557	0.694	0.700	0.165	0.656	0.780	0.766	0.626	0.672	0.427	0.430
DPEigenbackgroundBGS	0.692	0.471	0.893	0.597	0.488	0.645	0.547	0.426	0.581	0.780	0.642	0.584	0.601	0.399	0.345
DPMeanBGS	0.237	0.529	0.724	0.709	0.652	0.693	0.429	0.362	0.902	0.751	0.786	0.595	0.651	0.352	0.327
DPPratiMediodBGS	0.583	0.462	0.651	0.614	0.633	0.747	0.677	0.536	0.504	0.725	0.672	0.658	0.496	0.516	0.437
DPWrenGABGS	0.692	0.573	0.836	0.725	0.740	0.727	0.565	0.399	0.857	0.725	0.804	0.627	0.669	0.462	0.374
DPZivkovicAGMMBGS	0.466	0.509	0.801	0.749	0.634	0.698	0.619	0.192	0.631	0.771	0.779	0.697	0.668	0.420	0.405
FuzzyChoquetIntegral	0.581	0.543	0.721	0.996	0.609	0.704	0.710	0.552	0.998	0.677	0.780	0.631	0.687	0.450	0.247
FuzzySugenoIntegral	0.633	0.525	0.674	0.999	0.718	0.718	0.708	0.627	0.877	0.657	0.756	0.596	0.681	0.410	0.256
GMM	0.612	0.445	0.586	0.700	0.726	0.766	0.775	0.552	0.585	0.785	0.688	0.685	0.531	0.750	0.473
KDE	0.544	0.465	0.557	0.772	0.671	0.748	0.571	0.440	0.778	0.625	0.806	0.629	0.278	0.618	0.342
LBAdaptiveSOM	0.714	0.556	0.835	0.761	0.772	0.801	0.550	0.427	0.901	0.739	0.822	0.612	0.684	0.550	0.360
LBFuzzyAdaptiveSOM	0.729	0.511	0.871	0.734	0.726	0.763	0.544	0.449	0.837	0.765	0.815	0.595	0.659	0.522	0.357
LBFuzzyGaussian	0.725	0.517	0.705	0.736	0.737	0.810	0.478	0.296	0.870	0.775	0.815	0.687	0.637	0.418	0.327
LBmixtureOfGaussians	0.595	0.524	0.654	0.661	0.673	0.700	0.567	0.399	0.620	0.766	0.818	0.646	0.608	0.371	0.322
LBsimpleGaussian	0.546	0.566	0.757	0.682	0.724	0.725	0.257	0.405	0.746	0.785	0.794	0.652	0.664	0.410	0.355
MOGV1BGS	0.657	0.485	0.393	0.675	0.711	0.739	0.305	0.331	0.616	0.743	0.820	0.690	0.592	0.291	0.298
MOGV2BGS	0.634	0.556	0.779	0.698	0.616	0.720	0.609	0.414	0.578	0.765	0.735	0.692	0.746	0.603	0.333
MultiLayerBGS	0.669	0.568	0.627	0.674	0.730	0.805	0.768	0.584	0.630	0.790	0.807	0.742	0.773	0.691	0.409
SGMM-SOD	0.816	0.520	0.939	0.816	0.848	0.886	0.634	0.487	0.903	0.816	0.820	0.905	0.784	0.646	0.416
SigmaDeltaBGS	0.599	0.543	0.382	0.654	0.677	0.779	0.399	0.343	0.552	0.786	0.794	0.676	0.536	0.464	0.301
T2FGMM_UM	0.168	0.531	0.743	0.758	0.629	0.699	0.305	0.287	0.483	0.709	0.597	0.087	0.142	0.530	0.119
T2FGMM_UV	0.721	0.541	0.871	0.734	0.618	0.704	0.597	0.452	0.774	0.781	0.744	0.488	0.672	0.429	0.399
T2FMRF_UM	0.835	0.509	0.498	0.816	0.626	0.701	0.320	0.301	0.780	0.688	0.587	0.343	0.040	0.272	0.214
T2FMRF_UV	0.698	0.537	0.685	0.712	0.626	0.661	0.572	0.364	0.901	0.743	0.594	0.581	0.622	0.341	0.354
VuMeter	0.565	0.219	0.502	0.663	0.770	0.780	0.705	0.529	0.676	0.755	0.826	0.645	0.483	0.611	0.382

Tabla D.18: F1Score Frente Tercera Aproximación para cada Video

Apéndice E

Presupuesto

1. Ejecución Material

- Compra de ordenador personal (Software incluido)2.000 €
 - ◇ Alquiler de impresora láser durante 6 meses260 €
 - ◇ Material de oficina 150 €
 - ◇ Total de ejecución material 2.400 €

1. Gastos generales

- 16 % sobre Ejecución Material 352 €

2. Beneficio Industrial

- 6 % sobre Ejecución Material 132 €

3. Honorarios Proyecto

- 1800 horas a 15 € / hora27.000 €

4. Material fungible

- Gastos de impresión280 €
- Encuadernación 200 €

5. Subtotal del presupuesto

- Subtotal Presupuesto32.774 €

6. I.V.A. aplicable

- 21 % Subtotal Presupuesto6.882,5 €

7. Total presupuesto

- Total Presupuesto39.656,5 €

Madrid, Diciembre 2014

El Ingeniero Jefe de Proyecto

Fdo.: Raúl Lara Arranz

Ingeniero de Telecomunicación

Apéndice F

Pliego De Condiciones

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, Segmentación Persona-Fondo usando información de segmentación Frente-Fondo. 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 entorno. 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 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.