

UNIVERSIDAD AUTÓNOMA DE MADRID  
ESCUELA POLITÉCNICA SUPERIOR



PROYECTO FIN DE CARRERA  
Ingeniería de Telecomunicación

ESTIMACIÓN DE FONDO DE ESCENA EN SECUENCIAS DE VÍDEO

Jaime Gallo de Cal  
Julio 2016



# ESTIMACIÓN DE FONDO DE ESCENA EN SECUENCIAS DE VÍDEO

**Autor: Jaime Gallo de Cal**  
**Tutor: Diego Ortego Hernández**  
**Ponente: José María Martínez Sánchez**

email: {jaime.gallo@estudiante.uam.es, diego.ortego@uam.es, josem.martínez@uam.es}



**Video Processing and Understanding Lab**  
**Departamento de Tecnología Electrónica y de las Comunicaciones**  
**Escuela Politécnica Superior**  
**Universidad Autónoma de Madrid**  
**Julio 2016**

Trabajo parcialmente financiado por el gobierno español bajo el proyecto TEC2014-53176-R (HAVideo)





## Resumen

En este trabajo, se propone un algoritmo de estimación de fondo de escena en secuencias de vídeo con cámara fija basado en un análisis espacio-temporal y *online* a nivel de bloque de la escena.

En primer lugar, se realiza un estudio del estado del arte para familiarizarse con la tarea de estimación de fondo y especialmente con las técnicas utilizadas en el laboratorio VPULab previamente. Después, el trabajo se centra en incorporar un análisis temporal *online* propuesto en la literatura en el algoritmo desarrollado. A continuación, se mejora la etapa de estimación parcial del fondo o selección de semillas, que supone el punto de partida para una reconstrucción completa del fondo, mediante la selección de un mayor número de candidatos correctos. Además, se propone una nueva medida de continuidad espacial para la reconstrucción total del fondo que incorpora información tanto de los bordes del bloque como del interior del mismo.

Finalmente, se realizan experimentos para validar cada etapa del algoritmo desarrollado y el rendimiento obtenido frente a algoritmos recientes de la literatura, entendiendo así la utilidad de las técnicas propuestas.

## Palabras clave

Inicialización de fondo, Estimación de fondo, Multisemilla, Superbloque, Homogeneidad, *Clustering*, Continuidad espacial, Filtrado de movimiento, Reconstrucción Multicamino.



## Abstract

In this work, we propose a block-wise and online Background Initialization algorithm for static camera video sequences based in a spatio-temporal scene analysis.

First, the state-of-the-art is studied to understand the Background Initialization task and the challenges that it poses. In particular, previous techniques previously used in the laboratory VPULab where specially considered.

Then, the work focuses in incorporating a temporal online analysis proposed in the literature into the proposed algorithm. Furthermore, the initial stage of the spatial background reconstruction, the seed selection or partial background initialization stage, has been improved leading to more initial information before the entire background reconstruction. Additionally, a new measure of spatial continuity is proposed to perform the complete background reconstruction by using both information of the block borders and its interior.

Finally, experiments are carried out to validate each stage of the developed algorithm and the performance achieved against related literature, thus understanding the algorithm advantages and weaknesses see definition of weakness.

## Keywords

*Background Initialization, Background estimation, Multiseed, Superblock, Homogeneity, Clustering, Spatial continuity, Motion filtering, Multipath Reconstruction.*



# Agradecimientos

En primer lugar quería agradecerle a mi tutor, Diego Ortego, la implicación en el proyecto y la predisposición a ayudarme siempre que le he pedido ayuda, gracias de verdad. También agradecerles a Chema y a Jesús la posibilidad de llevar a cabo este proyecto.

Quiero darles las gracias a mis padres y a mis hermanos, Carlos y Pablo, sin vosotros esto no lo habría conseguido, gracias por estar siempre ahí. Gracias también a mis abuelos, que lo podrán celebrar allá donde estén.

Una mención especial a mi novia, Sofía, gracias por animarme siempre con tu sonrisa y confiar siempre en mí, por fin! este proyecto también es tu proyecto, que sé que ha sido un parto.

Por supuesto no puedo olvidarme de esos compañeros de carrera al principio, luego de fatigas, alguna vez de fiesta y ahora amigos, en especial a Cubero, Muse, Chuchi , Chema y como no Marabini, que lo que Roma une....

Darle las gracias también a mis amigos Juan, More, Álvaro, Sarmi, Yori, Quereda, David, Dudu, Osoro y Kike. A toda la gente que de alguna manera me ha ayudado a llegar hasta aquí. Gracias a todos!

Jaime Gallo de Cal

Julio 2016



# Índice general

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Motivación . . . . .	1
1.2	Objetivos . . . . .	2
1.3	Organización de la memoria . . . . .	2
<b>2</b>	<b>Estado del arte</b>	<b>5</b>
2.1	Introducción . . . . .	5
2.2	Trabajo relacionado: Inicialización de fondo . . . . .	6
2.3	Trabajo relacionado: Generación de candidatos a fondo . . . . .	8
<b>3</b>	<b>Algoritmo propuesto</b>	<b>11</b>
3.1	Esquema general . . . . .	11
3.2	Análisis temporal . . . . .	12
3.2.1	Filtrado de movimiento . . . . .	12
3.2.2	<i>Online Clustering</i> . . . . .	12
3.3	Análisis espacial . . . . .	14
3.3.1	Selección de semillas: Multisemilla . . . . .	14
3.3.2	Reconstrucción Secuencial Multicamino . . . . .	18
<b>4</b>	<b>Resultados experimentales</b>	<b>23</b>
4.1	Marco de evaluación . . . . .	23
4.1.1	Dataset . . . . .	23
4.1.2	Métricas . . . . .	24
4.2	Evaluación . . . . .	26
4.2.1	Análisis temporal . . . . .	26
4.2.2	Selección de semilla . . . . .	29
4.2.3	Evaluación estimación de fondo . . . . .	30

<b>5 Conclusiones y trabajo futuro</b>	<b>35</b>
5.1 Conclusiones . . . . .	35
5.2 Trabajo futuro . . . . .	36
<b>Bibliography</b>	<b>37</b>
<b>6 Presupuesto</b>	<b>45</b>
<b>7 Pliego de condiciones</b>	<b>47</b>

# Índice de figuras

1.1	Ejemplo de fondo ocluido y fondo a generar. . . . .	2
2.1	Ejemplo de espacios con multitudes. . . . .	6
3.1	Diagrama de bloques del algoritmo propuesto. . . . .	12
3.2	Ejemplo medida de <i>matching</i> basada en el ratio entre bloques. . . . .	14
3.3	Ejemplo de selección de semilla. . . . .	16
3.4	División del fondo en cuatro superbloques. . . . .	16
3.5	Ejemplo esquema multisevilla. . . . .	17
3.6	Mapa de semillas final. . . . .	17
3.7	Esquema de reconstrucción multicamino. . . . .	19
3.8	Funciones densidad de probabilidad de $d_e$ y $d_c$ . . . . .	20
3.9	Ejemplo de fondo reconstruido $B$ . . . . .	21
4.1	Dataset utilizado para la evaluación. . . . .	24
4.2	Ejemplo comparativo de GT y fondo ideal. . . . .	29
4.3	Ejemplo de evolución del algoritmo. . . . .	31
4.4	Comparativa de reconstrucciones con los mejores algoritmos del estado del arte. . . . .	32



# Índice de tablas

4.1	Descripción del dataset. . . . .	25
4.2	Tabla PRF para $W = 8$ píxeles. . . . .	27
4.3	Tabla PRF para $W = 16$ píxeles. . . . .	27
4.4	Tabla PRF para $W = 32$ píxeles. . . . .	28
4.5	Tabla comparativa en función de las muestras de cambios de iluminación y matriz de costes. . . . .	28
4.6	Tablas del porcentaje de error en el fondo reconstruido con los mejores candidatos del Clustering. . . . .	29
4.7	Comparación de resultados Selección de semilla. . . . .	30
4.8	Tabla comparativa con el método RMR. . . . .	33
4.9	Resultados de la reconstrucción de fondo, para todos los modelos. . . . .	34
4.10	Comparativa de resultados con otros algoritmos del estado del arte. . . . .	34



# Acrónimos

<b>AE</b>	<i>Average Error</i>
<b>ACE</b>	<i>Average Cluster Error</i>
<b>AGE</b>	<i>Average Gray-level Error</i>
<b>BB</b>	<i>Bounding-Box</i>
<b>BI</b>	<i>Background Initialization</i>
<b>BS</b>	<i>Background Substraction</i>
<b>CEP</b>	<i>Cluster Error Pixel</i>
<b>CQM</b>	<i>Color image Quality Measure</i>
<b>DCT</b>	<i>Discrete Cosine Transform</i>
<b>EP</b>	<i>Error Pixel</i>
<b>EB</b>	<i>EigenBackground</i>
<b>GT</b>	<i>Ground-truth</i>
<b>HSV</b>	<i>Hue, Saturation, value: Es un modelo de color</i>
<b>LAB</b>	<i>Es el modelo cromático usado normalmente para describir todos los colores que puede percibir el ojo humano. La <b>L</b> corresponde a la luminosidad y <b>A</b> y <b>B</b> corresponden a los colores complementarios</i>
<b>MAD</b>	<i>Mean of Absolute Differences</i>
<b>MRF</b>	<i>Markov Random Fields</i>
<b>MSE</b>	<i>Mean Square Error</i>
<b>MS-SSIM</b>	<i>Multi-Scale Structural Similarity Index</i>

<b>PSNR</b>	<i>Peak-Signal-to-Noise-Ratio</i>
<b>RPCA</b>	<i>Robust Principal Component Analysis</i>
<b>RGB</b>	<i>Red, Green, Blue. Es un modelo de color: El modelo aditivo de los colores rojo, verde y azul suma los canales para generar un array de colores</i>
<b>SLC</b>	<i>Similarity Criteria for Labels</i>
<b>SMR</b>	<i>Sequential Multipath Reconstruction</i>
<b>SVM</b>	<i>Support Vector Machine</i>
<b>YCbCr</b>	<i>Espacio de color basado en la luminancia y las crominancias</i>

# Capítulo 1

## Introducción

### 1.1 Motivación

Actualmente, el avance de la tecnología y la necesidad de seguridad han provocado la expansión de cámaras de vídeo-vigilancia en todo tipo de entornos públicos y privados. Este nuevo escenario ha provocado la necesidad de automatizar los procesos de monitorización, mediante algoritmos de visión por computador, para ayudar así a los operadores de seguridad a identificar posibles eventos de interés.

Una primera etapa en los algoritmos de monitorización, es la detección de las regiones de interés (objetos o personas). Es habitual realizar esta tarea mediante algoritmos de segmentación frente-fondo, conocidos como algoritmos de *Background Subtraction* (BS)[2]. La función de estos algoritmos es la de separar el frente del fondo mediante varias etapas de análisis: Modelado del fondo, Inicialización del fondo, Actualización del fondo y Detección de frente [2]. Este proyecto se centra en la etapa de Inicialización de fondo o *Background Initialization* (BI).

La etapa de BI consiste en generar el fondo de la escena a partir de una secuencia de entrenamiento en la que generalmente aparecen objetos de frente ocluyendo el fondo, como se puede ver en la Figura 1.1 . Este proceso es complejo y sufre problemas relacionados con la visibilidad del fondo y con factores fotométricos. Por un lado, la visualización del fondo es un factor clave debido a que ciertas regiones pertenecientes al fondo de la escena pueden ser visibles solo en breves instantes. Este aspecto, limita el rendimiento de técnicas tradicionales, donde el fondo se reconstruye atendiendo a criterios de repetitividad en la visualización. Por otro lado, factores fotométricos como las sombras y los cambios de iluminación dificultan la reconstrucción del fondo de escena, debido a que generan distintas representaciones válidas del mismo.

Observando el estado del arte reciente, existen distintas perspectivas para afrontar la tarea de inicializar el fondo de una escena. No obstante, ninguno de ellos ofrece un buen rendimiento en todas las situaciones, ya sea porque no son robustos cuando la secuencia de entrenamiento contiene objetos estáticos durante largos periodos de tiempo, porque tiene lugar una alta concurrencia



Figura 1.1: Ejemplo de fondo ocluido y fondo a generar. En (a), (b) y (c) se observa a varias personas ocluyendo el fondo y en (d) se puede observar el fondo libre de objetos.

de objetos o porque el fondo varía como consecuencia de sombras y cambios de iluminación

Por tanto, la motivación de este Proyecto Fin de Carrera es contribuir al campo de BI mediante el desarrollo de un algoritmo que se enfrente a las limitaciones actuales de esta tarea.

## 1.2 Objetivos

En este Proyecto Fin de Carrera el objetivo principal es desarrollar un algoritmo de Inicialización de fondo de escena para secuencias de vídeo con cámara estática que mejore las aproximaciones actuales del estado del arte. El desarrollo del objetivo principal se desarrollará mediante los siguientes sub-objetivos:

- 1) Estudio del estado del arte actual: Se analizarán los algoritmos de Inicialización de fondo de escena de la literatura. El objetivo es comprender los mecanismos de inicialización existentes y sus limitaciones.
- 2) Desarrollo de un algoritmo de Inicialización de fondo de escena: Se partirá de técnicas de visión por computador desarrolladas en el Grupo de Tratamiento e Interpretación de Vídeo (VPULab) de la Universidad Autónoma de Madrid (UAM) relacionadas con la tarea de inicialización. El objetivo es desarrollar un nuevo algoritmo que mejore las capacidades de los algoritmos previos de la literatura.
- 3) Evaluación comparativa de resultados: Se va a evaluar el rendimiento del algoritmo desarrollado desde distintas perspectivas sobre un *dataset* completo de secuencias. El objetivo es conocer el funcionamiento de las técnicas desarrolladas frente a los algoritmos de la literatura.

## 1.3 Organización de la memoria

La memoria consta de los siguientes capítulos:

- Capítulo 1: Introducción, motivación del proyecto y objetivos.

- Capítulo 2: Estudio del estado del arte en la Inicialización de fondo para secuencias de vídeo con cámara estática.
- Capítulo 3: Descripción del algoritmo desarrollado para Inicializar el fondo de escena de una secuencia de vídeo.
- Capítulo 4: Resultados experimentales.
- Capítulo 5: Conclusiones y trabajo futuro.



## Capítulo 2

# Estado del arte

En este capítulo se estudia el estado del arte relacionado con la Inicialización de fondo en secuencias de vídeo.

Este estudio se encuentra dividido en: la introducción a la Inicialización de fondo (Sección 2.1), el trabajo relacionado con la Inicialización de fondo (Sección 2.2) y por último el trabajo relacionado a la generación de candidatos para el *Clustering Online* a través del modelado de bloques (Sección 2.3).

### 2.1 Introducción

La Inicialización de fondo o *Background Initialization* (BI) consiste en la estimación de una imagen de fondo de escena a partir de una secuencia de entrenamiento, esta técnica es conocida también como *Bootstrapping*, Estimación de fondo, Generación de fondo o Reconstrucción de fondo [3]. BI suele entenderse como una etapa dentro de los algoritmos de sustracción de fondo o *Background Subtraction* (BS) [2][4]. No obstante, la Inicialización de fondo tiene otras aplicaciones además de la segregación del objetos en secuencias de vídeo [5], como por ejemplo, la representación de información redundante en la compresión de vídeo [6], la reparación de imágenes deterioradas en *inpainting* [7], la protección de la privacidad en vídeo [8] y para la obtención de imágenes libres de objetos en *computational photography* [9].

En el estado del arte de BI existen numerosos algoritmos [3] que llevan a cabo la generación del fondo atendiendo a criterios de dominancia temporal pues asumen que el fondo es fácilmente visible en las secuencias de entrenamiento, siendo los objetos de la escena minoritarios. Este tipo de técnicas que hacen usos de información temporal para inicializar el fondo, predominan en los algoritmos de *Background Subtraction* (BS). Estos algoritmos se encargan de obtener una máscara binaria de los objetos de interés (frente o *foreground*) que hay en la escena a partir de un modelo de fondo de la misma. En general, un algoritmo de BS se puede definir en función de la estrategia seguida para abordar las etapas de Modelado de fondo, Inicialización



Figura 2.1: Ejemplo de espacios con multitudes.

de fondo, Actualización de fondo y Detección de frente [2]. El Modelado de fondo genera un modelo de fondo de escena cuyo objetivo es capturar la multimodalidad de la misma (p.ej. árboles, escaleras mecánicas o fuentes), la Inicialización de fondo genera una imagen de fondo para construir un punto de partida adecuado para la etapa de Modelado, la Actualización de fondo adapta el modelo a las variaciones temporales y la Detección de frente define cómo detectar los objetos de la escena mediante comparaciones entre los frames bajo análisis y el modelo de fondo construido. La suposición de dominancia temporal del fondo no es válida en muchos escenarios de vídeo-vigilancia (p.ej. centros comerciales, aeropuertos o estaciones de tren) como se puede ver en la Figura 2.1, donde pueden existir muchos objetos en el frente debido a aglomeraciones u objetos estacionarios, factores que invalidan dicha suposición. Es importante diferenciar entre las etapas de Modelado de fondo, donde el principal objetivo es lidiar con fondos dinámicos, de la etapa de BI cuyo objetivo es generar una imagen de fondo libre de objetos de frente y donde, por tanto, la baja visibilidad del fondo es una limitación. En la literatura, los algoritmos con mejor rendimiento en entornos concurridos hacen uso tanto de información temporal como espacial. La información temporal analiza las variaciones de intensidad a lo largo de la secuencia de entrenamiento [10][11][12][13], generando una serie de candidatos posibles para el fondo, que son seleccionados en función de la distancia entre bloques. Es la etapa espacial la encargada de seleccionar los candidatos adecuados, para lo cual predomina la asunción de homogeneidad entre candidatos de fondo [5][14][15][13] como criterio para llevar a cabo la estimación del mismo.

## 2.2 Trabajo relacionado: Inicialización de fondo

En el estado del arte, se propone una clasificación de las aproximaciones de BI en función de la estrategia aplicada: temporal y espacial [16]. Estas estrategias pueden analizar la secuencia con un procesado *batch* u *online* y operar tanto a nivel de región (bloque) como de píxel.

Los algoritmos basados en estrategias temporales son comunes en la sustracción de fondo [3], donde se toma la primera imagen como un fondo a actualizar *frame a frame* en función de las nuevas representaciones de la escena analizadas [17][18][19]. Más allá de estas técnicas, las aproximaciones que se valen de *Robust Principal Component Analysis* (RPCA) [20] modelan el fondo de una secuencia de vídeo mediante un subespacio de bajo rango y el frente se define

como los *outliers* dispersos de dicho subespacio. Sin embargo, los métodos RPCA tienen una importante limitación para generar el fondo de una escena debido a que pierden la estructura temporal y espacial al representar cada *frame* como un vector columna. Existen otras aproximaciones temporales que hacen uso de técnicas de *EigenBackground* (EB), donde se calcula una base de autovectores de la secuencia de vídeo para modelar su fondo a nivel de escena [21] o bloque [22]. Los métodos EB requieren una consistencia temporal del fondo para operar correctamente, asumiéndose que el fondo solo sufre oclusiones temporalmente breves [23]. En definitiva, los métodos RPCA y EB no consideran múltiples bases que tengan en cuenta los distintos aspectos de la escena a lo largo de la secuencia de vídeo, así como las relaciones espaciales entre posiciones de la escena vecinas, disminuyendo así las capacidades para inicializar el fondo en presencia de objetos que se mueven lentamente o permanecen estáticos. Otra técnica temporal muy extendida para inicializar el fondo es la mediana temporal a nivel de píxel [24][25], pero tiene la importante limitación de no poder operar correctamente cuando existen objetos estacionarios que superen 50 % del tiempo de la secuencia a inicializar, pues estos son incluidos como parte del fondo. En cuanto a técnicas temporales que estiman el movimiento, como el flujo óptico [10][26][27] o las diferencias entre imágenes [11][27][28], pueden ser muy útiles para eliminar objetos de frente que no son parte del fondo. También se emplea la continuidad temporal de la intensidad de los píxeles para obtener distintos modelos de fondo en cada localización espacial [10][26][12], asumiendo de forma no necesariamente cierta que intervalos temporalmente no continuos describen representaciones del fondo distintas. Por tanto, una agrupación de intervalos no continuos puede solventar dicho problema [29][30][15][13][31].

Aunque, algunos algoritmos sólo utilizan el análisis temporal [25][12], es necesario un análisis espacial cuando existen objetos en movimiento o estáticos, pues asumir que la información temporal dominante describe el fondo no es una estrategia eficaz. Es habitual imponer restricciones de homogeneidad en el fondo para decidir si los nuevos píxeles o bloques que se analizan pertenecen al fondo [28]. En [31][13], la transformada discreta del coseno (DCT) empleada en un *framework* de MRF (*Markov Random Field*) introduce la asunción de homogeneidad del fondo entre localizaciones vecinas, usando además estimaciones iterativas del fondo para corregir posibles errores [13]. Alternativamente, [29] sustituye la DCT por la transformada de Hadamard para disminuir la complejidad computacional, usando además correcciones iterativas del fondo a partir de las características del gradiente entre candidatos vecinos. La homogeneidad también se puede utilizar para encontrar el mejor bloque entre candidatos parcialmente solapados y los ya establecidos en el fondo [15]. Recientemente, [16] propone un esquema de reconstrucción espacial multicamino para encontrar los trayectos espaciales óptimos para reconstruir los bloques de fondo mediante medidas de continuidad inter- e intra-bloque. Por otra parte, otros algoritmos modelan la homogeneidad espacial junto con la información temporal en *frameworks* de minimización de energía, tales como *Loopy Belief Propagation* [28][14], *Graph-cuts* [7], *Conditional Mixed-State*

MRFs [32], *Dynamic* MRFs [5].

En definitiva, existen múltiples estrategias donde los algoritmos recientes utilizan la información temporal junto con restricciones de homogeneidad sobre el fondo a estimar. La principal limitación de los enfoques actuales implica situaciones de baja visibilidad del fondo donde los esquemas de homogeneidad existentes no resuelven correctamente la tarea de inicialización.

## 2.3 Trabajo relacionado: Generación de candidatos a fondo

En la literatura, numerosos algoritmos de BI llevan a cabo una etapa de agrupamiento temporal o *clustering* a nivel de bloque [29][15][13][16] de las distintas representaciones de la escena para generar así un conjunto de candidatos espaciales a conformar el fondo. Estas técnicas emplean medidas de similitud o diferencia entre dos bloques  $R$  y  $R'$ . En esta sección se van a analizar distintas alternativas presentadas en la literatura para modelar distancias entre bloques, tanto empleadas en BI como en campos relacionados como seguimiento de objetos [33][34] o evaluación en segmentación de vídeo [35].

Es habitual en técnicas de seguimiento de objetos utilizar un rectángulo o *bounding-box* (BB) para definir cada objeto que se desea seguir en la escena, para después modelar dicho BB con distintas características. Ese seguimiento implica comparaciones entre el BB candidato a definir el objeto en cada *frame* y el BB definido como modelo de objeto a seguir, estrategia que puede aplicarse para modelar y comparar bloques de la escena. En [35], se modela cada BB empleando un histograma, de forma que la comparación entre dos BBs se realiza mediante:

$$d_h(R, R') = \chi(h(R), h(R')), \quad (2.1)$$

donde  $\chi(\cdot, \cdot)$  denota la distancia *Chi-square* [35] y  $h(\cdot)$  denota histograma. En particular, [35] trabaja sobre el espacio de color YCbCr. Por otro lado, en [36], se realiza una comparativa de distintas características útiles para modelar los BB, obteniendo como características más efectivas aquellas basadas en distancias de matrices de covarianzas. Las matrices de covarianza permiten modelar conjuntamente varias características sin necesidad de normalización y combinación, modelando como dichas características del bloque varían unas respecto a las otras. La distancia entre matrices de covarianza [36], se define como:

$$d_c(R, R') = \sqrt{\sum_i \ln^2 \lambda_i(\Sigma_R, \Sigma_{R'})}, \quad (2.2)$$

donde  $\Sigma_R$  y  $\Sigma_{R'}$  son las matrices de covarianza de las características que modelan cada bloque y  $\lambda_i$  denota los autovalores generalizados de las matrices de covarianza.

En cuanto a algoritmos diseñados para agrupar bloques, en [13] se utiliza el coeficiente de

correlación y la distancia euclídea.

$$d_r(i, j) = \rho(R, R'), \quad (2.3)$$

$$d_e = \frac{1}{n} \sum_{n=0}^{n-1} |R(i, j) - R'(i, j)|, \quad (2.4)$$

donde  $d_r$  es la similitud entre bloques basada en el coeficiente de correlación de Pearson  $\rho$ ,  $d_e$  es la distancia euclídea y  $n$  es el número de elementos. Estas dos ecuaciones evalúan respectivamente el coeficiente de correlación y la distancia euclídea entre los dos bloques.

Por último, se estudia la distancia entre bloques con el ratio, calculado entre dos bloques  $R$  y  $R'$ .

$$r_{max} = (R, R') = \max \{r(R, R')\}, \quad (2.5)$$

donde  $r$  corresponde al ratio. En este caso, se ha trabajado con diferentes espacios de color, RGB y RGBS. Se calcula de la siguiente manera:

$$r(R, R') = 1 - \frac{\min \{R, R'\} + m}{\max \{R, R'\} + m}, \quad (2.6)$$

siendo  $m$  una constante de corrección[37], utilizada para gestionar la inestabilidad del ratio con valores de baja intensidad. Cuando se comparan los valores de ambos píxeles se divide el valor máximo entre el mínimo para obtener  $H \in [0, 1]$ .

$$d_\mu(R, R') = \frac{1}{|R|} \sum_{p \in R} r_{max}(R, R'), \quad (2.7)$$

$$d_{\sigma^2}(R, R') = \frac{1}{|R| - 1} \sum_{p \in R} (r_{max}(R, R') - \mu(R, R'))^2, \quad (2.8)$$

donde  $d_\mu$  se corresponde con la media y  $d_{\sigma^2}$  es la varianza [16]. La media y la varianza serán los valores que posteriormente se utilicen para estimar, en la medida de lo posible, la diferencia entre los bloques.



## Capítulo 3

# Algoritmo propuesto

En este capítulo se propone un algoritmo para la Inicialización del fondo de escena en secuencias de vídeo con cámara fija, capaz de reconstruir el fondo en entornos con multitudes y con objetos estacionarios.

El capítulo se divide en tres secciones: un esquema general del algoritmo (Sección 3.1), la segunda se centra en el análisis temporal del algoritmo (Sección 3.2) y la última en el análisis espacial (Sección 3.3).

### 3.1 Esquema general

El algoritmo propuesto realiza un análisis espacio-temporal a nivel de bloque para cada *frame*  $I_t$  de una secuencia de entrenamiento  $\mathbb{F} = \{I_1, \dots, I_T\}$ , donde  $T$  es el número de *frames*, para generar un fondo  $B$  libre de objetos en movimiento y de objetos estacionarios. En primer lugar, el módulo encargado de dividir cada  $I_t$  en bloques no solapados  $R_t^s$  de tamaño  $W \times W$ , donde  $s$  es el índice bidimensional para la ubicación espacial de cada bloque. En segundo lugar, el módulo de análisis temporal crea una serie de candidatos para el fondo  $C_q^s$  para cada ubicación espacial  $s$ ,  $q \in \{1 \dots N^s\}$  y  $N^s \leq T$  es el número de candidatos. El análisis temporal consiste en dos etapas, una primera de filtrado de movimiento que se utiliza para descartar bloques  $R_t^s$  que pertenecen a objetos en movimiento y por tanto no son candidatos a formar el fondo y una etapa de *Online Clustering* que se encarga de generar el conjunto de candidatos a formar el fondo. Por último, el módulo de análisis espacial reconstruye el fondo de cada ubicación espacial  $s$ , parcialmente estimado en la etapa de selección de semillas. En particular, se emplea un esquema multisevilla con semillas en distintas zonas de la imagen que sirven de base para realizar varias reconstrucciones del fondo y generar las semillas finales a partir de dichas reconstrucciones. Finalmente, se realiza una reconstrucción del fondo partiendo de las semillas obtenidas. Todas las reconstrucciones realizadas emplean el esquema de Reconstrucción Multicamino que genera el fondo de forma iterativa para cada localización espacial  $s$  con el candidato  $C_*^s$ , haciendo uso

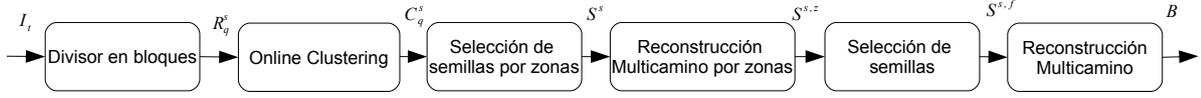


Figura 3.1: Diagrama de bloques del algoritmo propuesto.

de dos medidas de continuidad espacial entre bloques. En la Figura 3.1 se muestra un diagrama del flujo del programa.

## 3.2 Análisis temporal

El módulo de análisis temporal es el encargado de generar los candidatos del fondo para cada localización espacial  $\mathbf{s}$ . Este módulo se divide en 2 fases, filtrado de movimiento y *Online Clustering*.

### 3.2.1 Filtrado de movimiento

La etapa de filtrado de movimiento descarta los  $R_t^s$  bloques correspondientes a los objetos que no pueden ser candidatos para el fondo reconstruido  $B$ . Para toda la secuencia de entrenamiento se calcula el nivel de movimiento a nivel de bloque:

$$\lambda_t^s = \begin{cases} 1, & \exists \mathbf{p} \in \mathbf{s} : |I_t^{\mathbf{p}} - I_{t-k}^{\mathbf{p}}| > \eta, \\ 0, & \text{resto} \end{cases}, \quad (3.1)$$

donde  $k$  es la distancia temporal entre imágenes comparadas ( $k$  debe ser pequeña),  $\mathbf{p}$  es el índice bi-dimensional para las localizaciones de píxel en  $\mathbf{s}$  y  $\eta$  es el umbral calculado automáticamente [38] para detectar cambios de intensidad entre las imágenes analizadas  $I_t^{\mathbf{p}}$  e  $I_{t-k}^{\mathbf{p}}$ .  $\lambda_t^s$  toma valor 1 cuando se detecta movimiento y 0 en caso de que no, rechazando de esta manera el bloque  $R_t^s$  asociado en caso de detección de movimiento y guardando como posible candidato el bloque  $C_q^s$  en el que no se detecte el movimiento. La Ecuación 3.1 implica la visualización del fondo para  $k$  imágenes consecutivas, como a menudo se asume en la literatura [10].

### 3.2.2 Online Clustering

Una vez que cada  $I_t$  ha sido dividido en bloques no solapados, a su vez estos bloques, en el módulo de *Online Clustering* [39], se agrupan en *clusters*, según su posición y su similitud a lo largo del tiempo. Por lo tanto, las particiones de *clusters*  $\mathcal{L}^s = \{\mathbf{C}_{\mathbf{q}}^s\}_{\mathbf{q}=1:|\mathcal{L}^s|}$  se crean para cada localización de bloque  $\mathbf{s}$ , donde  $C_q^s$  es el bloque que representa el *cluster* y por tanto el candidato

a fondo. Con el fin de simplificar la notación, se omite en lo que resta de sección el índice  $\mathbf{s}$ , pues todas las operaciones tienen lugar en la misma ubicación.

Como el objetivo es identificar estacionaridad, los bloques que contienen objetos o personas en movimiento no son útiles para caracterizar el fondo. Por lo tanto, el *matching* entre bloques de imágenes consecutivas se realiza primero descartando los bloques  $R_t$  que no correspondan con  $R_{t-1}$ . Para cada  $R_t$  sin movimiento, esta fase determina los *matchings* del bloque con otro bloque de la misma posición almacenado en la partición  $\mathcal{L}$ . Cada *cluster* modela un patrón espacio-temporal de la escena y se describe a través de su repetibilidad  $\omega_q$  y el bloque que representa el *cluster*  $C_q$ . Para actualizar  $\mathcal{L}$ , si no se encuentra *matching* alguno con los candidatos almacenados se crea un nuevo *cluster*  $C_{q'} = R_t$  con el representante, donde  $q' = |\mathcal{L}| + 1$ . Por el contrario, si hay *matching*, el *cluster* es actualizado utilizando una media ponderada:

$$C_q = \frac{C_q^{match} \cdot \omega_q + R_t}{\omega_q + 1}, \quad (3.2)$$

donde  $C_q^{match}$  es el *cluster* coincidente y  $\omega_q$  es la repetibilidad del bloque que se incrementa en 1 cada vez que se visualiza dicho *cluster*. En ocasiones  $R_t$  puede coincidir con más de un candidato de  $\mathcal{L}$ , seleccionándose  $C_q^{match}$  de la siguiente manera:

$$C_q^{match} = \underset{\forall C_q \in \mathbb{C}}{\operatorname{argmax}} f(C_q, R_{t-T:t}), \quad (3.3)$$

donde  $\mathbb{C} = \{C_q : C_q \text{ matches } R_t \forall q\}$ . Por lo tanto, el *matching* seleccionado para usar en la ecuación 3.2 tiene la repetibilidad más alta.

La distancia entre bloques elegida para el *clustering* está basada en el ratio entre dos bloques, tal y como se detalla en la Ecuación 2.5 de la Sección 2.3. Con las medidas de la media y la varianza, se puede obtener, respectivamente, la intensidad y la homogeneidad de la variación entre los bloques  $R$  y  $R'$ . Como se observa en la Figura 3.2 a mayor valor de la media  $d_\mu(R, R')$ , la intensidad del cambio es más alta y a menor valor de la distancia  $d_{\sigma^2}(R, R')$  la homogeneidad entre los bloques será mayor. El *matching* entre bloques  $R$  y  $R'$  está modelado por una SVM, que ha sido entrenada previamente con diferentes casos entre bloques. Por lo tanto en el *Online clustering*, como se ha visto en la Sección 2.3  $d_\mu(R_t, R_{t-1})$  y  $d_{\sigma^2}(R_t, R_{t-1})$  se utilizan para encontrar *matchings* entre  $R_t$  y  $R_{t-1}$  cuando se descarta movimiento entre los bloques, mientras que  $d_\mu(R_t, C_q)$  y  $d_{\sigma^2}(R_t, C_q)$  explicadas en las Ecuaciones 2.7 y 2.8 respectivamente, se obtienen para determinar el *matching* entre  $R_t$  y los miembros  $C_q$  de  $\mathcal{L}$ .

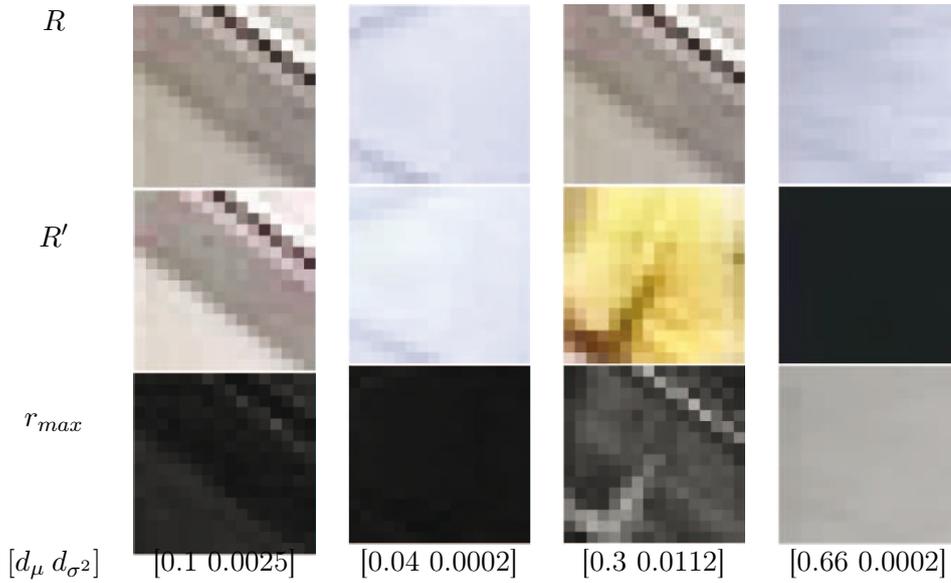


Figura 3.2: Ejemplo medida de *matching* basada en el ratio entre bloques.  $R$  y  $R'$  son dos bloques y  $d_\mu$  y  $d_{\sigma^2}$  son la distancias basadas en el ratio.

### 3.3 Análisis espacial

En este módulo se obtiene cada bloque del fondo  $B$  seleccionando el mejor candidato  $C_*^s$  entre el conjunto de candidatos del fondo  $C_q^s$ . Para cada posición, se estima una reconstrucción del fondo usando múltiples caminos, se fuerza a la reconstrucción a buscar una relación de homogeneidad y continuidad espacial en los bordes, con los bloques adyacentes del vecindario. El proceso de reconstrucción está dividido en dos etapas: selección de semillas, que utiliza múltiples reconstrucciones parciales, y reconstrucción multicamino que explora múltiples trayectos de reconstrucción[16].

#### 3.3.1 Selección de semillas: Multisemilla

##### 3.3.1.1 Cálculo de semillas

Esta fase del proceso de análisis espacial realiza una estimación parcial del fondo proporcionando así un conjunto de semillas para comenzar la reconstrucción del fondo. Estas semillas son candidatos de fondo altamente fiables temporalmente que se calculan basándose en la aproximación propuesta en [16].

Se hace un análisis unificado de la estacionariedad y la actividad de movimiento a lo largo de la secuencia de entrenamiento. Se detectan ubicaciones espaciales  $\mathbf{s}$  con una baja actividad de movimiento o sin objetos estacionarios que serán considerados candidatos adecuados para comenzar la reconstrucción del fondo como semillas. Para tal detección, se va a suponer que los

objetos estacionarios que ocuyen el fondo en  $I_1$  van a permanecer en el mismo lugar en  $I_T$ . Es razonable tener en cuenta esta suposición, ya que si un objeto no se mueve a lo largo de toda la secuencia de entrenamiento se podrá considerar como parte del fondo. Por lo tanto, se calculará una puntuación de actividad a nivel de bloque  $\xi^s$  de la siguiente manera:

$$\xi^s = \max \{ f(I_1^{\mathbf{P}}, \mathbb{F}^{\mathbf{P}} \setminus \{I_1^{\mathbf{P}}\}) + f(I_T^{\mathbf{P}}, \mathbb{F}^{\mathbf{P}} \setminus \{I_T^{\mathbf{P}}\}) \}_{\forall \mathbf{p} \in \mathbf{s}}, \quad (3.4)$$

donde  $\mathbf{p}$  es la localización espacial;  $\mathbb{F}^{\mathbf{P}}$   $I_1^{\mathbf{P}}$  y  $I_T^{\mathbf{P}}$  son los valores de píxel en los canales R, G y B en la localización  $\mathbf{p}$  de la secuencia de entrenamiento, la imagen inicial de la secuencia y la última imagen, respectivamente.  $\mathbb{F}^{\mathbf{P}} \setminus \{I_1^{\mathbf{P}}\}$  y  $\mathbb{F}^{\mathbf{P}} \setminus \{I_T^{\mathbf{P}}\}$  son la secuencia de entrenamiento sin la imagen inicial y sin la imagen final, respectivamente. Se puede ver como queda la puntuación del fondo en la Figura 3.3. La función  $f(\cdot, \cdot)$  calcula la media de los valores absolutos de las diferencias a nivel de píxel:

$$f(I_t^{\mathbf{P}}, \mathbb{I}^{\mathbf{P}}) = \frac{1}{|\mathbb{I}^{\mathbf{P}}|} \sum_{q=1}^{|\mathbb{I}^{\mathbf{P}}|} \begin{cases} 1 & \text{si } |I_t^{\mathbf{P}} - I_q^{\mathbf{P}}| > \tau \\ 0 & \text{resto} \end{cases}, \quad (3.5)$$

donde  $\mathbb{I}^{\mathbf{P}} = \{I_q^{\mathbf{P}}\}$  es un set genérico de píxeles con localización  $\mathbf{p}$  y  $\tau$  es un umbral de detección calculado automáticamente [38]. La siguiente puntuación de actividad  $f(I_1^{\mathbf{P}}, \mathbb{F}^{\mathbf{P}} \setminus \{I_1^{\mathbf{P}}\})$  compara los píxeles del primer *frame* con los píxeles de los demás *frames*, menos el último. Análogamente,  $f(I_T^{\mathbf{P}}, \mathbb{F}^{\mathbf{P}} \setminus \{I_T^{\mathbf{P}}\})$  compara los píxeles del último *frame* con los píxeles del resto de *frames*. Finalmente, la estimación inicial del fondo con semillas  $S^s$  se obtiene solo de las localizaciones con el mínimo  $\xi^s$ :

$$S^s = \begin{cases} \hat{C}_q^s & \text{si } \xi^s = \min \{ \xi^{s'} \}_{\forall s' \in q} \\ \emptyset & \text{resto} \end{cases}, \quad (3.6)$$

donde  $\hat{C}_q^s$  es el *cluster* temporalmente dominante en  $\mathbf{s}$ . La Figura 3.3 muestra un ejemplo de los mapas de actividad cuando las localizaciones con el mínimo  $\xi^s$  componen las semillas  $S^s$ . De esta manera se calculan semillas en las secuencias de vídeo, no obstante en la Subsección 3.3.1.2 se va a presentar una nueva alternativa para ampliar el número de semillas y por ende la información correcta disponible para comenzar la reconstrucción final.

### 3.3.1.2 Esquema Multisemilla

Esta etapa extiende el concepto de la reconstrucción multicamino presentado en [16] a la selección de semillas. La idea detrás de la reconstrucción multicamino se basa en la hipótesis de que la idoneidad de la reconstrucción puede ser altamente dependiente de la posición espacial por la que se llegue a reconstuir los bloques. El esquema multicamino propuesto en [16] se enfrenta a esta problemática localmente, realizando una reconstrucción espacial que explora diferentes caminos. De forma complementaria, lo técnica mutisemilla propuesta en este trabajo toma un enfoque

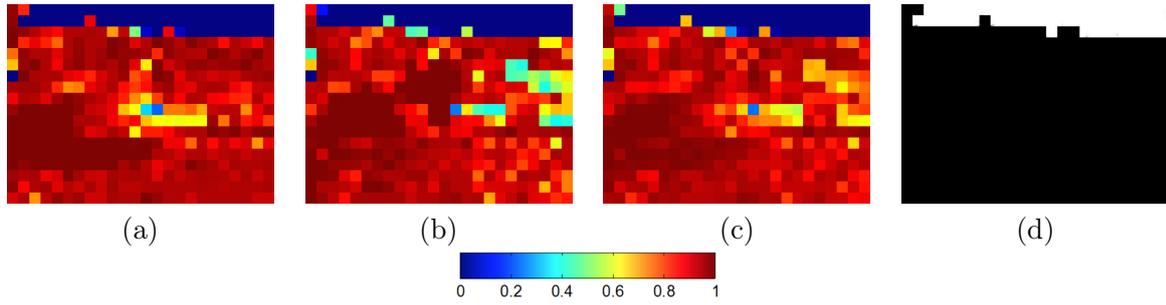


Figura 3.3: Ejemplo de selección de semilla. De izquierda a derecha, (a) corresponde al mapa de actividad  $f(I_1^{\mathbb{P}}, \mathbb{F}^{\mathbb{P}} \setminus \{I_1^{\mathbb{P}}\})$ , (b) corresponde al mapa de actividad  $f(I_T^{\mathbb{P}}, \mathbb{F}^{\mathbb{P}} \setminus \{I_T^{\mathbb{P}}\})$ , (c) es el mapa de actividad  $\xi^s$  y (d) es el mapa de semillas generado.

global de exploración de distintos caminos. Para ello, se ha realizado una división del mapa de semillas en superbloques, se puede ver en la Figura 3.4, en cuatro cuadrantes o superbloques que se adaptan al tamaño de la imagen. El objetivo del algoritmo es forzar a la reconstrucción espacial a utilizar únicamente las semillas pertenecientes a cada superbloque, creando así cuatro reconstrucciones diferentes que comienzan desde las diferentes posiciones de las semillas  $S^{s,z}$ , donde  $z$  denota el superbloque. Los bloques comunes en las cuatro reconstrucciones, forman un nuevo mapa de semillas con más información.

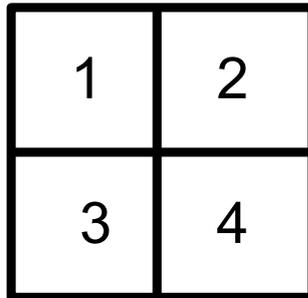


Figura 3.4: División del fondo en cuatro superbloques.

En primer lugar, se calculan las semillas globales  $S^s$  igual que en [16] y además se calculan los cuatro nuevos mapas de semillas  $\{S^{s,z}\}_{z=1\dots 4}$  para cada superbloque  $z$ . Todo lo que no corresponde al superbloque se mantiene vacío, si bien las semillas  $S^s$  se mantienen en un estado latente. Dicho estado latente consiste en mantener las posiciones  $s$  de las semillas globales vacías

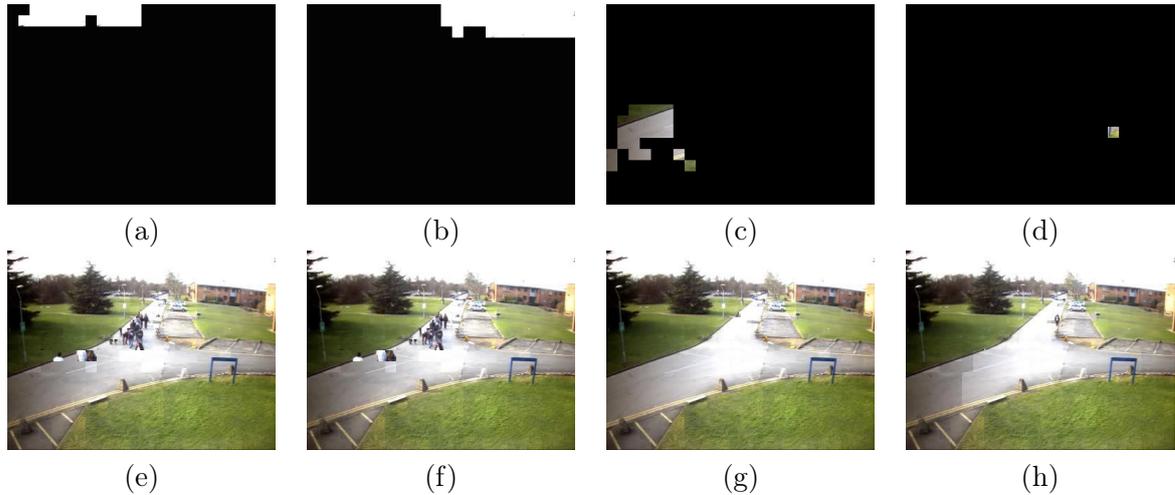


Figura 3.5: Ejemplo esquema multisevilla. De izquierda a derecha, primera fila: (a), (b), (c) y (d) representan, respectivamente, las semillas de cada superbloque  $S^{s,1}$ ,  $S^{s,2}$ ,  $S^{s,3}$  y  $S^{s,4}$ ; segunda fila: (e), (f), (g) y (h) representan, respectivamente, las estimaciones de fondo  $B^{s,1}$ ,  $B^{s,2}$ ,  $B^{s,3}$  y  $B^{s,4}$  calculadas a partir de las semillas de cada superbloque.

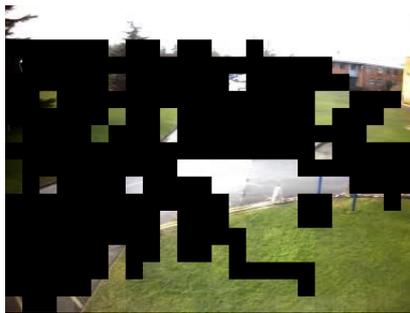


Figura 3.6: Mapa de semillas final  $S_*^s$ , formado por los bloques comunes en las cuatro reconstrucciones  $B^{s,z}$ .

en el fondo, pero modificar el conjunto de candidatos  $\mathcal{L}^s$ , de forma que  $|\mathcal{L}^s| = 1$  pues  $S^s$  es el único válido. Esta configuración permite que posiciones fiables no afecten el esquema multisevilla que busca forzar la reconstrucción desde distintos superbloques, a la vez que se aprovecha la información fiable disponible al máximo.

Para calcular las semillas de cada cuadrante, se aplica el algoritmo de la Subsección 3.3.1.1, que calcula  $S^s$ , a la región de cada superbloque. Es importante tener en cuenta, que las semillas de cada superbloque que no coinciden con las semillas globales son más susceptibles a ser erróneas y, por ende, menos fiables. La Figura 3.5 muestra en su primera fila un ejemplo de las 4 semillas seleccionadas para cada superbloque.

Una vez se tienen los superbloques con al menos una semilla cada uno, se reconstruye el fondo con el algoritmo detallado en la Sección de 3.3.2 de forma que se obtienen cuatro estimaciones

del fondo  $B^{s,z}$ , como se muestra en la segunda fila de la Figura 3.5. En esta fase del algoritmo se construye de un nuevo mapa de semillas  $S_*$  (ver Figura 3.6), en el que quedan incluidas las semillas iniciales y además se consideran semillas aquellos bloques que sean comunes en las cuatro reconstrucciones  $B^{s,1}$ ,  $B^{s,2}$ ,  $B^{s,3}$  y  $B^{s,4}$ .

### 3.3.2 Reconstrucción Secuencial Multicamino

Esta sección explica el funcionamiento de la Reconstrucción Secuencial Multicamino (*Sequential Multipath Reconstruction*, SMR) propuesta en [16]. Aunque se mantiene la operatividad de SMR, en este trabajo se han modificado las medidas de continuidad empleadas.

SMR toma como partida el mapa de semillas halladas en la Subsección 3.3.1.2, de forma que la estimación parcial del fondo  $\tilde{B} = S_*$ ,  $\forall s$ . Si se considera el índice de posición de bloque  $\mathbf{s}$  como un vector bidimensional ( $B^s = B^{(i,j)}$ ) el vecindario de los cuatro vecinos con los que colinda sería  $\mathbb{V}_4^s$  definido como:

$$\mathbb{V}_4^s = \left\{ B^{(i-1,j)}, B^{(i,j+1)}, B^{(i+1,j)}, B^{(i,j-1)} \right\}, \quad (3.7)$$

y donde el vecindario de ocho bloques es:

$$\mathbb{V}_8^s = \left\{ B^{(i-1,j)}, B^{(i-1,j+1)}, B^{(i,j+1)}, B^{(i+1,j+1)}, \right. \\ \left. B^{(i+1,j)}, B^{(i+1,j-1)}, B^{(i,j-1)}, B^{(i-1,j-1)} \right\}. \quad (3.8)$$

SMR lleva a cabo una reconstrucción iterativa, en la que en cada iteración se elige un bloque del fondo  $\tilde{B}^s$  con el máximo número de bloques no vacíos en el vecindario  $\mathbb{V}_8^s$ , donde los bloques vacíos serán reconstruidos empleando  $\mathbf{m} = 1 \dots 8$  caminos o hipótesis. Cada camino comienza desde un bloque  $\tilde{B}^s$  y sale de dicho bloque hacia arriba, derecha, izquierda o abajo, empleando una dirección (a favor o en contrade las agujas del reloj) y secuencialmente va completando las localizaciones vacías  $\mathbf{s}' \in \mathbb{V}_8^s$  con los candidatos almacenados en el *clustering*  $C_q^{s'}$ . La Figura 3.7 (a) muestra la selección del bloque  $\tilde{B}^s$  cuyo vecindario  $\mathbb{V}_8^s$  se explora utilizando los ocho caminos descritos en la Figura 3.7(b). Algunos de los bloques del trayecto ya existen, por lo que no se reconstruirán. La Figura 3.7 (c) muestra un ejemplo donde algunos vecinos de  $\tilde{B}^s$  todavía están vacíos.

Para cada uno de los trayectos de  $\mathbf{m}$ , se seleccionan los candidatos idóneos para ocupar espacios vacíos mediante el empleo de una función de continuidad espacial basada en dos medidas diferentes de distancia entre bloques que se combinan para tener en cuenta ambas informaciones. Por un lado, se calcula la distancia euclídea  $d_e$  entre los bordes de cada bloque adyacente en el vecindario. Por otro lado, se calcula la distancia de covarianza  $d_c$  (ver Ecuación 2.2) entre el bloque fijo y el , para hacer una comparación que tenga en cuenta información interna del bloque. Para poder combinar las dos distancias de una manera real, estas se normalizan de manera que  $d_c, d_e \in [0, 1]$ . Una vez que se tienen las dos distancias se calcula la función densidad

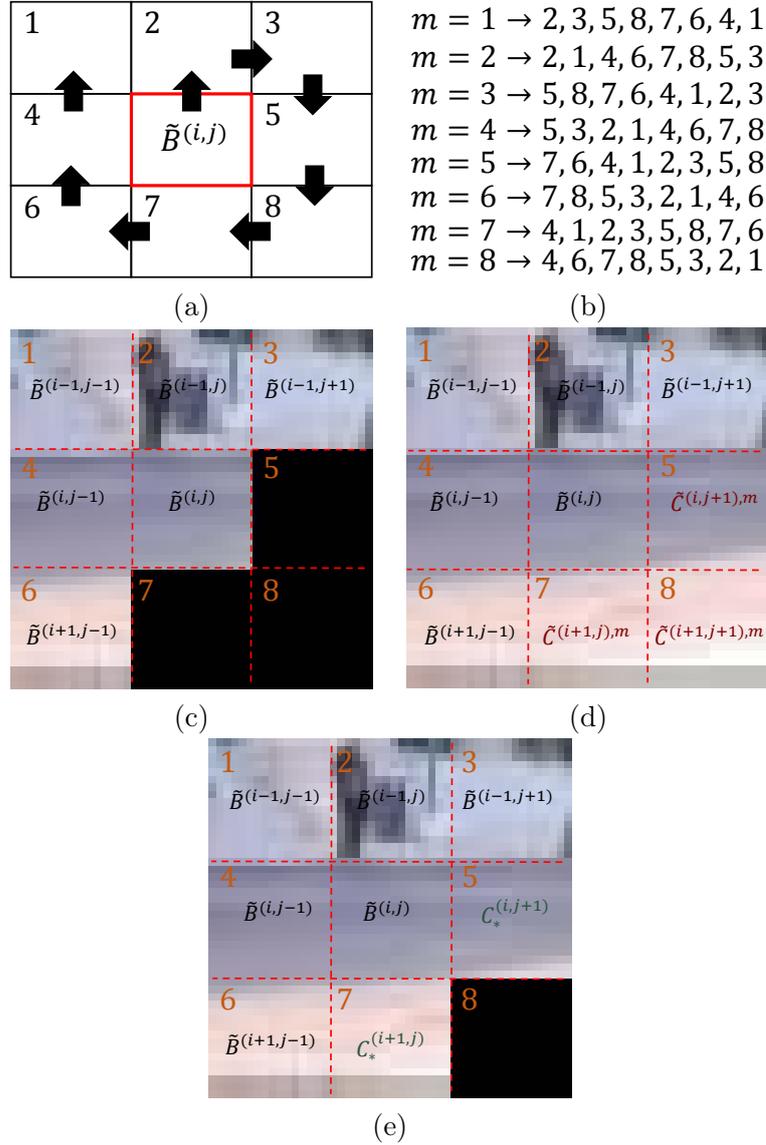


Figura 3.7: Esquema de reconstrucción multicamino. Se describe la imagen de izquierda a derecha y de arriba a abajo. La figura (a) muestra el primer camino para  $\mathbf{m} = 1$ . (b) los trayectos que realiza para reconstruir un bloque. La figura (c) muestra un ejemplo de reconstrucción para un bloque y su  $\mathbb{V}_8^s$ . (d) ejemplifica la reconstrucción de  $\mathbb{V}_8^s$  para (c) con  $\mathbf{m} = 1$ . (e) es la reconstrucción final del bloque para  $\mathbb{V}_4^s$ .

de probabilidad, de manera *offline*, para cada una de ellas, únicamente de los inliers, para después utilizar dichas distribuciones en el algoritmo y de esta manera combinar ambas distancias:

$$\Phi \left( C_q^{s'} \right) = \frac{1}{|\mathbb{V}_4^s|} \sum_{s'' \in \mathbb{V}_4^s} \left( \alpha \cdot fdp \left( \bar{d}_e \left( \mathbf{p}, \mathbf{p}' \right) \right) + \beta \cdot fdp \left( \bar{d}_c \left( R, R' \right) \right) \right), \quad (3.9)$$

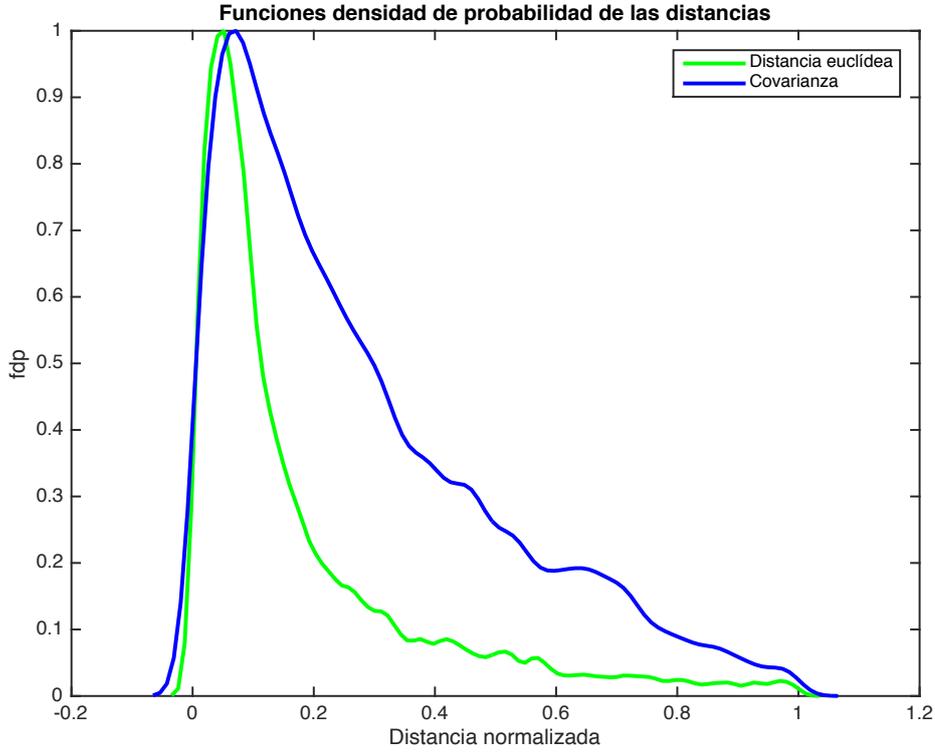


Figura 3.8: Funciones densidad de probabilidad de  $d_e$  y  $d_c$ .

donde  $\Phi$  es la nueva distancia combinada,  $\alpha$  y  $\beta$  son dos factores, donde  $\alpha + \beta = 1$ , utilizados para calcular el peso que tiene cada distancia en la distancia total combinada  $d_e^{\bar{}}(\cdot)$  y  $d_c^{\bar{}}(\cdot)$  son la distancia eculídea a nivel de pixel en la frontera normalizada y la distancia de covarianza a nivel de bloque normalizada, respectivamente. En la Figura 3.8 se muestran las dos funciones densidad de probabilidad.

Para cada trayecto  $\mathbf{m}$ , el candidato  $\tilde{C}_q^{s',m}$  se selecciona minimizando  $\Phi$ :

$$\tilde{C}_q^{s',m} = \underset{\forall q \in \{1, \dots, N^{s'}\}}{\operatorname{argmin}} \Phi \left( C_q^{s'} \right), \quad (3.10)$$

donde  $\mathbf{m} \in \{1, \dots, 8\}$  y  $C_q^{s'}$ ,  $\forall q$  son los candidatos disponibles. La Figura 3.7 (e) muestra la reconstrucción de  $\mathbb{V}_4^s$  comenzando desde la estimación inicial en la Figura 3.7 (c) donde la Figura 3.7 (d) muestra un paso temporal  $\mathbb{V}_8^s$  de la reconstrucción de un solo camino.

Finalmente, se obtiene la estimación del vecindario  $\mathbb{V}_4^s$  usando los ocho caminos  $\mathbf{m} = 1, \dots, 8$ . Para ello, se seleccionan los mejores candidatos  $C_*^{s'}$  entre los bloques temporales  $\tilde{C}^{s',m}$ .



Figura 3.9: Ejemplo de fondo reconstruido  $B$ .

$$C_*^{s'} = \underset{\forall m \in \{1, \dots, 8\}}{\operatorname{argmin}} \Phi(\tilde{C}^{s', m}), \quad (3.11)$$

donde  $\Phi(\tilde{C}^{s', m})$  es el valor de  $\Phi$  obtenido por los candidatos durante la reconstrucción multicamino (Ecuación 3.10). Como en los bloques temporales  $\mathbb{V}_4^s$  (centro-arriba, centro-derecha, centro-izquierda y centro-abajo) emplean tres fronteras y los bloques temporales en  $\mathbb{V}_8^s$  (arriba-izquierda, arriba-derecha, abajo-izquierda y abajo-derecha) emplean dos, sólo se van a seleccionar los candidatos  $C_*^{s'}$  de  $\mathbb{V}_4^s$  por tener mayor fiabilidad. Así que este proceso de selección de  $\tilde{B}^s$  y de reconstrucción de  $\mathbb{V}_4^s$  se repite hasta que se regenera el fondo completamente. En la Figura 3.9 se puede apreciar el resultado de la reconstrucción final.



## Capítulo 4

# Resultados experimentales

En este capítulo se exponen los experimentos realizados para el estudio y análisis del funcionamiento del algoritmo propuesto. El capítulo se divide en dos secciones: en primer lugar se expone el dataset utilizado y se explica la métrica utilizada para el análisis de resultados (Sección 4.1) y en segundo lugar se evalúa el algoritmo para las diferentes configuraciones y se expone frente a varias aproximaciones del estado del arte (Sección 4.2).

### 4.1 Marco de evaluación

#### 4.1.1 Dataset

Para la evaluación de resultados se han utilizado 29 secuencias y sus imágenes de fondo (*ground-truth*) propuestas en [16], de datasets públicos (TRECVID<sup>1</sup>, PBI [15], AVSS 2007<sup>2</sup>, LIRIS 2012 [40], CDNET [41], Wallflower [42], LIMU<sup>3</sup>, CUHK [43], COST211<sup>4</sup>, IDIAP [44], PETS 2009 [45], SAIVT-Campus [46]) que cubren diferentes escenarios y dificultades (Ver en la Figura 4.1), principalmente con objetos estacionarios y multitudes, que hace más compleja la detección del fondo, ya que estará visible cortos espacios de tiempo.

---

<sup>1</sup><http://trecvid.nist.gov/trecvid.data.html>

<sup>2</sup>[http://www.eecs.qmul.ac.uk/~andrea/avss2007\\_d.html](http://www.eecs.qmul.ac.uk/~andrea/avss2007_d.html)

<sup>3</sup><http://limu.ait.kyushu-u.ac.jp/dataset/en/>

<sup>4</sup><http://www.csd.uoc.gr/~tziritas/cost.html>

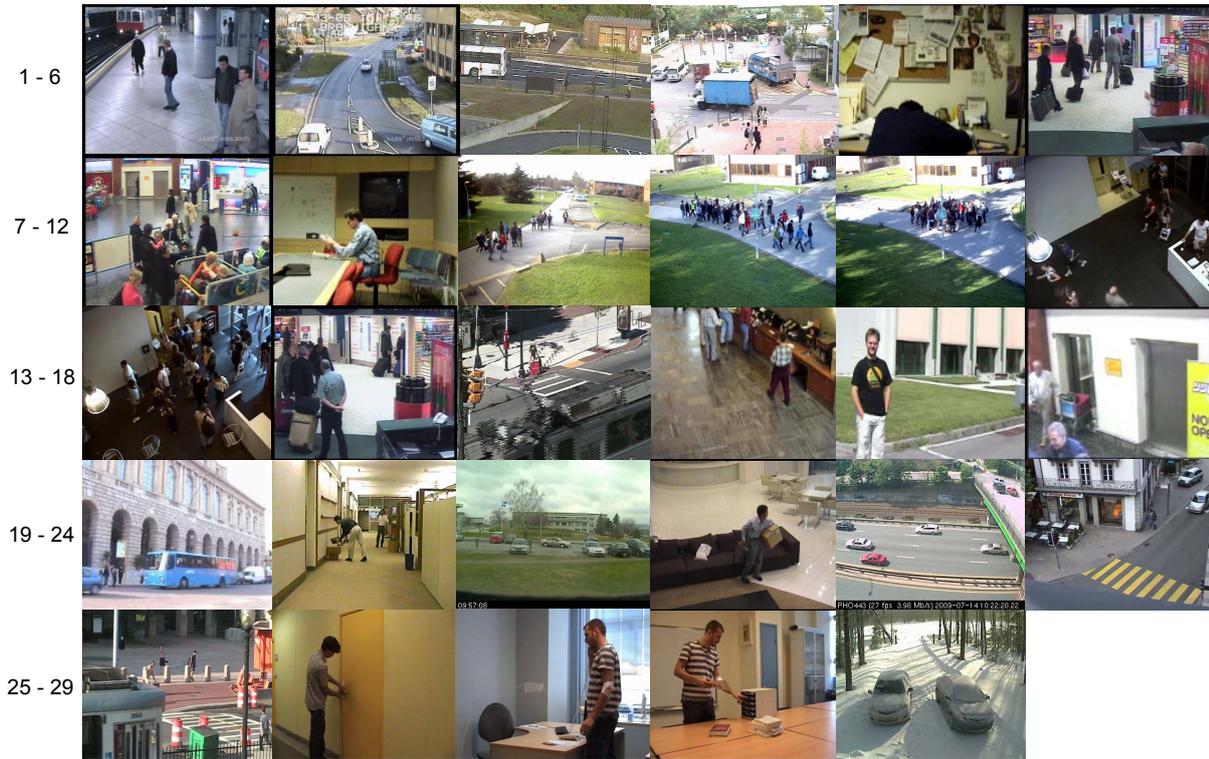


Figura 4.1: Dataset utilizado para la evaluación.

En la Tabla 4.1 se describen las propiedades de los vídeos en términos de: tamaño y duración; estacionaridad del frente; visibilidad; duración y tiempo en el que el fondo está libre de objetos; sombras y cambios de iluminación; factores fotométricos.

#### 4.1.2 Métricas

Para evaluar los resultados se han utilizado los seis indicadores propuestos en [1] para evaluar la tarea de Inicialización de fondo. Las métricas se definen como sigue:

- *Average Gray-level Error (AGE)*: Es el promedio de la diferencia absoluta a nivel de grises entre las imágenes de GT y de fondo estimado. Sus valores oscilan en  $[0, L - 1]$ , donde  $L$  es el número máximo de niveles de gris; cuanto menor sea el valor de AGE, mejor es la estimación de fondo.
- *Average Error Pixels (AE)*: Un error de píxel (EP) es un píxel del donde reconstruido cuyo valor difiere del valor del píxel correspondiente a GT por más de un umbral  $th$  (en los experimentos se ha sugerido el valor  $th = 20$ ). AE es la relación entre los EPs y el número de píxeles de la imagen. Sus valores oscilan en  $[0, 1]$ ; cuanto menor sea el valor de las PEP, mejor es la estimación de fondo.

ID	Video	Dataset	#f	T	S	V	SI
1	<i>AB_H</i>	AVSS 2007	400	I	H	M	M
2	<i>PV_E</i>	AVSS 2007	500	I	H	L	M
3	<i>BSM</i>	LIMU	400	O	H	L	L
4	<i>SQ</i>	CUHK	500	O	H	L	L
5	<i>FGA</i>	Wallflower	400	I	H	L	L
6	<i>TREC1</i>	TRECVID	498	I	H	H	M
7	<i>TREC2</i>	TRECVID	699	I	L	H	M
8	<i>MO</i>	Wallflower	300	I	H	L	L
9	<i>PETS1</i>	PETS 2009	221	O	L	H	H
10	<i>PETS2</i>	PETS 2009	240	O	M	H	H
11	<i>PETS3</i>	PETS 2009	378	O	H	H	M
12	<i>Test</i>	SAIVT Campus	500	I	L	M	M
13	<i>Train</i>	SAIVT Campus	500	I	L	H	H
14	<i>TREC3</i>	TRECVID	400	I	M	M	M
15	<i>AB_Box</i>	CDNET	500	O	H	M	L
16	<i>bootstrap</i>	Wallflower	294	I	L	L	H
17	<i>ca_vignal</i>	PBI	258	O	M	L	L
18	<i>cam4</i>	TRECVID	300	I	M	L	L
19	<i>guardia</i>	PBI	400	O	H	M	L
20	<i>hall_m</i>	COST	300	I	M	M	L
21	<i>parking</i>	CDNET	400	O	H	L	L
22	<i>sofa</i>	CDNET	400	I	H	L	L
23	<i>st_light</i>	CDNET	400	O	H	H	L
24	<i>traffic</i>	IDIAP	500	O	H	L	L
25	<i>tramp</i>	CDNET	400	O	H	H	L
26	<i>vid16</i>	LIRIS 2012	380	I	H	L	L
27	<i>vid22</i>	LIRIS 2012	345	I	M	M	L
28	<i>vid36</i>	LIRIS 2012	128	I	M	M	L
29	<i>winter</i>	CDNET	500	O	H	L	M

Tabla 4.1: Descripción del dataset. #f: Número de *frames*. T: Tipo. I: Interior. O: Exterior. S: Dificultad de la región estacionaria. V: Complejidad para visualizar el fondo libre de objetos. SI: Complejidad de las sombras y los cambios de iluminación. L, M y H significan niveles bajo, medio y alto, respectivamente.

- *Average Clustered Error Pixels* (ACE): Un error de píxel en el *cluster* (CEP) se define como cualquier píxel de error cuyos cuatro vecinos adyacentes también son píxeles erróneos. ACE es la relación entre los CEPs y el número de píxeles de la imagen. Sus valores oscilan en  $[0, 1]$ ; cuanto menor es el valor ACE, mejor es la estimación de fondo.
- *Peak-Signal-to-Noise-Ratio* (PSNR): Se define como  $PSNR = 10 \log_{10} \left( \frac{2^{(L-1)}}{MSE} \right)$ , donde L es el número máximo del nivel de gris y el MSE es el error cuadrático medio entre el GT

y las imágenes de difusión. Se asumen valores en decibelios; cuanto mayor sea el valor de PSNR, mejor es la estimación de fondo.

- *Multi-Scale Structural Similarity Index* (MS-SSIM): Esta es la métrica definida por [47], que utiliza la distorsión estructural como una estimación de la distorsión visual percibida. Se asume valores entre  $[0, 1]$ ; cuanto mayor sea el valor de MS-SSIM, mejor es el fondo estimado.
- *Color image Quality Measure* (CQM): Esta es la métrica propuesta por [48], basado en una transformación reversible del espacio de color YUV y en la PSNR calculada en las bandas de YUV individuales. En cuanto a la PSNR, asume valores en decibelios; cuanto mayor sea el valor CQM, mejor es la estimación de fondo.

Adicionalmente, para evaluar la selección de semilla se ha utilizado *Reconstructed Percentage* (RP) que indica el porcentaje del fondo reconstruido para una reconstrucción parcial del mismo.

## 4.2 Evaluación

### 4.2.1 Análisis temporal

#### 4.2.1.1 Medidas de comparación entre bloques

Una parte importante del proyecto es explorar las capacidades de las medidas presentadas en la Sección 2.3 del del Capítulo 2 para medir distancias entre bloques. Para ello, se han entrenado diversos modelos con el fin de distinguir si dos bloques son iguales o no, en concreto se ha trabajado con SVM (*Support Vector Machine*) como algoritmo de entrenamiento y test.

En primer lugar, para poder entrenar una SVM ha sido necesario etiquetar manualmente numerosas muestras de comparaciones entre bloques, fijando la clase distintos (0) y la clase iguales (1). Las muestras etiquetadas han sido recopiladas de distintas situaciones entre las que se encuentran: sombras y cambios de iluminación, camuflajes, objetos en movimiento o casos sustancialmente diferentes. La selección de las muestras se ha dividido, para los casos citados, en bloques de diferentes tamaños, donde  $W = \{8, 16, 32\}$ , obteniendo 1500 muestras para  $W = 16$  (muestras de entrenamiento) y 500, 1500 y 500 para test de tamaño 8, 16 y 32, respectivamente. La SVM se ha entrenado con las muestras de  $W = 16$ , para posteriormente y con el modelo entrenado, testear las muestras para los diferentes tamaños de bloque. En cuanto a los modelados de bloque utilizados, se ha empleado el ratio con el espacio de color RGB [39]; la distancia de histogramas [35][36] para los espacios de color RGB, HSV y LAB; la distancia de covarianza usando la mejor configuración propuesta en [36] (nivel de grises, información de la primera y segunda derivada y módulo y fase del gradiente); y por último la correlación y la distancia euclídea a nivel de bloques [13], también para el espacio de color RGB. Para este experimento,

Modelo para $W = 8$	P	R	F	ED	EI
Ratio RGB	0.932	0.698	0.798	0.056	0.301
Histograma RGB	0.603	0.996	0.751	0.731	0.003
Histograma HSV	0.565	0.992	0.7203	0.853	0.007
Histograma LAB	0.554	0.989	0.710	0.890	0.010
Covarianza	0.875	0.789	0.829	0.126	0.210
SCL	0.969	0.349	0.513	0.012	0.650

Tabla 4.2: Tabla PRF para  $W = 8$  píxeles.

Modelo para $W = 16$	P	R	F	ED	EI
Ratio RGB	0.913	0.801	0.853	0.094	0.198
Histograma RGB	0.894	0.617	0.730	0.090	0.382
Histograma HSV	0.823	0.586	0.685	0.156	0.413
Histograma LAB	0.853	0.556	0.674	0.118	0.443
Covarianza	0.887	0.906	0.897	0.143	0.093
SCL	0.967	0.570	0.717	0.023	0.429

Tabla 4.3: Tabla PRF para  $W = 16$  píxeles.

se han considerado diferentes métricas de evaluación: *Precision* ( $P$ ), *Recall* ( $R$ ), *F-score* ( $F$ ), error de clase 1 clasificado como 0 (EI), y error de clase 0 clasificado como 1 (ED). Se distingue entre los errores de ambas clases, debido a que es interesante desglosar el error, pues clasificar un bloque distinto como igual puede llevar a importantes errores.

En las tablas 4.2, 4.3 y 4.4 se observa el estudio para bloques de tamaño 8, 16 y 32 píxeles, respectivamente. Aunque atendiendo a  $F$  la distancia que mejor funciona es la basada en covarianza, seguida muy de cerca por el Ratio RGB, si se mira el ED se puede ver que es superior en la distancia basada en matrices de covarianza. El error de tipo ED es crítico en esta tarea y se ha verificado como integrando la medida en un algoritmo completo de inicialización de fondo los errores eran numerosos. Además, se puede apreciar bastante estabilidad entre los distintos tamaños de bloque para los rendimientos obtenidos.

Debido a los buenos resultados del ratio RGB, se ha decidido trabajar con esta medida para un tamaño de bloque de  $16 \times 16$  e intentar disminuir el error ED. Para la disminución del error se han explorado dos vías de forma individual y conjunta: introducir costes en los errores al entrenar la SVM y eliminar las muestras asociadas a sombras y cambios de iluminación que son las muestras de clase 1 (igual) más cercanas a la clase 0 (distinto). La matriz de costes empleada es la siguiente  $M_c = \begin{pmatrix} 0 & C_{ED} \\ C_{EI} & 0 \end{pmatrix}$ , donde para penalizar el error ED se ha fijado  $C_{ED} = 10$ . En la Tabla 4.5 se muestran los resultados obtenidos para las combinaciones con/sin muestras y con/sin costes. Se puede observar que el valor más bajo de ED, manteniendo un valor razonable de EI, tiene lugar para el modelo sin muestras de cambios de iluminación y con costes. Por tanto,

Modelo para $W = 32$	P	R	F	ED	EI
Ratio RGB	0.918	0.951	0.934	0.100	0.048
Histograma RGB	0.587	0.840	0.691	0.703	0.160
Histograma HSV	0.885	0.444	0.591	0.068	0.555
Histograma LAB	0.910	0.360	0.515	0.042	0.640
Covarianza	0.728	1.000	0.842	0.444	0.000
SCL	0.861	0.831	0.846	0.158	0.168

Tabla 4.4: Tabla PRF para  $W = 32$  píxeles.

Ratio RGB, $W = 16$	P	R	F	ED	EI
Con muestras, sin costes	0.901	0.837	0.868	0.114	0.162
Sin muestras sin costes	0.861	0.986	0.919	0.195	0.014
Sin muestras, con costes	0.940	0.830	0.882	0.064	0.169
Con muestras, con costes	0.997	0.5811	0.734	0.0016	0.418

Tabla 4.5: Tabla comparativa en función de las muestras de cambios de iluminación y matriz de costes.

se ha utilizado el anterior modelo para llevar a cabo un *clustering* temporal de bloques.

#### 4.2.1.2 Reconstrucción ideal

Para un análisis exhaustivo de la fase de *clustering*, en esta sección se ha realizado una extracción de los candidatos del fondo de todas las secuencias para reconstruir un fondo a partir de la comparativa de cada candidato del *clustering* con el bloque correspondiente del GT, eligiendo de esta forma el candidato que obtenga una puntuación menor a nivel de bloque, en la diferencia entre bloques. En este experimento se busca analizar si la fase de *Online Clustering* funciona correctamente, es decir, si entre los candidatos almacenados en el *cluster*, se encuentra el correcto para formar parte del fondo. . El resultado visual se muestra en la Figura 4.2 y el cualitativo en la Tabla 4.6.

A la vista de los resultados, a nivel visual se aprecian errores de generación del fondo ideal, aunque en algunos casos no parecen excesivamente significativos. Sin embargo, a nivel cualitativo, en la Tabla 4.6 se muestra como hay porcentajes de error muy elevados con los mejores candidatos para el fondo almacenados en el *cluster*. Comenzar la etapa de selección de semillas con un error medio de base de  $AE(\%) = 4,13$  es un problema, ya que este error es el mínimo que se conseguirá en caso de que la reconstrucción funcione perfecta. Por lo que se puede concluir que el funcionamiento del *Online Clustering* umbralizado con una SVM entrenada no selecciona correctamente los candidatos y este error se propaga por todo el algoritmo.



Figura 4.2: Ejemplo comparativo de GT y fondo ideal, en primera y segunda fila, respectivamente.

Vídeo	AE(%)	Vídeo	AE(%)
1	5.92	16	9.97
2	15.98	17	0.00
3	9.97	18	15.50
4	3.59	19	0.32
5	1.95	20	3.18
6	4.68	21	4.66
7	1.60	22	0.25
8	7.52	23	4.19
9	0.99	24	1.60
10	1.27	25	7.55
11	2.03	26	1.14
12	0.84	27	1.24
13	1.57	28	1.19
14	1.05	29	3.10
15	6.38	MEDIA	4.13

Tabla 4.6: Tablas del porcentaje de error en el fondo reconstruido con los mejores candidatos del *Clustering*. La columna de Vídeo corresponde al número de la secuencia vídeo.

#### 4.2.2 Selección de semilla

En esta sección se han realizado experimentos para estimar la calidad de la selección de semillas al haber utilizado el esquema multisevilla. Es relevante saber cuál de las configuraciones posibles ofrece un mapa de semillas con mayor información y precisión. Para ello se ha hecho un estudio comparativo de resultados en el que se calcula el mapa de semillas de cada configuración, y de esta manera compararlo con un mapa de semillas GT, adaptado a cada uno de ellos. En la Tabla 4.7 se comparan las diferentes configuraciones de la selección de semilla (multisevilla) en función de dos características: el número de superbloques en los que se divida la imagen y el

Configuración	Factores		AGE	AE	ACE	MS-SSIM	PSNR	CQM	RP
	$\alpha$	$\beta$							
1x1	-	-	0.149	0.0005	0	0.9996	55.628	84.576	4.62
2x2	1	0	4.3926	0.0328	0.0108	0.9796	29.289	39.290	38.74
2x2	0.5	0.5	3.838	0.0317	0.0097	0.9848	30.427	40.373	36.93
2x2	0	1	5.7304	0.0562	0.0226	0.9624	26.577	37.099	44.84
3x3	1	0	2.6792	0.0181	0.0061	0.9866	31.875	41.813	26.03
3x3	0.5	0.5	2.186	0.0138	0.0025	0.9923	33.974	44.228	24.48
3x3	0	1	3.292	0.026	0.0089	0.9802	29.783	40.340	30.73

Tabla 4.7: Comparación de resultados Selección de semilla.

peso que tiene cada uno de los factores en la distancia combinada de ambos. Se han probado varias configuraciones, para la característica del número de superbloques se han realizado tres configuraciones diferentes: la primera, de  $1 \times 1$  que no introduce ninguna novedad en el estado del arte de [39], otra de  $2 \times 2$  y una última de  $3 \times 3$ . Se puede observar que sin ninguna división del espacio, la precisión es muy elevada, pero esto conlleva un inicio de reconstrucción con mucha menos información, es decir, el porcentaje reconstruido es muy bajo; sin embargo en una división  $2 \times 2$  del espacio, se comienza la reconstrucción con más de un tercio del fondo reconstruido, aunque el error medio (AE); así mismo, la división en nueve superbloques del espacio reduce el porcentaje reconstruido y aumenta la precisión de reconstrucción, esto se debe a que es menos probable que un bloque que añada error sea reconstruido nueve veces, por diferentes caminos, por lo que es más fiable, pero también más lento, ya que computacionalmente realiza muchas más reconstrucciones que los otros modelos. Por otra parte es importante el peso que tiene cada uno de los factores a la hora de realizar las reconstrucciones parciales, excepto con la configuración  $1 \times 1$  que no realiza ninguna reconstrucción parcial, ya que dependiendo del valor de estos se le dará más importancia a la continuidad espacial a nivel de frontera del bloque ( $\alpha$ ) o a nivel interno del bloque ( $\beta$ ). En el experimento se observa que el valor de los factores para la distancia de bloques es óptimo cuando  $\alpha = \beta = 0,5$ .

Al igual que en [16], se ha calculado el AE sobre el total de fondo reconstruido y el RP para evaluar el error y porcentaje de fondo reconstruido. En la Tabla 4.8 se muestran los resultados obtenidos para cada secuencia tanto para el algoritmo del PFC como para RMR [16], observándose una mejora sustancial tanto en RP como en AE en las semillas seleccionadas en el algoritmo propuesto. Esta mejora significa que es posible comenzar la reconstrucción espacial del fondo con mayor cantidad de información correcta.

### 4.2.3 Evaluación estimación de fondo

Este experimento se realiza para elegir el mejor modelo teniendo en cuenta las configuraciones del esquema multisevilla empleadas y para realizar una comparativa frente algoritmos del estado

del arte. En la Tabla 4.9 se muestra el rendimiento para las distintas configuraciones del esquema multisevilla y de las combinaciones de las medidas de continuidad empleadas en la reconstrucción multicamino SMR (ver Subsección 3.3.2). Por un lado las medidas de continuidad espacial se combinan usando  $d_\rho = \alpha \cdot d_e + \beta \cdot d_c$ , donde  $\alpha + \beta = 1$ . Por otro lado, el esquema multisevilla emplea configuraciones de suprabloques de  $1 \times 1$ ,  $2 \times 2$  y  $3 \times 3$ . La mejor combinación de medidas de distancia entre bloques es en la que los dos coeficientes tienen el mismo peso, es decir,  $\alpha = \beta = 0,5$ , empleándose así tanto información de los bordes de bloques adyacentes como del los bloques completos. Atendiendo a factores con el mismo peso, la configuración multisevilla que ofrece un mejor resultado es la división del espacio en  $2 \times 2$  superbloques. Por lo tanto, esta es la configuración que se ha elegido para el algoritmo propuesto en este PFC.

Por último, se ha comparado el algoritmo propuesto (PFC), frente a algoritmos relevantes de la literatura. En la Tabla 4.10 se presenta el algoritmo propuesto en un ranking junto con 14 algoritmos del estado del arte. Puede observarse que se ha conseguido un rendimiento cercano a los mejores algoritmos de la literatura. No obstante, los errores en el *clustering* y en las medidas de continuidad empleadas llevan a obtener un rendimiento peor. En la Figura 4.3 se muestra como a pesar de obtener una gran cantidad de información en la selección de semilla, la reconstrucción falla debido a las medidas de continuidad empleadas para reconstruir. Finalmente, en la Figura 4.4 se muestra una comparativa junto con los mejores algoritmos del estado del arte.

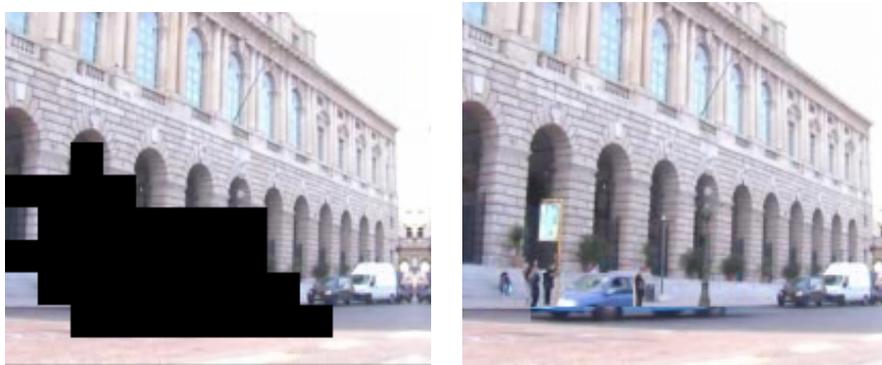


Figura 4.3: Ejemplo de evolución del algoritmo. De izquierda a derecha, la primera imagen es un mapa de semillas y la segunda imagen es el fondo reconstruido.

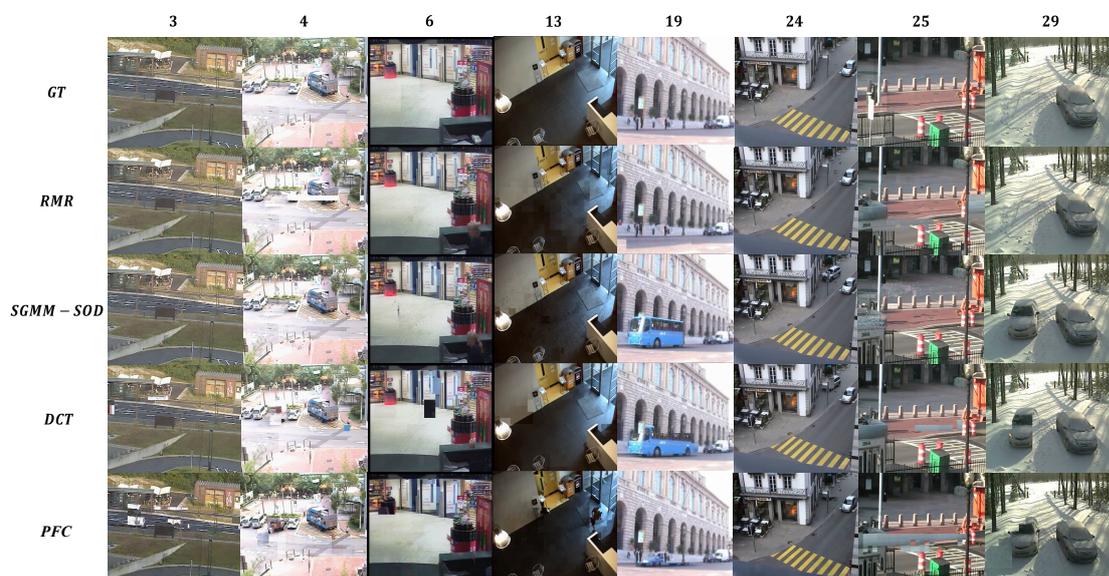


Figura 4.4: Comparativa con los mejores algoritmos del estado del arte. De izquierda a derecha, por columnas, las secuencias 3 (BSM), 4 (CUHK), 6 (TREC1), 19 (guardia), 24 (traffic), 29 (winter). Cada fila corresponde a un algoritmo (la primera corresponde al GT).

VÍdeo	RP		AE	
	RMR	PFC	RMR	PFC
1	12.80	31.37	1.92	0.11
2	4.11	17.64	3.40	0.35
3	13.33	29.35	0.00	0.15
4	34.30	41.78	0.2	0.07
5	1.25	30.00	0.017	0.04
6	24.88	30.35	0.00	0.18
7	2.17	18.87	0.00	0.20
8	1.25	31.25	0.00	0.04
9	56.71	46.80	0.00	0.09
10	26.62	24.30	0.16	0.05
11	18.29	20.13	0.00	0.09
12	18.18	47.72	0.00	0.00
13	10.10	14.64	0.00	0.18
14	15.22	32.95	0.89	0.04
15	12.35	30.03	0.00	0.07
16	3.75	14.99	0.00	0.20
17	11.11	92.30	3.13	0.00
18	5.00	80.00	0.00	0.20
19	46.15	74.82	0.00	0.00
20	31.82	18.78	0.10	0.05
21	59.00	56.38	0.00	0.02
22	45.00	53.33	0.00	0.00
23	14.00	45.35	0.00	0.03
24	17.87	37.68	0.00	0.04
25	0.21	19.95	0.00	0.12
26	1.69	8.21	0.00	0.16
27	40.58	18.35	0.00	0.50
28	0.24	41.78	0.00	0.04
29	23.67	62.00	0.00	0.11
MEDIA	19.02	36.93	0.33	0.10

Tabla 4.8: Tabla comparativa con el método RMR. En esta tabla se puede comprobar la relación entre el porcentaje de fondo reconstruido para ser utilizado como semilla y el porcentaje de error. El número de la primera columna equivale al número de la secuencia de vídeo correspondiente. Los colores verde, rojo y negro denotan un resultado mejor, peor e igual respectivamente.

Configuración	Factores		AGE	AE	ACE	MS-SSIM	PSNR	CQM
	$\alpha$	$\beta$						
1x1	1	0	14.67	0.152	0.0718	0.935	22.878	33.026
1x1	0.5	0.5	14.51	0.146	0.0651	0.895	22.898	33.199
1x1	0	1	16.69	0.189	0.0937	0.804	20.576	31.172
2x2	1	0	14.36	0.152	0.0709	0.9114	22.942	33.102
<b>2x2</b>	<b>0.5</b>	<b>0.5</b>	<b>13.85</b>	<b>0.142</b>	<b>0.0607</b>	<b>0.9085</b>	<b>23.223</b>	<b>33.621</b>
2x2	0	1	16.59	0.190	0.0948	0.8125	20.739	31.485
3x3	1	0	14.34	0.145	0.0637	0.9113	23.020	33.230
3x3	0.5	0.5	14.29	0.141	0.0607	0.9051	22.887	33.256
3x3	0	1	16.22	0.185	0.0906	0.8205	20.925	31.394

Tabla 4.9: Resultados de la reconstrucción de fondo, para todos los modelos. En negrita el algoritmo propuesto.

Algoritmo	AGE	AE	ACE	MS-SSIM	PSNR	CQM
RMR[16]	9.75	5.21	3.61	0.964	28.52	39.54
DCT[13]	12.81	10.47	7.09	0.905	26.25	37.50
SGMM-SOD[49]	16.19	13.34	9.83	0.884	25.73	35.52
<b>PFC</b>	<b>13.85</b>	<b>14.27</b>	<b>6.07</b>	<b>0.908</b>	<b>23.22</b>	<b>33.62</b>
RSM[12]	17.00	15.96	10.55	0.816	23.30	35.13
IMBS-1[50]	19.40	16.57	8.85	0.831	22.78	33.67
IMBS-2[51]	20.72	19.25	10.32	0.795	22.37	33.60
LOBSTER[52]	19.06	24.52	14.86	0.812	20.99	31.66
3dSOBS+[25]	22.17	25.95	20.78	0.772	21.92	35.94
MED[3]	21.31	27.19	22.39	0.806	23.41	37.27
SuBSENSE[18]	20.89	31.79	23.46	0.845	22.63	37.09
Fuzzy[53]	18.87	32.28	26.44	0.882	24.46	40.23
SC-SOBS[17]	22.91	35.26	24.91	0.810	21.00	36.77
FPCP[54]	22.53	46.34	40.84	0.891	21.59	-
LRGeomCG[55]	22.90	47.37	40.26	0.885	21.41	-

Tabla 4.10: En términos de comparación según las medidas de error propuestas en [1]. Para AGE, AE y ACE cuanto menor sea el valor, mejor es el método, mientras que para MS-SSIM, PSNR y CQM el mejor modelo es el que tenga el valor más alto. Los métodos se presentan en un ranking con un orden descendente en función de AE. En negrita, el algoritmo propuesto.

## Capítulo 5

# Conclusiones y trabajo futuro

### 5.1 Conclusiones

En este PFC se ha presentado un algoritmo de estimación de fondo de escena en secuencias de vídeo que trabaja a nivel de bloque y emplea criterios de continuidad espacial para reconstruir la imagen de fondo.

En primer lugar se estudian diferentes medidas para calcular la distancia entre bloques a la hora de agrupar los candidatos en el *Clustering Online*, donde finalmente se elige la distancia del ratio para el espacio de color RGB y se modela a través del entrenamiento de una SVM para evitar usar un umbral calculado empíricamente. Sin embargo, después de calcular los resultados de la evaluación, tanto a nivel visual como cuantitativo, se observa que este modelado resulta no ser todo lo eficiente que se esperaba, siendo las muestras etiquetadas como casos especiales (cambios de iluminación y sombras en concreto) un problema en el contexto de la inicialización de fondo. La generación de candidatos erróneos como consecuencia de la errónea agrupación de bloques, lleva a la reconstrucción a estar abocada al error por no existir el candidato correcto de fondo.

En segundo lugar, para el módulo de selección de semilla, se exporta la idea de la Reconstrucción Multicamino [16], desarrollando un esquema Multisemilla que obliga al algoritmo a reconstruir el fondo varias veces y desde distintas zonas o superbloques de la escena, para en última instancia obtener un mapa de semillas final compuesto por las semillas iniciales y los bloques comunes a todas las reconstrucciones. Esta técnica, tal y como se demuestra en los experimentos, permite mejorar el trabajo previo realizado, obteniendo una estimación parcial del fondo inicial con más información correcta.

Por último, para la Reconstrucción Multicamino se hace una variación con respecto a [16] a la hora de buscar la continuidad espacial, proponiendo una distancia combinada de la distancia euclídea a nivel de píxel en las fronteras entre bloques adyacentes y la distancia entre las matrices de covarianza que describen el interior de dos bloques adyacentes. En la evaluación de los resul-

tados, se observa como la fase de reconstrucción espacial comienza la reconstrucción partiendo del mapa de semillas, con un porcentaje de error bajo y después de esta fase el porcentaje de error se incrementa. Cabe destacar que la fase de reconstrucción multicamino opera también en la fase de selección de semillas, ya que el mapa de semillas final se obtiene de hacer tantas reconstrucciones previas como superbloques haya. No obstante, si se consiguiese una pequeña mejora en ese apartado se mejoraría sustancialmente la eficiencia del algoritmo, no habiendo conseguido los resultados esperados en esta etapa.

## 5.2 Trabajo futuro

La tarea de Inicialización de fondo de escena para secuencias de vídeo continúa siendo un reto para futuras líneas de investigación.

En lo que respecta al PFC, se pueden incluir mejoras en la parte de *Clustering online*, por ejemplo, fijando un umbral hallado empíricamente siendo muy restrictivo o incluso umbralizando a través del entrenamiento de una SVM pero siendo mucho más selectivo con las muestras que se van a etiquetar considerando la tarea en la que se está trabajando (evitar casos de cambios de iluminación, sombras y camuflajes). Por otro lado, se puede seguir investigando otras medidas para el cálculo de la distancia entre bloques diferentes al ratio RGB empleado.

En cuanto al aspecto de la continuidad espacial, también hay gran margen de mejora. Una forma sería proponer un tamaño de bloque dinámico en función del tamaño de cada *frame* para seguir guardando la proporción del bloque con respecto a la imagen. Se puede considerar la opción de utilizar bloques solapados, y de esta forma tener más información de los bloques con los bloques del vecindario. Y por último probar con diferentes medidas de distancia combinada manteniendo el concepto de distancia de los bordes y distancia intra-bloque.





# Bibliography

- [1] L. Maddalena and A. Petrosino. Towards benchmarking scene background initialization. In *New Trends in Image Analysis and Processing (ICIAP Workshops)*, volume 9281 of *LNCS*, pages 469–476. 2015. [24](#), [34](#)
- [2] T. Bouwmans. Traditional and recent approaches in background modeling for foreground detection: An overview. *Computer Science Review*, 11-12:31–66, 2014. [1](#), [5](#), [6](#)
- [3] L. Maddalena and A. Petrosino. Background model initialization for static cameras. chapter 3, pages 1–16. Chapman and Hall/CRC 2014, 2014. [5](#), [6](#), [34](#)
- [4] Andrews Sobral and Antoine Vacavant. A comprehensive review of background subtraction algorithms evaluated with synthetic and real videos. *Computer Vision and Image Understanding*, 122:4–21, 2014. [5](#)
- [5] D. Park and H. Byun. A unified approach to background adaptation and initialization in public scenes. *Pattern Recognition*, 46(7):1985–1997, 2013. [5](#), [6](#), [8](#)
- [6] M. Paul. Efficient video coding using optimal compression plane and background modelling. *IET Image Processing*, 6(9):1311–1318, 2012. [5](#)
- [7] X. Chen, Y. Shen, and Y.H. Yang. Background estimation using graph cuts and inpainting. In *Proceedings of Graphics Interface (GI)*, pages 97–103, 2010. [5](#), [7](#)
- [8] Y. Nakashima, N. Babaguchi, and F. Jianping. Automatic generation of privacy-protected videos using background estimation. In *Proceedings of IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6, 2011. [5](#)
- [9] M. Granados, H.P. Seidel, and H.P.A. Lensch. Background estimation from non-time sequence images. In *Proceedings of Graphics Interface (GI)*, pages 33–40, 2008. [5](#)
- [10] C. Chia-Chih and J.K. Aggarwal. An adaptive background model initialization algorithm with objects moving at different depths. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pages 2664–2667, 2008. [6](#), [7](#), [12](#)
- [11] H. Hsiao and J. Leou. Background initialization and foreground segmentation for bootstrapping video sequences. *EURASIP Journal on Image and Video Processing*, 12:1–19, 2013. [6](#), [7](#)
- [12] H. Wang and D. Suter. A novel robust statistical method for background initialization and visual surveillance. In *Proceedings of Asian Conference on Computer Vision (ACCV)*, volume 3851 of *Lecture Notes in Computer Science*, pages 328–337. 2006. [6](#), [7](#), [34](#)

- [13] V. Reddy, C. Sanderson, and B.C. Lovell. A low-complexity algorithm for static background estimation from cluttered image sequences in surveillance contexts. *EURASIP Journal on Image and Video Processing*, pages 1–14, 2011. 6, 7, 8, 26, 34
- [14] X. Xun and T. S. Huang. A loopy belief propagation approach for robust background estimation. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–7, 2008. 6, 7
- [15] A. Colombari and A. Fusiello. Patch-based background initialization in heavily cluttered video. *IEEE Transactions on Image Processing*, 19(4):926–933, 2010. 6, 7, 8, 23
- [16] Diego Ortego. Rejection based multipath reconstruction for background estimation in video sequences with stationary objects. *Computer vision and image understanding*, 147:22–37, 2016. 6, 7, 8, 9, 14, 15, 16, 18, 23, 30, 34, 35
- [17] L. Maddalena and A. Petrosino. The SOBS algorithm: What are the limits? In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 21–26, 2012. 6, 34
- [18] P.L. St-Charles, G.A. Bilodeau, and R. Bergevin. Subsense: A universal change detection method with local adaptive sensitivity. *IEEE Transactions on Image Processing*, 24(1):359–373, 2015. 6, 34
- [19] C. Zezhi and T. Ellis. A self-adaptive gaussian mixture model. *Computer Vision and Image Understanding*, 122:35–46, 2014. 6
- [20] T. Bouwmans and E. H. Zahzah. Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance. *Computer Vision and Image Understanding*, 122:22–34, 2014. 6
- [21] Oliver N.M., Rosario B., and Pentland A.P. A bayesian computer vision system for modeling human interactions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8):831–843, 2000. 7
- [22] Z. Hu, G. Ye, G. Jia, X. Chen, Q. Hu, K. Jiang, Y. Wang, L. Qing, Y. Tian, X. Wu, and W. Gao. Pku@ trecvid2009: Single-actor and pair-activity event detection in surveillance video. In *Proceedings of TRECVID Workshop*, 2009. 7
- [23] Y. Tian, Y. Wang, Z. Hu, and T. Huang. Selective eigenbackground for background modeling and subtraction in crowded scenes. *IEEE Transactions on Circuits and Systems for Video Technology*, 23(11):1849–1864, 2013. 7
- [24] Howlung Eng, K.-A. Toh, A.H. Kam, J. Wang, and Wei-Yun Yau. An automatic drowning detection surveillance system for challenging outdoor pool environments. In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 532–539, 2003. 7
- [25] L. Maddalena and A. Petrosino. The 3dSOBS+ algorithm for moving object detection. *Computer Vision and Image Understanding*, 122:65–73, 2014. 7, 34
- [26] Gutchess D., Trajkovics M., Cohen-Solal E., Lyons D., and Jain A K. A background model initialization algorithm for video surveillance. In *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 733–740, 2001. 7

- [27] Horng-Horng Lin, Tyng-Luh Liu, and Jen-Hui Chuang. Learning a scene background model via classification. *IEEE Transactions on Signal Processing*, 57(5):1641–1654, 2009. 7
- [28] C. Guo, S. Gao, and D. Zhang. Belief propagation algorithm for background estimation based on local maximum weight matching. In *Proceedings of International Congress on Image and Signal Processing (CISP)*, pages 82–85, 2012. 7
- [29] D. Baltieri, R. Vezzani, and R. Cucchiara. Fast background initialization with recursive hadamard transform. In *Proceedings of IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pages 165–171, 2010. 7, 8
- [30] Mohcene Benalia and Samy Ait-Aoudia. An improved basic sequential clustering algorithm for background construction and motion detection. In *Image Analysis and Recognition*, volume 7324 of *Lecture Notes in Computer Science*, pages 216–223. 2012. 7
- [31] V. Reddy, C. Sanderson, and B.C. Lovell. An efficient and robust sequential algorithm for background estimation in video surveillance. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pages 1109–1112, 2009. 7
- [32] Simultaneous motion detection and background reconstruction with a conditional mixed-state markov random field. *International Journal of Computer Vision*, 94(3):295–316, 2011. 8
- [33] Yuru Wang, Longkui Jiang, Qiaoyuan Liu, and Minghao Yin. Optimal appearance model for visual tracking. *PLoS ONE*, 11(1):1–15, 2016. 8
- [34] Arnold W. M. Smeulders, Dung M. Chu, Rita Cucchiara, Simone Calderara, Afshin Dehghan, and Mubarak Shah. Visual tracking: An experimental survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(1446-1468), 2014. 8
- [35] Çigdem Eroglu Erdem, A. Murat Tekalp, and Bülent Sankur. Performance measures for video object segmentation and tracking. *IEEE Transactions on Image Processing*, 13(7), 2004. 8, 26
- [36] Alexandre Alahi, Pierre Vandergheynst, Michel Bierlaire, and Murat Kunt. Cascade of descriptors to detect and track objects across any network of cameras. *Computer Vision and Image Understanding*, pages 624–640, 2010. 8, 26
- [37] Julien Pilet, Christoph Strecha, and Pascal Fua. Making background subtraction robust to sudden illumination changes. *European conference on Computer Vision*, 5305:555–563, 2008. 9
- [38] J. Kapur, P. Sahoo, and A. Wong. A new method for graylevel picture thresholding using the entropy of the histogram. *Computer Graph and Image Process*, 29(3):273–285, 1985. 12, 15
- [39] Diego Ortego, Juan C. San Miguel, and José María Martínez. Long-term stationary object detection based on spatio-temporal change detection. *IEEE Signal Processing Letters*, 22(12), 2015. 12, 26, 30
- [40] C. Wolf, J. Lombardi, E. Mille, O. Celiktutan, M. Jiu, E. Dogan, G. Eren, M. Bacouche, E. Delandrea, C. Bichot, C. García, and B. Sankur. Evaluation of video activity localizations integrating quality and quantity measurements. *CVIU*, 127:14–30, 2014. 23

- [41] Y. Wang, P.M. Jodoin, F. Porikli, J. Konrad, Y. Benezeth, and P. Ishwar. Cdnet 2014: An expanded change detection benchmark dataset. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, pages 393–400, 2014. 23
- [42] K. Toyama, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: principles and practice of background maintenance. In *In proceedings of IEEE International Conference on Computer Vision (ICCV)*, volume 1, pages 255–261, 1999. 23
- [43] M. Wang, W. Li, and X. Wang. Transferring a generic pedestrian detector towards specific scenes. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3274–3281, 2012. 23
- [44] J. Varadarajan and J. Odobez. Topic models for scene analysis and abnormality detection. In *Proceedings of IEEE International Conference on Computer Vision Workshops (ICCVW)*, pages 1338–1345, 2009. 23
- [45] A. Ellis, A. Shahrokni, and J.M. Ferryman. PETS2009 and winter-PETS 2009 results: A combined evaluation. In *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS)*, pages 1–8, 2009. 23
- [46] J. Xu, S. Denman, S. Sridharan, and C. Fookes. Activity analysis in complicated scenes using dft coefficients of particle trajectories. In *IEEE International Conference on Advanced Video and Signal-Based Surveillance (AVSS)*, pages 82–87, 2012. 23
- [47] Z. Wang, E. P. Simoncelli, and A. C. Bovik. Multiscale structural similarity for image quality assessment. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Seventh Asilomar Conference on*, volume 2, pages 1398–1402. IEEE, 2003. 26
- [48] Y. Yalman and I. Erturk. A new color image quality measure based on yuv transformation and psnr for a new color image quality measure based on yuv transformation and psnr for human vision system. *Turkish Journal of Electrical Engineering and Computer Sciences*, 2013. 26
- [49] R.H. Evangelio, M. Patzold, I Keller, and T. Sikora. Adaptively splitted gmm with feedback improvement for the task of background subtraction. *IEEE Transactions on Information Forensics and Security*, 9(5):863–874, 2014. 34
- [50] D.D. Bloisi, A. Pennisi, and L. Iocchi. Background modeling in the maritime domain. *Machine Vision and Applications*, 25(5):1257–1269, 2014. 34
- [51] D.D. Bloisi, A. Grillo, A. Pennisi, L. Iocchi, and C. Passaretti. Multi-modal background model initialization. In *New Trends in Image Analysis and Processing – ICIAP 2015 Workshops*, volume 9281 of *LNCS*, pages 485–492. 2015. 34
- [52] P.L. St-Charles and G.A. Bilodeau. Improving background subtraction using local binary similarity patterns. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 509–515, 2014. 34
- [53] F. El Baf, T. Bouwmans, and B. Vachon. Fuzzy integral for moving object detection. In *IEEE International Conference on Fuzzy Systems (Fuzz-IEEE)*, pages 1729–1736, 2008. 34

- [54] P. Rodriguez and B. Wohlberg. Fast principal component pursuit via alternating minimization. In *Proceedings of IEEE International Conference on Image Processing (ICIP)*, pages 69–73, 2013. 34
- [55] B. Vandereycken. Low-rank matrix completion by Riemannian optimization. *SIAM Journal on Optimization*, 23(2):1214–1236, 2013. 34



## Capítulo 6

# Presupuesto

### 1) Ejecucion Material

- Compra de ordenador personal (Software incluido) ..... 2.000 €
- Alquiler de impresora láser durante 6 meses ..... 260 €
- Material de oficina ..... 150 €
- Total de ejecución material ..... 2.400 €

### 2) Gastos generales

- 16 % sobre Ejecucion Material ..... 352 €

### 3) Beneficio Industrial

- 6 % sobre Ejecucion Material ..... 132 €

### 4) Honorarios Proyecto

- 1800 horas a 15 €/ hora ..... 27.000 €

### 5) Material fungible

- Gastos de impresión ..... 280 €
- Encuadernación ..... 200 €

### 6) Subtotal del presupuesto

- Subtotal Presupuesto ..... 32.774 €

### 7) I.V.A. aplicable

- 21 % Subtotal Presupuesto.....6.882,2 €

**8) Total presupuesto**

---

- Total Presupuesto.....39.656,5 €

Madrid, 04/07/2016

El Ingeniero Jefe de Proyecto

Fdo.: Jaime Gallo de Cal

Ingeniero de Telecomunicación

## Capítulo 7

# Pliego de condiciones

Este documento contiene las condiciones legales que guiarán la realización, en este proyecto, de un “**Sistema de estimación de fondo de escena para secuencias de vídeo**” para ser visto en pantallas de baja resolución. En lo que sigue, se supondrá que el proyecto ha sido encargado por una empresa cliente a una empresa consultora con la finalidad de realizar dicho sistema.

Dicha empresa ha debido desarrollar una línea de investigación con objeto de elaborar el proyecto. Esta línea de investigación, junto con el posterior desarrollo de los programas está amparada por las condiciones particulares del siguiente pliego.

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

### Condiciones generales

- 1) La modalidad de contratación será el concurso. La adjudicación se hará, por tanto, a la proposición más favorable sin atender exclusivamente al valor económico, dependiendo de las mayores garantías ofrecidas. La empresa que somete el proyecto a concurso se reserva el derecho a declararlo desierto.
- 2) El montaje y mecanización completa de los equipos que intervengan será realizado totalmente por la empresa licitadora.
- 3) En la oferta, se hará constar el precio total por el que se compromete a realizar la obra y el tanto por ciento de baja que supone este precio en relación con un importe límite si este se hubiera fijado.
- 4) La obra se realizará bajo la dirección técnica de un Ingeniero Superior de Telecomunicación, auxiliado por el número de Ingenieros Técnicos y Programadores que se estime preciso para el desarrollo de la misma.

- 5) Aparte del Ingeniero Director, el contratista tendrá derecho a contratar al resto del personal, pudiendo ceder esta prerrogativa a favor del Ingeniero Director, quien no estará obligado a aceptarla.
- 6) El contratista tiene derecho a sacar copias a su costa de los planos, pliego de condiciones y presupuestos. El Ingeniero autor del proyecto autorizará con su firma las copias solicitadas por el contratista después de confrontarlas.
- 7) Se abonará al contratista la obra que realmente ejecute con sujeción al proyecto que sirvió de base para la contratación, a las modificaciones autorizadas por la superioridad o a las órdenes que con arreglo a sus facultades le hayan comunicado por escrito al Ingeniero Director de obras siempre que dicha obra se haya ajustado a los preceptos de los pliegos de condiciones, con arreglo a los cuales, se harán las modificaciones y la valoración de las diversas unidades sin que el importe total pueda exceder de los presupuestos aprobados. Por consiguiente, el número de unidades que se consignan en el proyecto o en el presupuesto, no podrá servirle de fundamento para entablar reclamaciones de ninguna clase, salvo en los casos de rescisión.
- 8) Tanto en las certificaciones de obras como en la liquidación final, se abonarán los trabajos realizados por el contratista a los precios de ejecución material que figuran en el presupuesto para cada unidad de la obra.
- 9) Si excepcionalmente se hubiera ejecutado algún trabajo que no se ajustase a las condiciones de la contrata pero que sin embargo es admisible a juicio del Ingeniero Director de obras, se dará conocimiento a la Dirección, proponiendo a la vez la rebaja de precios que el Ingeniero estime justa y si la Dirección resolviera aceptar la obra, quedará el contratista obligado a conformarse con la rebaja acordada.
- 10) Cuando se juzgue necesario emplear materiales o ejecutar obras que no figuren en el presupuesto de la contrata, se evaluará su importe a los precios asignados a otras obras o materiales análogos si los hubiere y cuando no, se discutirán entre el Ingeniero Director y el contratista, sometiéndolos a la aprobación de la Dirección. Los nuevos precios convenidos por uno u otro procedimiento, se sujetarán siempre al establecido en el punto anterior.
- 11) Cuando el contratista, con autorización del Ingeniero Director de obras, emplee materiales de calidad más elevada o de mayores dimensiones de lo estipulado en el proyecto, o sustituya una clase de fabricación por otra que tenga asignado mayor precio o ejecute con mayores dimensiones cualquier otra parte de las obras, o en general, introduzca en ellas cualquier modificación que sea beneficiosa a juicio del Ingeniero Director de obras, no tendrá derecho sin embargo, sino a lo que le correspondería si hubiera realizado la obra con estricta sujeción a lo proyectado y contratado.

- 12) Las cantidades calculadas para obras accesorias, aunque figuren por partidaalzada en el presupuesto final (general), no serán abonadas sino a los precios de la contrata, según las condiciones de la misma y los proyectos particulares que para ellas se formen, o en su defecto, por lo que resulte de su medición final.
- 13) El contratista queda obligado a abonar al Ingeniero autor del proyecto y director de obras así como a los Ingenieros Técnicos, el importe de sus respectivos honorarios facultativos por formación del proyecto, dirección técnica y administración en su caso, con arreglo a las tarifas y honorarios vigentes.
- 14) Concluida la ejecución de la obra, será reconocida por el Ingeniero Director que a tal efecto designe la empresa.
- 15) La garantía definitiva será del 4% del presupuesto y la provisional del 2%.
- 16) La forma de pago será por certificaciones mensuales de la obra ejecutada, de acuerdo con los precios del presupuesto, deducida la baja si la hubiera.
- 17) La fecha de comienzo de las obras será a partir de los 15 días naturales del replanteo oficial de las mismas y la definitiva, al año de haber ejecutado la provisional, procediéndose si no existe reclamación alguna, a la reclamación de la fianza.
- 18) Si el contratista al efectuar el replanteo, observase algún error en el proyecto, deberá comunicarlo en el plazo de quince días al Ingeniero Director de obras, pues transcurrido ese plazo será responsable de la exactitud del proyecto.
- 19) El contratista está obligado a designar una persona responsable que se entenderá con el Ingeniero Director de obras, o con el delegado que éste designe, para todo relacionado con ella. Al ser el Ingeniero Director de obras el que interpreta el proyecto, el contratista deberá consultarle cualquier duda que surja en su realización.
- 20) Durante la realización de la obra, se girarán visitas de inspección por personal facultativo de la empresa cliente, para hacer las comprobaciones que se crean oportunas. Es obligación del contratista, la conservación de la obra ya ejecutada hasta la recepción de la misma, por lo que el deterioro parcial o total de ella, aunque sea por agentes atmosféricos u otras causas, deberá ser reparado o reconstruido por su cuenta.
- 21) El contratista, deberá realizar la obra en el plazo mencionado a partir de la fecha del contrato, incurriendo en multa, por retraso de la ejecución siempre que éste no sea debido a causas de fuerza mayor. A la terminación de la obra, se hará una recepción provisional previo reconocimiento y examen por la dirección técnica, el depositario de efectos, el interventor y el jefe de servicio o un representante, estampando su conformidad el contratista.

- 22) Hecha la recepción provisional, se certificará al contratista el resto de la obra, reservándose la administración el importe de los gastos de conservación de la misma hasta su recepción definitiva y la fianza durante el tiempo señalado como plazo de garantía. La recepción definitiva se hará en las mismas condiciones que la provisional, extendiéndose el acta correspondiente. El Director Técnico propondrá a la Junta Económica la devolución de la fianza al contratista de acuerdo con las condiciones económicas legales establecidas.
- 23) Las tarifas para la determinación de honorarios, reguladas por orden de la Presidencia del Gobierno el 19 de Octubre de 1961, se aplicarán sobre el denominado en la actualidad "Presupuesto de Ejecución de Contrata" y anteriormente llamado "Presupuesto de Ejecución Material" que hoy designa otro concepto.

### **Condiciones particulares**

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

- 1) La propiedad intelectual de los procesos descritos y analizados en el presente trabajo, pertenece por entero a la empresa consultora representada por el Ingeniero Director del Proyecto.
- 2) La empresa consultora se reserva el derecho a la utilización total o parcial de los resultados de la investigación realizada para desarrollar el siguiente proyecto, bien para su publicación o bien para su uso en trabajos o proyectos posteriores, para la misma empresa cliente o para otra.
- 3) Cualquier tipo de reproducción aparte de las reseñadas en las condiciones generales, bien sea para uso particular de la empresa cliente, o para cualquier otra aplicación, contará con autorización expresa y por escrito del Ingeniero Director del Proyecto, que actuará en representación de la empresa consultora.
- 4) En la autorización se ha de hacer constar la aplicación a que se destinan sus reproducciones así como su cantidad.
- 5) En todas las reproducciones se indicará su procedencia, explicitando el nombre del proyecto, nombre del Ingeniero Director y de la empresa consultora.
- 6) Si el proyecto pasa la etapa de desarrollo, cualquier modificación que se realice sobre él, deberá ser notificada al Ingeniero Director del Proyecto y a criterio de éste, la empresa consultora decidirá aceptar o no la modificación propuesta.

- 7) Si la modificación se acepta, la empresa consultora se hará responsable al mismo nivel que el proyecto inicial del que resulta el añadirla.
- 8) Si la modificación no es aceptada, por el contrario, la empresa consultora declinará toda responsabilidad que se derive de la aplicación o influencia de la misma.
- 9) Si la empresa cliente decide desarrollar industrialmente uno o varios productos en los que resulte parcial o totalmente aplicable el estudio de este proyecto, deberá comunicarlo a la empresa consultora.
- 10) La empresa consultora no se responsabiliza de los efectos laterales que se puedan producir en el momento en que se utilice la herramienta objeto del presente proyecto para la realización de otras aplicaciones.
- 11) La empresa consultora tendrá prioridad respecto a otras en la elaboración de los proyectos auxiliares que fuese necesario desarrollar para dicha aplicación industrial, siempre que no haga explícita renuncia a este hecho. En este caso, deberá autorizar expresamente los proyectos presentados por otros.
- 12) El Ingeniero Director del presente proyecto, será el responsable de la dirección de la aplicación industrial siempre que la empresa consultora lo estime oportuno. En caso contrario, la persona designada deberá contar con la autorización del mismo, quien delegará en él las responsabilidades que ostente.