

# Posprocesamiento morfológico adaptativo basado en algoritmos genéticos y orientado a la detección robusta de humanos

Enrique Carmona, Javier Martínez-Cantos y José Mira

Dpto. de Inteligencia Artificial, ETSI Informática, UNED,  
Juan del Rosal 16, 28040, Madrid, Spain.  
{ecarmona, jmira}@dia.uned.es  
javiermc@info-ab.uclm.es

**Resumen.** Existen distintas aproximaciones para la detección de objetos móviles basadas en el denominado método de background. En secuencias reales, el gran inconveniente de este método, que comparte con otros métodos de segmentación, es la forma de eliminar el ruido inherente tanto al foreground como al background. Una aproximación muy utilizada para resolver este problema es la aplicación de una secuencia fija de operadores morfológicos pero que, al tenerse que decidirse a priori, no siempre está garantizado el éxito de la restauración del foreground. En este trabajo, se propone un método de posprocesamiento que, para cada frame, ofrece automáticamente una secuencia de operadores morfológicos obtenida a partir de la salida de un algoritmo genético cuya búsqueda está fuertemente sesgada hacia la restauración de siluetas humanas. Finalmente, se propone el uso de este método dentro de un sistema de detección robusta de humanos.

*Palabras clave:* Segmentación, Morfología, Algoritmos genéticos, Detección de Humanos

## 1 Introducción

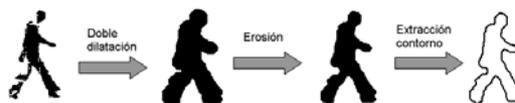
La detección de objetos móviles en secuencias de video es el primer paso relevante en la extracción de información en muchas aplicaciones de visión por computador, incluyendo, por ejemplo, la video-vigilancia, el control de tráfico y el seguimiento de personas. En esta primera etapa de detección, cuanto más fiable sea la forma y posición del objeto que se mueve, más fiable será la identificación del mismo.

En la literatura existen diferentes métodos para la detección de objetos móviles basados, por ejemplo, en métodos estadísticos [1], [2], en la substracción de frames consecutivos [3], en el flujo óptico [4] o en aproximaciones híbridas [5] que combinan algunas de estas técnicas mencionadas. No obstante, por su rapidez y sencillez de implementación, una de las aproximaciones más utilizadas, utilizando cámara fija, se basa en el denominado método de substracción del background y sus múltiples variantes [6], [7], [8], [9], [10], [11], [12]. Básicamente, este método permite detectar regiones en movimiento restando pixel a pixel la imagen actual de una imagen de background tomada como referencia (modelo de background) y creada mediante el

promediado de imágenes a lo largo del tiempo durante un periodo de inicialización

Las salidas producidas por los algoritmos de detección mencionados anteriormente, sobre todo si se trabaja con escenas reales, contienen ruido generalmente. Las causas de ruido son debidas principalmente al ruido intrínseco de la propia cámara, a reflejos indeseados, a objetos cuyo color total o parcial coincide con el background y a la existencia de sombras y de cambios (artificiales o naturales) repentinos en la iluminación. El efecto total de estos factores puede ser doble: por un lado, pueden provocar que zonas no pertenecientes a objetos en movimiento se incorporen al foreground (ruido de foreground) y, de otro lado, que determinadas zonas pertenecientes a objetos en movimiento dejen de aparecer en el foreground (ruido de background). En ambos casos, se produce un deterioro en la segmentación de los objetos que es necesario reparar o al menos mejorar si se quiere tener unas garantías mínimas de éxito en las etapas siguientes relacionadas con el *tracking* o con el reconocimiento de objetos.

El recurso empleado en muchos trabajos que se enfrentan con el problema del ruido consiste en la aplicación de una etapa de posprocesamiento basada en la utilización de operadores morfológicos [1], [5], [13], [14], [15]. Por ejemplo, tal y como se indica en la Figura 1, es muy típica la secuencia formada por una fase de dilatación seguida de otra fase de erosión. La primera va encaminada a rellenar todos los huecos que pudieran existir en la silueta de los objetos y/o a conectar distintos fragmentos de un mismo objeto. La segunda es utilizada para invertir la expansión de los límites del objeto producida por la dilatación. Obsérvese que esta decisión tomada anticipadamente puede, en algunos casos, ser crítica. Efectivamente, el gran inconveniente de un esquema fijo de posprocesamiento morfológico radica en su no adaptatividad puesto que, independientemente del mapa de foreground obtenido en un determinado instante del proceso, es necesario haber decidido previamente, en tiempo de compilación, el número de operadores, el orden de ejecución de la secuencia de operadores y el tipo y tamaño del elemento estructurante que utilizará cada uno de ellos.



**Figura 1.** Ejemplo de posprocesamiento morfológico de un objeto de foreground para mejorar su segmentación.

En este trabajo, se propone una nueva aproximación basada en el método de sustracción de background combinándolo con técnicas de aprendizaje evolutivas. Básicamente, la idea consiste en hacer evolucionar una población de secuencias de operadores morfológicos mediante un algoritmo genético con el objeto de encontrar la mejor secuencia que permita eliminar el ruido para restaurar el foreground. Para orientar y facilitar la búsqueda, se utiliza una métrica que tiene la propiedad de minimizar su valor cada vez que la secuencia morfológica se aplica a todas aquellas zonas del foreground con siluetas humanas. El resultado final es la construcción de un sistema que abarca simultáneamente la tarea de detección, la de posprocesamiento y la de reconocimiento. Eso sí, las dos últimas, siempre orientadas a la búsqueda de actividad humana.

El resto de este trabajo se organiza de la siguiente manera. En la Sección 2 se aborda la descripción del método de posprocesamiento morfológico evolutivo propuesto, haciendo hincapié en la codificación cromosómica utilizada y en la descripción de la función de adaptación construida. En base al proceso de minimizar esta función, se seleccionará finalmente la secuencia morfológica más apta para restaurar el foreground. La sección 3 describe, mediante un diagrama de bloques la interacción entre la etapa de aprendizaje y el resto de las distintas etapas del sistema propuesto para la detección de humanos. Los resultados experimentales se recogen en la Sección 4 y, finalmente, las conclusiones en la Sección 5.

## **2 Posprocesamiento morfológico adaptativo basado en algoritmos genéticos**

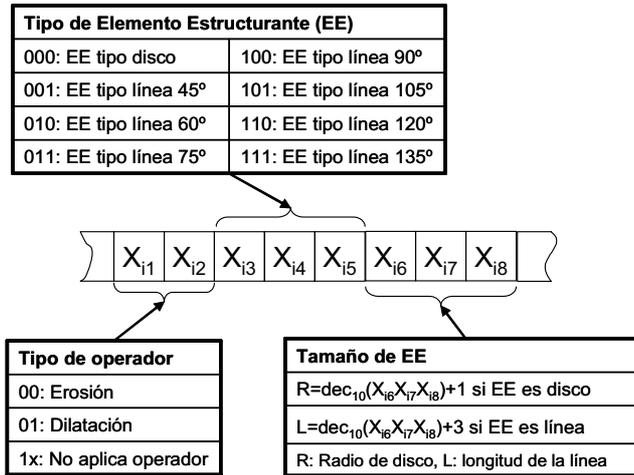
El método de posprocesamiento adaptativo basado en algoritmos genéticos propuesto en este trabajo consiste, básicamente, en buscar en un espacio de secuencias de operadores morfológicos, aquella secuencia óptima que permita mejorar el mapa de foreground en cada instante. Esta búsqueda está fuertemente sesgada (*bias*) debido a que siempre está orientada a la detección de objetos móviles que cumplen unas determinadas características. Concretamente, en el trabajo que aquí nos ocupa, características humanas.

Hay que tener en cuenta que, en general, el modelado de un problema utilizando AG's requiere, en primer lugar, definir la forma de representación de cada individuo (cromosoma), en segundo lugar, construir la función de adaptación necesaria para evaluar la bondad de cada individuo respecto de la solución del problema y, finalmente, elegir los distintos operadores, parámetros y criterios que definen la estructura del AG. Aunque estas tres etapas están fuertemente interrelacionadas, las dos primeras constituyen las etapas más críticas en el proceso de encajar un problema de optimización en el marco de un AG. Así, para la codificación de los individuos, se debe utilizar una representación mínima que sea lo más expresiva posible, de forma tal que sea capaz de representar cualquier solución al problema. En cuanto a la función de adaptación elegida, ésta debe modelar de forma adecuada el problema de optimización que se quiere resolver. El éxito de la solución del problema dependerá en gran medida de ello.

### **2.1 Codificación cromosómica de la secuencia de operadores**

Para cada frame de entrada, tras la fase de detección del foreground, el método aquí propuesto decidirá, en tiempo de ejecución, la secuencia de operadores morfológicos más adecuada a aplicar y orientada a la restauración y detección de todos aquellos objetos que pudieran estar asociados a un humano. En cada instante, esta decisión corresponde a la decodificación del mejor cromosoma perteneciente a la población resultante de la ejecución de un algoritmo genético modelado y parametrizado para tal efecto. Puesto que para definir una secuencia de operadores morfológicos es necesario declarar el tipo de cada operador, su número, el tipo de elemento estructurante

asociado a cada uno de ellos y, finalmente, el orden de aplicación, en esta sección se describe la codificación realizada para definir la estructura cromosómica de cada uno de los individuos que conformarán la población a evolucionar en el AG.



**Figura 2.** Detalle de la codificación de una secuencia de operadores morfológicos mediante un cromosoma de 64 bits. De izquierda a derecha, cada subcadena de 8 bits del cromosoma contiene la información necesaria para codificar un operador morfológico.

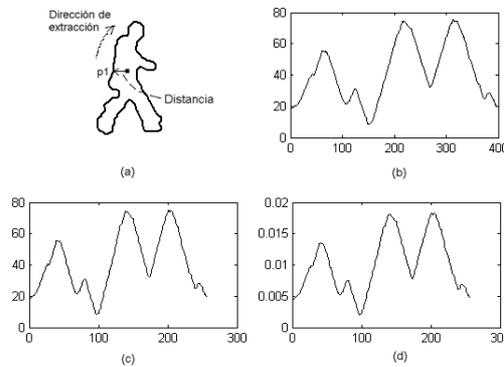
Así, la codificación elegida para cada individuo se hace de acuerdo a la agrupación de 8 subcadenas consecutivas de 8 bits cada una, para formar finalmente una cadena binaria de 64 bits. Por tanto, cada cromosoma codifica una secuencia como máximo de 8 operadores morfológicos. Los dos primeros bits de cada subcadena codifican el tipo de operador morfológico: erosión, dilatación o no aplica operador<sup>1</sup>. De esta forma se consigue que la secuencia de operadores pueda tener un número mínimo de cero operadores y un número máximo de ocho. Los tres siguientes bits codifican el tipo de elemento estructurante: en forma de disco o en forma de línea con diferentes ángulos de inclinación (medidos respecto a la horizontal). La elección del rango de variación de estos ángulos, 45° a 135°, no es arbitraria ya que trata de tener en cuenta el ángulo que puede llegar a formar los brazos o las piernas de un individuo al caminar respecto a su vertical. Finalmente, los tres bits restantes codifican el tamaño del elemento estructurante. En el caso de un elemento estructurante en forma de disco, el radio del disco viene determinado por la decodificación en base decimal de los tres dígitos binarios, sumada en una unidad (no tiene sentido un elemento estructurante de radio 0). Si el elemento estructurante es de tipo línea, la decodificación en base decimal de la terna de bits, aumentada en 3, da el valor de su longitud (no tienen sentido líneas de longitud menor de tres). La Figura 2 muestra esquemáticamente la codificación de un operador morfológico en una subcadena de 8 bits.

<sup>1</sup> Si los dos primeros bits de la subcadena codifican la no aplicación de operador, entonces se ignoran los seis bits restantes de la subcadena.

### 2.2 Modelado de la función de adaptación

La función de adaptación utilizada por el algoritmo genético va encaminada a evaluar el grado de aproximación entre la silueta de un objeto de foreground restaurado mediante una secuencia de operadores morfológicos y la silueta correspondiente a un humano. Para cuantificar el grado de aproximación entre ambas siluetas, se requiere tanto una métrica adecuada como una base de datos de siluetas humanas patrón con las que realizar la comparación.

La métrica aquí utilizada está basada en la similaridad del contorno de los objetos. Aunque existen numerosos métodos en la literatura para comparar formas [16], [17], aquí nos hemos decantado por la métrica definida en base a la distancia euclídea existente entre la denominada señal de distancia [13] asociada a cada uno de los dos objetos a comparar. Esta elección obedece a la importante propiedad que posee la señal de distancia de ser invariante a las rotaciones, traslaciones y cambio de escala.



**Figura 3.** Ejemplo de contorno de un objeto (a) y su correspondiente señal de distancia original (b) escalada (c) y normalizada (d), mostrada en el eje  $y$  para cada punto del contorno (eje  $x$ ).

Sea  $S=\{p_1, p_2, \dots, p_n\}$  los puntos que determinan el perímetro de un objeto que consta de  $n$  puntos ordenados, sea  $p_1$  el punto central más a la izquierda de la silueta y el resto obtenidos en la dirección de las manecillas del reloj y sea  $c_m$  su centro de gravedad (ver Figura 3). La señal de distancia  $SD=\{d_1, d_2, \dots, d_n\}$ , ver ejemplo en Figura 3b, se genera calculando la distancia euclídea entre  $c_m$  y cada  $p_i$ , es decir,

$$d_i = Dist_{eucl}(c_m, p_i), \quad \forall i \in [1..n]. \tag{1}$$

Para definir una métrica efectiva que permita la comparación de dos formas basada en la similitud de sus señales distancia respectivas, tiene que coincidir la posición relativa, respecto del c.d.g., del primer punto a comparar, el sentido de recorrido del perímetro y el número de puntos en las dos formas a comparar. Cumpliendo estas tres condiciones se consigue que ante dos formas parecidas pero de tamaño diferente, la posición relativa de cada pareja de puntos a comparar sea aproximadamente la misma.

Por tanto, si  $N$  es el tamaño del vector señal de distancia  $SD$  y a  $C$  es una constante usada para fijar el nuevo tamaño del vector a un valor siempre fijo, la señal de distan-

cia escalada, ver ejemplo en Figura 3c, se calcula entonces sub/super muestreando la señal de distancia original  $SD$  como sigue:

$$\overline{SD}[i] = SD \left[ i * \frac{N}{C} \right], \quad \forall i \in [1..C]. \quad (2)$$

En el siguiente paso, la señal de distancia escalada se normaliza para tener integral de área unidad (ver Figura 3c) y garantizar así la invarianza al tamaño. La señal de distancia escalada y normalizada se calcula de acuerdo a la siguiente ecuación:

$$\overline{\overline{SD}}[i] = \frac{\overline{SD}[i]}{\sum_{i=1}^C \overline{SD}[i]}, \quad \forall i \in [1..C]. \quad (3)$$

La métrica de clasificación compara la similaridad entre la forma del perímetro de dos objetos,  $A$  y  $B$ , calculando la distancia entre sus correspondientes señales de distancia escaladas,  $SD_A$  y  $SD_B$ . Esta distancia se calcula en [13] como el sumatorio del valor absoluto de las diferencia de las componentes de ambas señales. Aquí se calculará como la distancia euclídea entre las dos señales, es decir:

$$Sim_{AB} = \sqrt{\sum_{i=1}^{i=C} \left( \overline{\overline{SD}}_A[i] - \overline{\overline{SD}}_B[i] \right)^2}. \quad (4)$$

Para completar la definición de la función de adaptación, hay que tener en cuenta que la comparación de la señal de distancia del objeto restaurado por la secuencia de operadores morfológicos que se hace evolucionar, se realiza sobre todas las  $M$  señales de distancia almacenadas previamente en una base de datos de siluetas humanas. Finalmente, la función de adaptación,  $f_{adp}$ , se define como:

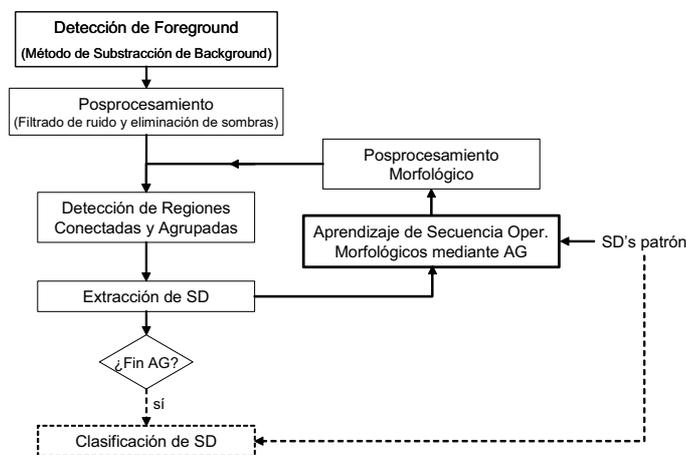
$$f_{adp} = \min_{i=1..M} (Sim_{AB[i]}). \quad (5)$$

Es decir, el AG evoluciona para minimizar la distancia euclídea existente entre la señal de distancia del objeto analizado y la señal de distancia más similar a esta última, existente en la base de datos.

### 3 Sistema para el procesamiento morfológico adaptativo

El diagrama de bloques de nuestra aproximación para el procesamiento morfológico adaptativo basado en algoritmos genéticos se muestra en la Figura 4. A cada frame de entrada de la secuencia de video, se aplica el método de substracción de background adaptativo para identificar todos aquellos objetos que pertenecen al foreground. La existencia de ruido se manifiesta tanto en el foreground como en el background y, por tanto, se hace necesaria la aplicación de una etapa de posprocesamiento. En nuestro caso, se dividirá en dos fases: la primera encaminada a eliminar el ruido de foreground y, la segunda, a eliminar el ruido de background.

Una de las principales causas que originan ruido de foreground son las sombras que proyectan los objetos y que tienen unos efectos altamente indeseables. Por tanto, se hace necesaria la utilización de algún método [18], [19], [20] para eliminar este tipo de ruido porque, de lo contrario, el fracaso en una posible posterior etapa de reconocimiento de objetos está casi asegurado. En nuestro sistema usamos un esquema de detección y eliminación de sombra basada en el trabajo presentado en [20] que en la práctica demuestra trabajar bien. La primera fase de posprocesado finaliza con la aplicación de un filtro para eliminar todo aquel ruido de foreground asociado a regiones de tamaño muy pequeño y que puede ser debido tanto al propio proceso de segmentación como al resultado de aplicar el filtro de eliminación de sombras.



**Figura 4.** Diagrama de bloques de un sistema de detección robusta de actividad humana en secuencias de vídeo basado en posprocesamiento morfológico adaptativo.

La segunda fase de posprocesado está relacionada con la eliminación de ruido de background expresado por la aparición de huecos en la silueta del objeto, por la disminución de tamaño y/o por la fragmentación del mismo. En este último caso, para determinar las distintas agrupaciones con las que trabajar, se implementa una etapa de detección de regiones conectadas y agrupadas.

En definitiva, la reconstrucción se consigue mediante la aplicación automática del posprocesamiento morfológico adaptativo aquí propuesto y aplicado a cada región de agrupaciones obtenidas en la etapa anterior. En la Figura 4, esto se representa mediante el lazo de realimentación. Así, la misión del módulo de aprendizaje es buscar, desde un punto de vista evolutivo, de forma automática y para cada frame, la mejor secuencia de operadores morfológicos que permita transformar cada agrupación de manchas del foreground, si ello es posible, en una silueta perteneciente a un conjunto de siluetas preestablecido. En nuestro caso, siluetas humanas asociadas a distintas posturas. Una vez que finaliza la búsqueda evolutiva de la secuencia óptima, ésta se aplica.

Es importante resaltar que el método aquí propuesto sólo hace hincapié en la fase de identificación y restauración de objetos móviles en la escena, es decir, no aborda la

tarea de clasificación. Esta es la razón por la que el último módulo de la Figura 4 aparece en línea discontinua. No obstante, con el método propuesto, dicha tarea resultaría relativamente fácil de abordar. Así, se podría medir la similaridad de la señal de distancia del objeto restaurado con las que son de interés y están almacenadas en la base de datos. Para ello, por ejemplo, se podría volver a utilizar la métrica definida en la expresión (5) o, en otro caso, el grado de correlación entre señales distancia. Finalmente, la decisión de pertenencia a la clase se realizaría mediante el uso de un umbral adecuado.

#### 4 Resultados experimentales

Aunque de partida parece más adecuado usar la señal de distancia escalada y normalizada por su invarianza al tamaño, los resultados experimentales demuestran que los mejores resultados se obtienen cuando se trabaja sólo con la señal de distancia escalada sin normalizar. Esto es debido a que se producen restauraciones indeseadas obtenidas como consecuencia de una aplicación excesiva del operador de erosión, que tiende a reducir la silueta del humano hasta una pequeña mancha regular y que, aunque nada tiene que ver con la silueta real, su señal de distancia es muy similar a la que se obtiene para una postura humana con brazos y piernas pegadas al cuerpo. En este sentido, el uso de la señal de distancia escalada sin normalizar implica que la base de datos creada ya sólo es válida para analizar escenas captadas por una cámara que se encuentre a una distancia similar a la que se utilizó para obtener dicha base de datos, es decir, debe cumplirse que los individuos de las secuencias a analizar tengan tamaños similares a los almacenados en la base de datos.

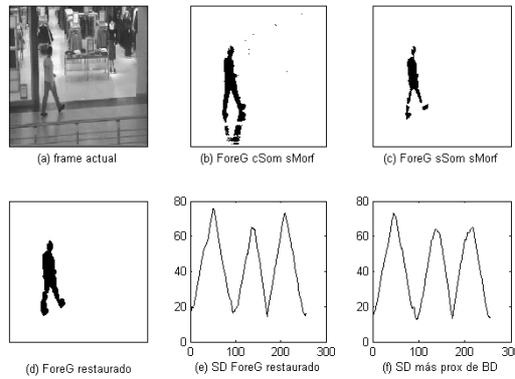
Con la restricción mencionada, para demostrar la validez del método propuesto se utiliza como entrada un conjunto de subsecuencias extraídas de la secuencia denominada *WalkByShop1front.mpg*, perteneciente al almacén<sup>2</sup> público de secuencias de vídeo relacionadas con el denominado proyecto CAVIAR<sup>3</sup>. Así, a modo de ejemplo, la Figura 5 muestra los sucesivos pasos seguidos para procesar uno de los frames perteneciente a lo que hemos denominado *secuencia A*, compuesta de 56 frames. Esta secuencia muestra cómo un hombre, que proyecta sombra, se adentra en el interior de una tienda desplazándose en dirección noreste. Es típico en esta escena el ruido debido a la similaridad de color entre ropa y background. Así, la Figura 5a muestra el ejemplo de frame de entrada seleccionado para esta secuencia y la Figura 5b el foreground resultante de aplicar el método de substracción de background a dicha entrada. Como puede comprobarse, el resultado es altamente ruidoso debido a la existencia de sombras (ruido de foreground). A continuación, la aplicación del filtro de eliminación de sombras produce el resultado mostrado en la Figura 5b. Obsérvese que se ha conseguido reducir el ruido de foreground a costa de producir ruido de background (el objeto disminuye de tamaño y queda fragmentado). Seguidamente, se procede a reconstruir el objeto mediante el método aquí propuesto y cuyo resultado es el que se muestra en la Figura 5d. Finalmente, las Figura 5e-f muestran, respectivamente, la

<sup>2</sup> <http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/>

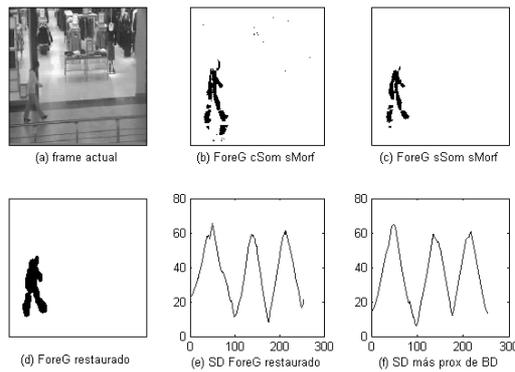
<sup>3</sup> <http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>

señal de distancia de la silueta obtenida tras la restauración y la que de forma automática se selecciona de la base de datos durante el proceso de aprendizaje como señal de distancia más similar. El parecido entre ambas resulta evidente.

Análogamente, en la Figura 6 se muestra el procesado equivalente de uno de los frames pertenecientes, esta vez, a lo que hemos denominado *secuencia B*. La descripción de las distintas subfiguras es totalmente similar a la realizada anteriormente para la Figura 5. Ahora, la *secuencia B* consta de 75 frames y conserva el mismo escenario de background que la anterior. Sin embargo, ahora es una mujer la que se desplaza en dirección este, delante del escaparate de la tienda. En este caso, las causas de ruido son también similares a las de la *secuencia A*: sombras y coincidencia de colores entre objeto y background.

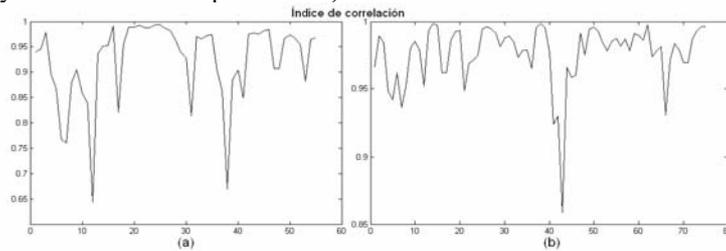


**Figura 5.** Ejemplo gráfico de la salida de las distintas etapas implicadas en el sistema de procesamiento morfológico adaptativo en la *secuencia B* de vídeo: frame de entrada (a), detección de objetos móviles (b), eliminación de sombras (c), restauración del foreground (d). También se muestra la señal de distancia del humano restaurado (e) y la del humano más similar de la base de datos (f).

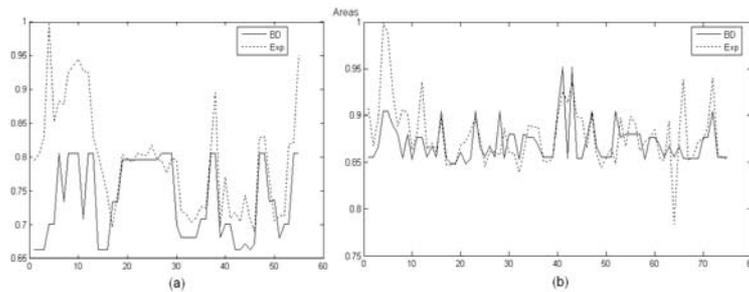


**Figura 6.** Ejemplo gráfico de la salida de las distintas etapas implicadas en el sistema de procesamiento morfológico adaptativo en la *secuencia A* de vídeo: frame de entrada (a), detección de objetos móviles (b), eliminación de sombras (c), restauración del foreground (d). También se muestra la señal de distancia del humano restaurado (e) y la del humano más similar de la base de datos (f).

Ante la imposibilidad, por motivos de espacio, de mostrar el resultado del procesamiento de todos los frames de las secuencias utilizadas, hemos recurrido a calcular el índice de correlación entre la señal de distancia de la silueta restaurada respecto de la mejor señal de distancia de la base de datos. Los resultados se muestran en la Figura 7 y en ella se puede apreciar, para ambas secuencias, los buenos resultados obtenidos y expresados por el alto grado de correlación conseguido (superior al 0.8 en la mayoría de los frames procesados).



**Figura 7.** Gráficas que muestran, para cada frame (eje x) de la secuencia, el máximo índice de correlación entre la señal de distancia de la silueta restaurada respecto de la mejor señal de distancia de la base de datos para la *secuencia A* (a) y la *secuencia B* (b).



**Figura 8.** Comparación, para cada frame (eje x) de la secuencia, del área de la silueta restaurada respecto de aquella de la base de datos para la que sus señales distancia respectivas presentan máxima correlación. En cada caso, *secuencia A* (a) y *secuencia B* (b), las áreas se han normalizado respecto al valor máximo obtenido experimentalmente.

Para añadir más información, en la Figura 8 se compara, también para cada frame de cada secuencia, el área de la silueta restaurada (trazo discontinuo) y la correspondiente a la silueta más parecida de la base de datos (trazo sólido), es decir, aquella cuya señal de distancia presenta mayor correlación. Teniendo en cuenta ambos resultados, grado de correlación y área, se observa que en los inicios de ambas secuencias existe un cierto efecto de inercia que impide obtener siluetas restauradas con un valor de área cercano al deseado pero que, sin embargo, permiten ser identificadas como siluetas humanas dado su alto índice de correlación con la base de datos. Es decir, corresponden a siluetas cuya restauración es reconocible pero en las que predomina el efecto de operadores morfológicos de dilatación frente a los de erosión (el área de las siluetas que se obtienen es mayor que el real). La explicación de este fenómeno está motivada por cómo se inicializa la población de operadores morfológicos al analizar

cada frame. En el primer frame de cada secuencia, la población de operadores se escoge de forma aleatoria pero, a partir de aquí, la población final de secuencias de operadores morfológicos obtenida en el frame  $i$ -ésimo se utilizará como población inicial para procesar el frame  $(i+1)$ -ésimo. La ventaja de esta política es doble, no sólo ofrece buenos resultados, tal y como lo evidencia el hecho de que las áreas obtenidas se acercan paulatinamente a las áreas reales sino que, además, permite acelerar la convergencia del algoritmo genético en el procesamiento de cada frame.

Finalmente, resaltar que el caso en el que coincide un valor de correlación relativamente bajo con una discrepancia grande de áreas, suele estar asociado a una mala restauración. Este hecho se puede apreciar, principalmente, en el procesado de ciertos frames de la subsecuencia B (Figura 8b).

## 5 Conclusiones y trabajo futuro

En secuencias de vídeo no captadas en condiciones ideales de laboratorio sino a partir de situaciones reales, el ruido inherente a la propia escena y al originado como resultado de aplicar etapas de procesado a la misma, puede deteriorar la salida resultante de la tarea de segmentación asociada a un sistema de visión de detección de humanos. Aquí se presenta un método novedoso de posprocesamiento morfológico adaptativo basado en algoritmos genéticos y en la idea de señal de distancia. Hay que tener en cuenta que aunque el método presentado está orientado a la restauración de siluetas humanas, el método no pierde generalidad si se quiere aplicar a otro tipo de objetos: simplemente basta utilizar una base de datos de siluetas que sea representativa de la clase de objeto a detectar. Creemos que la aplicabilidad del método radica en facilitar la implementación de una etapa de restauración posterior a cualquier etapa de segmentación y anterior a cualquier etapa de *tracking* o de clasificación. Los prometedores resultados obtenidos avalan la eficacia del método propuesto.

Por otro lado, es conocida la penalización en tiempo asociada al uso de algoritmos genéticos: debería lanzarse un proceso de búsqueda evolutivo por cada frame y por cada región de agrupación de manchas de foreground que haya en la escena. La solución pasa por paralelizar el procedimiento, favorecida además por la naturaleza paralelizable inherente a todo algoritmo genético. Trabajos en este sentido pueden consultarse en [21], [22], [23]. Esta cuestión no ha sido abordada aquí y se propone como una línea de trabajo futura.

## Agradecimientos

Esta investigación ha sido soportada por la *CICYT*, a través del proyecto *TIN2004-07661*. También queremos agradecer al *EC Funded CAVIAR project/IST 2001 37540* al hacer público su almacén de secuencias de vídeo, algunas de las cuales han sido utilizadas en este trabajo

## Referencias

1. Haritaoglu, I., Harwood, D., Davis, L. S.: "W4: A real time system for detecting and tracking people". In *Computer Vision and Pattern Recognition*, (1998), 962-967.
2. Stauffer, C., Grimson, W.: "Adaptive background mixture models for realtime tracking". In *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, (1999), 246-252.
3. Lipton, A. J., Fujiyoshi, H., Patil, R.S.: "Moving target classification and tracking from real-time video". In *Proc. of Workshop Applications of Computer Vision*, (1998), 129-136
4. Wang, L., Hu, W., Tan, T.: "Recent developments in human motion analysis". *Pattern Recognition*, 36 (3), March, (2003), 585-601
5. Collins, R. T. et al.: "A system for video surveillance and monitoring: VSAM final report". Technical report CMU-RI-TR-00-12, Robotics Institute, Carnegie Mellon University, May (2000).
6. Haritaoglu, I., Harwood, D., Davis, L.S.: "W4: Real-Time Surveillance of People and Their Activities". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, Aug. (2000), 809-830.
7. Amamoto, N., Fujii, A.: "Detecting Obstructions and Tracking Moving Objects by Image Processing Technique". *Electronics and Comm. Japan, Part 3*, vol. 82, no. 11, (1999), 28-37
8. Stauffer, C., Grimson, W.: "Learning Patterns of Activity Using Real-Time Tracking". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, Aug. (2000), 747-757.
9. McKenna, S.J., Jabri, S., Duric, Z., Rosenfeld, A., Wechsler, H.: "Tracking Groups of People". *Computer Vision and Image Understanding*, vol. 80, no. 1, Oct. (2000), 42-56.
10. Wren, C., Azarbayejani, A., Darrell, T., Pentland, A.P.: "Pfinder: Real-Time Tracking of the Human Body". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, July (1997), 780-785
11. Seki, M., Fujiwara, H., Sumi, K.: "A Robust Background Subtraction Method for Changing Background". *Proc. IEEE Workshop Applications of Computer Vision*, (2000), 207-213
12. Cucchiara, R., Piccardi, M., Prati, A.: "Detecting Moving Objects, Ghost and Shadows in Video Streams". *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 25, no. 10, Oct., (2003), 1337-1342.
13. Dedeoglu, Y.: "Moving Object Detection, Tracking and Classification for Smart Video Surveillance". Ph.D. Thesis, Department of Computer Engineering and The Institute of Engineering and Science of Bilkent University, (2004).
14. Ekinici, M., Gedikli, E.: "Silhouette Based Human Motion Detection and Analysis for Real-Time Automated Video Surveillance". *Turk J Elec Engin*, vol.13, no. 2, (2005), 199-229
15. Heijden, F.: "Image Based Measurement Systems: Object Recognition and Parameter Estimation". Wiley, January, (1996).
16. Loncaric, S.: "A survey of shape analysis techniques". *Pattern Recognition*, vol. 31, no. 8, August (1998), 983-1001.
17. Veltkamp, R.C., Hagedoorn, M.: "State-of-the-art in shape matching". *Principles of Visual Information Retrieval*, Springer, (2001), 87-119.
18. Cucchiara, R., Grana, C., Piccardi, M., Prati, A., Sirotti, S.: "Improving Shadow Suppression in Moving Object Detection with HSV Color Information". *Proc. IEEE Int'l Conf. Intelligent Transportation Systems*, Aug., (2001), 334-339.
19. Prati, A., Cucchiara, R., Mikic, I., Trivedi, M.M.: "Analysis and Detection of Shadows in Video Streams: A Comparative Evaluation". *Proc. IEEE Int'l Conf. Computer Vision and Pattern Recognition*, (2001).
20. Horprasert, T., Harwood, D., Davis, L.S.: "A statistical approach for realtime robust background subtraction and shadow detection". In *Proc. of IEEE Frame Rate Workshop*, Kerkyra, Greece, (1999), 1-19.

21. Bevilacqua, A., "Optimizing parameters of a motion detection system by means of a distributed genetic algorithm". *Image and Vision Computing*, vol. 23, (2005), 815-829.
22. Bevilacqua, A., Campanini, R., Lanconelli, N., "A Distributed Genetic Algorithm for Parameters Optimization to Detect Microcalcifications in Digital Mammograms". In *Proc. of Third European Workshop on Evolutionary Computation in Image Analysis and Signal Processing (EvoIASP)*, Como, Italy, LNCS Springer-Verlag, vol. 2037, (2001), 278-287.
23. Cantú-Paz, E., "A Survey of Parallel Genetic Algorithms". Report No. 97003, [cite-seer.ist.psu.edu/155991.html](http://cseer.ist.psu.edu/155991.html).