

Umbral adaptativo para la detección de objetos en movimiento usando visión computacional

J.F.Vazquez, C.A.Luna

Departamento de Telecomunicaciones,
Universidad de Oriente

Santiago de Cuba, Cuba

jfvglez@yahoo.es

M.Mazo, J.L.Lázaro, J.Ureña, J.J.García

Departamento de Electrónica, Universidad de
Alcalá

Alcalá de Henares, España

mazo@depeca.uah.es

Abstract—En este trabajo se presenta un método de umbralización adaptativa para la detección de objetos en movimiento en entornos exteriores, usando visión computacional. Esta alternativa trata de resolver el problema que presentan varios métodos de umbralización, cuando existe solo una clase en las imágenes capturadas. La solución propuesta utiliza el histograma de una matriz de distancias y la varianza de una imagen promedio para obtener el umbral adaptativo, esta propuesta tiene en cuenta las variaciones de las condiciones climatológicas y de iluminación, presentes en los entornos exteriores. Se utilizaron varios bancos de imágenes para comparar el método propuesto con otro descrito en este trabajo, mostrándose con los resultados alcanzados, la robustez de nuestra alternativa.

I. INTRODUCCIÓN

La umbralización es una técnica importante en la segmentación de imágenes, usada para identificar y separar objetos del fondo, en base a la distribución de los niveles de gris o la textura de los objetos en las imágenes. Muchas técnicas de umbralización se basan en la estadística del histograma de los niveles de gris o en la matriz de co-ocurrencia de una imagen [1-9]. La determinación de los umbrales puede ser mediante procedimientos paramétricos o no paramétricos [1],[4],[13]. En los procedimientos paramétricos, la distribución de los niveles de gris de un objeto permite obtener el umbral. Estudios como los descritos en [3] y en [5], son un ejemplo de este tipo de procedimiento. En los procedimientos no paramétricos, los umbrales se obtienen de una forma óptima, de acuerdo a algún criterio. Por ejemplo, el método de Otsu selecciona los umbrales óptimos maximizando la varianza entre clases mediante una búsqueda exhaustiva [2]. En los métodos descritos en [7] y [8], el umbral se obtiene maximizando la entropía del histograma de los niveles de gris de las clases resultantes. Algunas técnicas de umbralización en una dimensión (1D) aplican métodos de umbralización multiniveles como en [1-5]. A diferencia de los métodos de umbralización en una dimensión, los métodos en dos dimensiones (2D) segmentan la imagen usando información espacial [10-15]. Kirby y Rosenfeld

proponen un método de umbralización 2D que considera simultáneamente los niveles de gris de los píxeles y la estadística local de los píxeles vecinos [12]. Un método particular 2D es el que umbraliza mediante la entropía, este método utiliza la entropía espacial para obtener los umbrales óptimos. En [13] y [14] se propone que los umbrales óptimos pueden ser seleccionados maximizando la entropía a posteriori de dos clases. Sin embargo, su método consume mucho tiempo en determinar la entropía 2D total de las dos clases resultantes. Como resultado, en [10] se propone una variante que reduce la complejidad computacional. Lo mismo se propone en [11], aunque la propuesta no logra la suficiente eficiencia para aplicarse en la umbralización 1D multinivel.

Los estudios de umbralización global descritos en [6] consideran que el método de Otsu es uno de los mejores métodos de selección del umbral, para imágenes del mundo real. Sin embargo el método de Otsu usa una búsqueda exhaustiva para evaluar el criterio de la maximización de la varianza entre clases. En la medida que el número de clases en una imagen aumenta, el método de Otsu toma demasiado tiempo para que se pueda utilizar como una solución práctica en la selección de umbrales multiniveles.

Existen algunas aplicaciones donde, en condiciones normales, solo existe una clase (el fondo). En estos casos, los métodos citados anteriormente no son apropiados porque asumen que existen al menos dos clases. Una solución a este problema se describe en [16], que es el punto de partida para nuestra propuesta.

La referencia [16] propone un método de umbralización adaptativa para la segmentación de imágenes, que es efectiva, aun cuando exista solo una clase (el fondo), pero en condiciones adversas como por ejemplo en presencia de neblina, el algoritmo puede fallar. Para minimizar este problema proponemos introducir una modificación en [16], que reduce el porcentaje de errores en la clasificación de objetos en movimiento.

En la sección II y en la sección III, se describe el método propuesto en [16]. En la sección IV, se propone la modificación al método de umbralización descrito en la sección III. En la sección V se muestran los resultados

experimentales. Finalmente, la sección VI resume las principales conclusiones del trabajo.

II. DETECCION DE MOVIMIENTO

En [16] se propone un método de umbralización adaptativa para la detección de objetos en movimiento. Se obtiene una matriz de transformación U (usando la técnica del Análisis de las Componentes Principales) a partir de imágenes que solo contienen las partes estáticas de la escena vigilada (aunque dichas imágenes pueden tener diferentes condiciones de iluminación). Esta matriz U modela la parte estática de la escena (el fondo). Una nueva imagen captada Φ_j , se proyecta en el espacio transformado y posteriormente es recuperada $\hat{\Phi}_j$.

Determinando la distancia entre la imagen Φ_j y su proyección recuperada $\hat{\Phi}_j$, es posible determinar si existe o no movimiento .

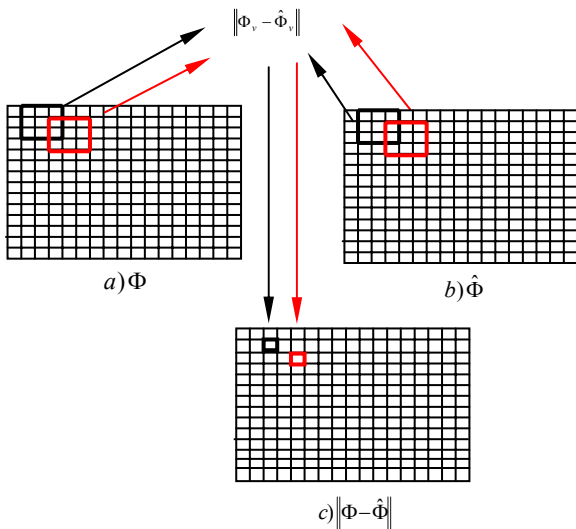


Figura 1. Representación gráfica de la distancia Euclídea d_v para cada ventana.

Para minimizar la influencia del ruido, la distancia Euclídea se obtiene para cada uno de los píxeles de las imágenes, definiendo una ventana centrada en esos píxeles (figura 1a y 1b). El resultado es un “mapa de distancias” donde cada valor del mapa se corresponde con la distancia Euclídea entre las correspondientes ventanas en la imagen original Φ_j y en la imagen recuperada $\hat{\Phi}_j$.

Si se definen ventanas de 3x3 píxeles, y se identifican como Φ_v y $\hat{\Phi}_v$ para la imagen original y recuperada respectivamente, la distancia Euclídea d_v , asociada al píxel central de esas ventanas, queda como:

$$d_v = \|\Phi_v - \hat{\Phi}_v\| = \sqrt{\sum_{i \in V} (\phi_{vi} - \hat{\phi}_{vi})^2} \quad (1)$$

donde V representa la ventana de 3x3 píxeles, y $\phi_{vi}, \hat{\phi}_{vi}$ son los valores de los píxeles dentro de las ventanas para la imagen original y recuperada, respectivamente.

III. UMBRAL ADAPTATIVO

Una vez obtenido el “mapa de distancias” para cada imagen capturada I_j , se deben clasificar las diferentes distancias d_v , como pertenecientes a objetos en movimiento dentro de I_j , o como zonas estáticas. Para esto se debe definir un umbral de manera que las distancias mayores que este umbral se consideren como objetos en movimiento y las distancias menores o iguales a este umbral se consideren como pertenecientes al fondo o parte estática de la imagen. Como resultado, se construye una imagen binaria BW , donde cada píxel toma un valor de 255 si la distancia es mayor que el umbral T_h , o cero en caso contrario.

$$BW = 255 \quad \text{if} \quad \|\Phi_v - \hat{\Phi}_v\| > T_h \quad (2)$$

$$BW = 0 \quad \text{if} \quad \|\Phi_v - \hat{\Phi}_v\| \leq T_h \quad (3)$$

La determinación del umbral T_h es un aspecto de gran importancia, sobre todo para aplicaciones en entornos exteriores, donde las condiciones de iluminación son variables y no se pueden controlar. Si se utilizara un umbral fijo, las variaciones en las distancias Euclídeas debidas a variaciones de las condiciones ambientales produciría errores importantes en el proceso de clasificación.

Para dar respuesta a la problemática anterior, en [16] se propone una solución que tiene en cuenta que la escena a vigilar, tiene dimensiones mucho mayores que los posibles objetos en movimiento. Esto implica que, al dividir las imágenes Φ y $\hat{\Phi}$ en ventanas y calcular sus respectivas distancias Euclídeas, las distancias d_v de las ventanas sin movimiento tendrán mayor frecuencia de aparición y valores relativamente parecidos, mientras que las ventanas con movimiento presentarán valores de distancias, mucho mayores que las ventanas sin movimiento. Con estas premisas se propone obtener el histograma de las distancias Euclídeas $\|\Phi_v - \hat{\Phi}_v\|$ de las diferentes ventanas (figura 2).

A partir del histograma del “mapa de distancias”, se obtiene la distancia de mayor frecuencia de aparición d_{mf} (que debe corresponder con las ventanas Φ_v sin movimiento). Si el histograma del “mapa de distancias” es (h) y el valor de cada distancia es π_i , entonces la distancia de mayor frecuencia de aparición es:

$$d_{mf} = \arg \max_{pi} h(pi) \quad (4)$$

por último se fija el umbral T_h como:

$$T_h = d_{mf} + \beta \quad (5)$$

donde β es una constante que se obtiene en la fase de entrenamiento del sistema.

Para determinar el valor de β , se considera que el

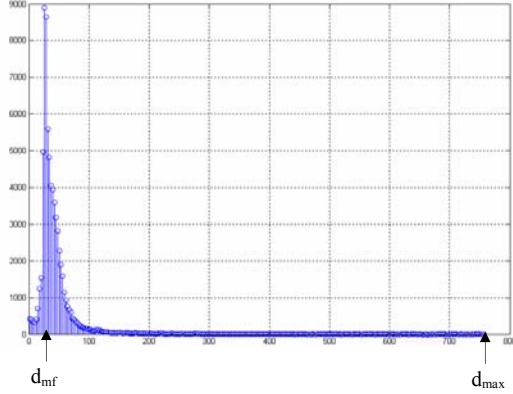


Figura 2. Histograma del "mapa de distancias" de las diferentes ventanas.

universo de trabajo está formado por dos clases, la clase α_1 que representa a las imágenes sin objetos en movimiento y la clase α_2 que describe a las imágenes con objetos en movimiento. Para que estas clases representen lo mejor posible, las variadas condiciones que tienen lugar en las aplicaciones en espacios exteriores, se captaron imágenes con y sin objetos en movimiento, en diferentes condiciones de iluminación.

Una vez escogidas las imágenes pertenecientes a las clases α_1 y α_2 , se aplica a dichas imágenes, el método de detección de movimiento descrito en la sección 3. Como resultado se obtiene para cada imagen, el "mapa de distancias" (figura 1c). Tomando como referencia el histograma de estos "mapas de distancias", se determina, para cada una de las imágenes pertenecientes a las clases α_1 y α_2 , la diferencia Δd entre la máxima distancia y la distancia de mayor frecuencia de aparición, es decir:

$$\Delta d = d_{max} - d_{mf} \quad (6)$$

donde d_{max} es la máxima distancia Euclidea del "mapa de distancias". Como resultado se obtiene un conjunto de valores pertenecientes a las imágenes sin objetos en movimiento Δd_{α_1} y otro conjunto de valores que pertenecen a las imágenes con objetos en movimiento Δd_{α_2} .

$$\Delta d_{\alpha_1} : \{\Delta d_{\alpha_11}, \Delta d_{\alpha_12}, \dots, \Delta d_{\alpha_1R}\}$$

$$\Delta d_{\alpha_2} : \{\Delta d_{\alpha_21}, \Delta d_{\alpha_22}, \dots, \Delta d_{\alpha_2Q}\}$$

siendo R y Q, el número de imágenes pertenecientes a las clases α_1 y α_2 respectivamente, que se emplean en la fase de entrenamiento.

Los centroides de las clases α_1 y α_2 coincidirán con sus medias, es decir:

Siendo R y Q, el número de imágenes pertenecientes a las clases α_1 y α_2 respectivamente, que se emplean en la fase de entrenamiento.

Los centroides de las clases α_1 y α_2 coincidirán con sus medias, es decir:

$$\mu_{\alpha_1} = \frac{1}{R} \sum_{i=1}^R \Delta d_{\alpha_1 i} \quad (7)$$

$$\mu_{\alpha_2} = \frac{1}{Q} \sum_{i=1}^Q \Delta d_{\alpha_2 i} \quad (8)$$

La constante β se escoge como la media de ambos centroides.

$$\beta = \frac{\mu_{\alpha_1} + \mu_{\alpha_2}}{2} \quad (9)$$

IV. UMBRAL ADAPTATIVO MODIFICADO

El método de umbralización adaptativa descrita en la sección III, obtiene la constante β en la fase de entrenamiento del sistema. Debido a que en la fase de entrenamiento no se pueden captar todas las posibles condiciones presentes en los entornos exteriores, se generan algunos falsos positivos, particularmente cuando las condiciones son adversas. Para minimizar el número de falsos positivos, proponemos obtener β a partir de las imágenes usadas para crear el modelo, específicamente, de la imagen media Ψ (ver [16]). Esto permite ajustar β en función de las características particulares del entorno exterior.

Un estudio de varios bancos de imágenes indica que existe una relación directa entre el valor óptimo de β y la varianza de la imagen media Ψ , como se muestra en la figura 3.

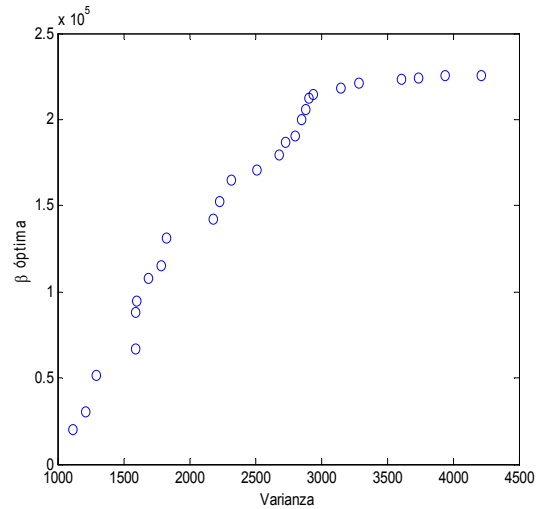


Figura 3. Relación entre el valor óptimo de β y la varianza de la imagen media Ψ .

Por tanto se propone obtener una relación analítica usando un polinomio para ajustar la curva mostrada en la figura 3.

El polinomio propuesto es:

$$p(x) = p_1x^n + p_2x^{n-1} + \dots + p_nx + p_{n+1} \quad (10)$$

donde n es el grado del polinomio.

El objetivo es encontrar los coeficientes del polinomio $p(x)$ en el sentido de los mínimos cuadrados.

En este estudio se usó $n=1$ y $n=2$ para obtener la curva que mejor se ajusta a los datos. Como se muestra en la figura 4, el mejor ajuste es para $n=2$, en este caso la expresión para β es:

$$\beta = p_1(S)^2 + p_2S + p_3 \quad (11)$$

donde S es la varianza de la imagen media Ψ , y p_1, p_2 y p_3 son los coeficientes del polinomio.

$$p_1 = -0.0277; p_2 = 213; p_3 = -182132.98$$

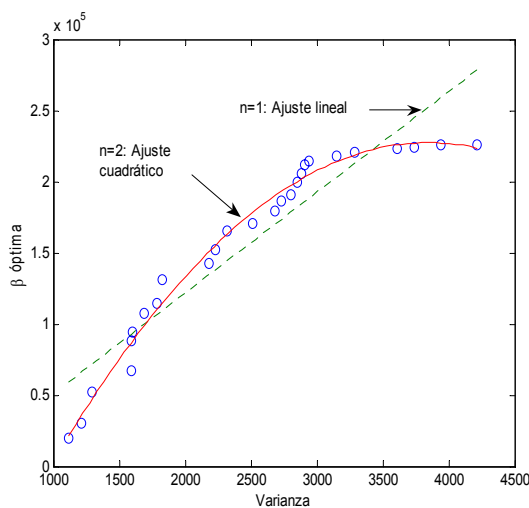


Figura 4. Ajuste de curvas con el polinomio de primer y segundo orden.

V. RESULTADOS EXPERIMENTALES

Para evaluar nuestra propuesta, se capturaron imágenes reales en entornos ferroviarios, en condiciones ambientales y de iluminación variadas, algunas de ellas en condiciones muy adversas como es el caso de imágenes con neblina (figura 6a). Se utilizaron 26 bancos de imágenes para comparar el método descrito en la sección III con la modificación propuesta en este trabajo. Los bancos de imágenes se procesaron por ambos métodos y los resultados alcanzados se muestran en la figura 5. Como se puede apreciar, la alternativa descrita en la sección IV reduce el porcentaje de errores en la clasificación de objetos en movimiento.

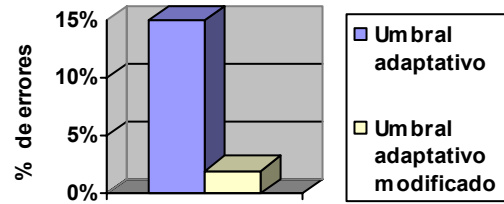
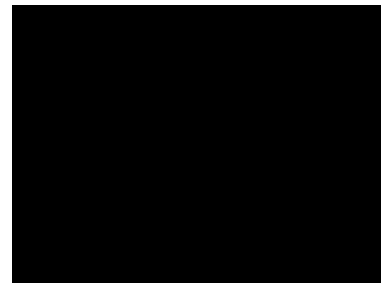


Figura 5. Comparación entre el método de umbral adaptativo y umbral adaptativo modificado.

Un ejemplo de la mejora del método de umbral adaptativo modificado (sección IV), respecto al método de umbral adaptativo (sección III) se muestra en la figura 6. En esta figura existe un peatón en la imagen captada (figura 6a), el cual no es detectado por el método de umbral adaptativo (figura 6b) pero si es detectado por el método de umbral adaptativo modificado (figura 6c).



a)



b)



c)

Figura 6. a) Imagen original con un peatón.
b) Resultado usando el método de umbral adaptativo.
c) Resultado usando el método de umbral adaptativo modificado.

VI. CONCLUSIONES

El método de umbralización adaptativa modificado propuesto en este trabajo minimiza el problema que tienen varios métodos de umbralización, cuando solo existe una clase presente en las imágenes capturadas.

El uso del histograma del “mapa de distancias” y la varianza de la imagen media Ψ para seleccionar el umbral, es otra mejora del método propuesto, porque de esta manera, la selección del umbral tiene en cuenta las variadas condiciones climatológicas y de iluminación presentes en las aplicaciones en entornos exteriores.

AGRADECIMIENTOS

Este trabajo ha sido realizado en el marco del proyecto PARMEI, financiado por el Ministerio de Ciencia y Educación (DIP 2003-08715-C02-01); y por las empresas *RENFE* y *LOGYTEL* (Cátedra de “Control Electrónico” aplicado al transporte, del Departamento de Electrónica de la Universidad de Alcalá).

REFERENCIAS

- [1]. D. M. Tsai and Y. H. Chen, “A fast histogram-clustering approach for multilevel thresholding,” *Pattern Recognition Letters*, Vol. 13, No. 4, 1992, pp. 245-252.
- [2]. N. Otsu, “A threshold selection method from gray-level histogram,” *IEEE Transactions on System Man Cybernetics*, Vol. SMC-9, No. 1, 1979, pp. 62-66.
- [3]. W. H. Tsai, “Moment-preserving thresholding: a new approach,” *Computer Vision, Graphics, and Image Processing*, Vol. 29, 1985, pp. 377-393.
- [4]. J. C. Yen, F. J. Chang, and S. Chang, “A new criterion for automatic multilevel thresholding,” *IEEE Transactions on Image Processing*, Vol. 4, No. 3, 1995, pp. 370-378.
- [5]. S. Wang and R. Haralick, “Automatic multithreshold selection,” *Computer Vision, Graphics, and Image Processing*, Vol. 25, 1984, pp. 46-67.
- [6]. P. K. Sahoo, S. Soltani, A. K. C. Wong, and Y. Chen, “A survey of thresholding techniques,” *Computer Vision Graphics Image Processing*, Vol. 41, 1988, pp. 233-260.
- [7]. T. Pun, “A new method for gray-level picture thresholding using the entropy of the histogram,” *Signal Processing*, Vol. 2, 1980, pp. 223-237.
- [8]. J. N. Kapur, P. K. Sahoo, and A. K. C. Wong, “A new method for gray-level picture thresholding using the entropy of the histogram,” *Computer Vision Graphics Image Processing*, Vol. 29, 1985, pp. 273-285.
- [9]. S. U. Lee and S. Y. Chung, “A comparative performance study of several global thresholding techniques for segmentation,” *Computer Vision Graphics Image Processing*, Vol. 52, 1990, pp. 171-190.
- [10]. W. T. Chen, C. H. Wen, and C. W. Yang, “A fast two-dimensional entropic thresholding algorithm,” *Pattern Recognition*, No. 27, No. 7, 1994, pp. 885-893.
- [11]. J. Gong, L. Li, and W. Chen, “Fast recursive algorithms for two-dimensional thresholding,” *Pattern Recognition*, Vol. 31, No. 3, 1998, pp. 295-300.
- [12]. R. L. Kirby and A. Rosenfeld, “A note on the use of (gray, local average gray level) space as an aid in thresholding selection,” *IEEE Transactions on System Man Cybernetics* Vol. SMC-9, No. 12, 1979, pp. 860-864.
- [13]. A. S. Abutaleb, “Automatic thresholding of gray-level pictures using two-entropy,” *Computer Vision Graphics Image Processing*, Vol. 47, 1989, pp. 22-32.
- [14]. N. R. Pal and S. K. Pal, “Entropic thresholding,” *Signal Processing*, Vol. 16, 1989, pp. 97-108.
- [15]. A. D. Brink, “Thresholding of digital images using two-dimensional entropies,” *Pattern Recognition*, Vol. 25, No. 8, 1992, pp. 803-808.
- [16]. J. Vázquez et. al., “Detection of moving objects in railway using vision,” *IEEE Intelligent Vehicles Symposium*, Parma, Italy, June 14-17, 2004.