

## Reconocimiento de objetos parcialmente ocultos con filtros de correlación no lineales

Martínez-Díaz S.\*, Barranco-Gutiérrez A\*.

\* Instituto Tecnológico de La Paz, La Paz BCS, México  
(e-mail: [saulmd@itlp.edu.mx](mailto:saulmd@itlp.edu.mx), [abarrancog0900@ipn.mx](mailto:abarrancog0900@ipn.mx)).

---

**Resumen:** En este artículo se propone el diseño de filtros no lineales compuestos para el reconocimiento de objetos parcialmente ocultos. Los filtros propuestos se combinan con la operación de correlación morfológica para lograr un sistema de reconocimiento robusto al ruido y con una buena capacidad de discriminación. El desempeño de la técnica propuesta es comparado con el del filtrado lineal clásico mediante simulaciones por computadora.

*Palabras claves:* Filtros compuestos, correlación morfológica, reconocimiento de patrones, filtros no lineales, sistemas robustos al ruido.

---

### 1. INTRODUCCIÓN

Uno de los grandes retos para un sistema de reconocimiento de imágenes es la detección de objetos aún cuando estos se encuentran incompletos o parcialmente ocultos por otros objetos. Desde hace algunos años los filtros de correlación han sido usados ampliamente para el reconocimiento de imágenes. Esto se debe principalmente a que sus fundamentos matemáticos son sólidos. Además, la correlación es invariante a traslación y explota toda la información contenida en la imagen. Bajo este enfoque el reconocimiento se lleva a cabo calculando una función de correlación entre la escena de prueba y una plantilla (llamada filtro) previamente diseñada. El filtro más básico se forma únicamente con la imagen del objeto buscado (VanderLugt, 1964). Desafortunadamente este tipo de filtro puede fallar cuando el objeto buscado dentro de la escena tiene algún tipo de distorsión geométrica, como pueden ser la rotación o el escalamiento. En esos casos es preferible utilizar un filtro compuesto por varias imágenes. Una manera de diseñar filtros compuestos se basa en el uso de Funciones Discriminantes Sintéticas (Hester and Casasent, 1980). El filtro se diseña para generar un valor de correlación previamente especificado para cada imagen, y puede contener imágenes representativas de las distorsiones esperadas en la escena. Adicionalmente, para evitar

la formación de lóbulos laterales en la salida, se puede minimizar la energía promedio en el plano de correlación (Mahalanobis et al., 1987). Este último filtro es conocido como filtro de Mínima Energía Promedio de Correlación (MACE por sus siglas en inglés). Utilizando estos principios se propuso el diseño de filtros compuestos para la detección de objetos parcialmente ocultos (González-Fraga et al., 2005). Los filtros mencionados anteriormente utilizan técnicas lineales para su diseño y el cálculo de correlación. Una desventaja de esta técnica es que los filtros son sensibles al ruido con distribuciones de probabilidad no gaussiana, el cual suele encontrarse en imágenes reales.

Por otro lado, existen también diferentes tipos de filtros no lineales para reconocimiento de imágenes (Doh et al. 2004; Wang et al., 2006). Recientemente se propusieron filtros no lineales compuestos, cuyo diseño se basa en operaciones lógicas (Martínez-Díaz and Kober, 2008). La técnica de filtrado utiliza una operación no lineal llamada Correlación Morfológica la cual resulta de minimizar el error medio absoluto (Maragos, 1989). Este tipo de filtrado es robusto al ruido y ha demostrado buena capacidad de discriminación entre objetos de diferentes clases. Una desventaja de estas técnicas es su alto costo computacional, no obstante, es posible reducir su costo con sistemas paralelos (Martínez-Díaz 2012).

En este artículo se propone el uso de filtros no lineales compuestos para el reconocimiento de objetos parcialmente ocultos y sometidos a ruido de naturaleza no gaussiana. El desempeño de los filtros es probado y comparado con el de su contraparte lineal en términos de capacidad de discriminación y robustez al ruido. El artículo se organiza como sigue: La sección 2 es una breve revisión de los filtros lineales compuestos. La sección 3 presenta el método propuesto. En la sección 4 se presentan y discuten los resultados de las simulaciones por computadora. La sección 5 contiene las conclusiones.

### 2. FILTROS LINEALES COMPUESTOS

Los filtros lineales compuestos basados en Funciones Discriminantes Sintéticas se diseñan de la siguiente manera: Suponga que se tienen  $N$  imágenes de entrenamiento, cada una con  $d$  píxeles. Cada imagen en 2D se convierte en un vector columna, en orden lexicográfico. Cada vector se inserta en una matriz  $\mathbf{R}$  de dimensiones  $dxN$ . También se genera un vector columna  $\mathbf{u}$  con  $N$  elementos. Cada elemento de  $\mathbf{u}$  es el valor de correlación deseado en la salida, correspondiente a cada imagen de entrenamiento. De este modo se puede discriminar entre múltiples clases de objetos, asignando a todos los objetos de la misma clase el mismo valor de salida. Si la matriz  $(\mathbf{R}^+\mathbf{R})$  es no singular, se puede obtener el filtro compuesto  $\mathbf{h}_{SDF}$  con:

$$\mathbf{h}_{SDF} = \mathbf{R}(\mathbf{R}^+\mathbf{R})^{-1}\mathbf{u} \quad (1)$$

Donde el superíndice  $+$  significa transpuesta conjugada. Sin embargo, este filtro tiene la desventaja de generar lóbulos laterales en la salida de correlación, lo que disminuye su capacidad de discriminación. Esto se debe a que no se controla el valor de salida en todo el plano de correlación. Una solución a este problema es minimizar la energía promedio de correlación, satisfaciendo al mismo tiempo las condiciones de correlación impuestas para cada imagen de entrenamiento. Con ello se obtiene el llamado filtro MACE, el cual es uno de los más efectivos filtros lineales compuestos. Para este caso se procede como sigue: Se obtiene la transformada de Fourier de cada imagen de entrenamiento y se convierte en un vector columna. Con los vectores columna se forma una matriz  $\mathbf{X}$ . La respuesta en frecuencia del filtro MACE puede expresarse como:

$$\mathbf{h}_{MACE} = \mathbf{D}^{-1}\mathbf{X}(\mathbf{X}^+\mathbf{D}^{-1}\mathbf{X})^{-1}\mathbf{u} \quad (2)$$

Donde  $\mathbf{D}$  es una matriz diagonal de dimensiones  $dx d$  que contiene el espectro promedio de potencia de las imágenes de entrenamiento.

### 3. MÉTODO PROPUESTO

El método propuesto utiliza una ventana deslizante que contiene la señal de referencia (filtro). La correlación se calcula para el pixel central de la ventana móvil, considerando una vecindad espacial  $W$  de píxeles que lo rodean. La forma de la vecindad es similar a la región de soporte (píxeles con un valor diferente de cero) del filtro. El número de píxeles dentro de la vecindad es representado por  $|W|$ , y es aproximadamente del mismo tamaño de la región de soporte del filtro.

Para diseñar el filtro a cada imagen  $I(k,l)$  con  $Q$  niveles de gris se le aplica una descomposición por umbral (Fitch et al., 1984) de la siguiente forma:

$$I^q(k,l) = \begin{cases} 1, & \text{si } I(k,l) \geq q \\ 0, & \text{en otro caso} \end{cases} \quad (3)$$

Donde  $(k,l)$  son las coordenadas del pixel y  $\{I^q(k,l), q=1\dots Q\}$  es un conjunto de imágenes binarias obtenidas de la descomposición en cada nivel  $q$ . La imagen original se puede recuperar con:

$$I(k,l) = \sum_{q=1}^{Q-1} I^q(k,l) \quad (4)$$

Ahora supóngase que se tienen  $N$  objetos de la clase que se desea reconocer  $\{T_i(k,l), i=1\dots N\}$ , llamada la clase verdadera, y  $M$  objetos de una clase a rechazar  $\{F_i(k,l), i=1\dots M\}$ , llamada la clase falsa. El  $i$ -ésimo filtro no lineal compuesto  $H_i$  se obtiene con:

$$H_i(k,l) = \sum_{q=1}^{Q-1} T_i^q(k,l) \cap \left[ \bigcup_{j=1}^M F_j^q(k,l) \right], \quad i=1\dots N \quad (5)$$

En este caso  $\{T_i^q(k,l), q=1,\dots,Q-1, i=1,\dots,N\}$  y  $\{F_i^q(k,l), q=1,\dots,Q-1, i=1,\dots,M\}$  representan las imágenes binarias obtenidas de la descomposición por umbral aplicada a las imágenes de entrenamiento de las clases verdadera y falsa, respectivamente.  $\cup$  y  $\cap$  representan las operaciones lógicas de unión e intersección, respectivamente.

Ahora definimos la operación de correlación morfológica normalizada entre la escena de prueba  $S$  y un filtro  $H$  como:

$$C(k,l) = \sum_{m,n \in W} \text{MIN}[a_{k,l}S(m+k,n+l) + b_{k,l}, H(m,n)] \quad (6)$$

Donde  $\text{MIN}(X,Y)$  representa el mínimo entre  $X$  e  $Y$ . La suma es tomada dentro de la vecindad  $W$ . Los coeficientes locales  $a_{k,l}$  y  $b_{k,l}$  se incluyen para tomar en cuenta las diferencias en iluminación de las imágenes. Estos coeficientes pueden obtenerse minimizando el error cuadrático medio entre la escena de prueba y el filtro:

$$a_{k,l} = \frac{\sum_{m,n \in W} H(m,n)S(m+k,n+l) - |W|\bar{H}\bar{S}_{k,l}}{\sum_{m,n \in W} (S(m+k,n+l))^2 - |W|(\bar{S}_{k,l})^2} \quad (7)$$

$$b(k,l) = \bar{H} - a(k,l) \cdot \bar{S}(k,l) \quad (8)$$

Aquí  $\bar{H}$  y  $\bar{S}_{k,l}$  son los valores promedio del filtro y la escena dentro de la ventana local calculados respecto a  $W$  en la  $(k,l)$ -ésima posición, respectivamente. El valor final de correlación se obtiene aplicando la siguiente expresión:

$$\hat{C}(k,l) = \text{MAX} \left\{ \left\{ \frac{u}{t_i} C_i(k,l), i = 1 \dots N \right\} \right\} \quad (9)$$

Donde  $C_i(k,l)$  es la  $i$ -ésima correlación obtenida entre el  $i$ -ésimo filtro no lineal  $H_i$  (ecuación 5) y la escena de prueba, al aplicar la ecuación 8.  $\text{Max}(X_i)$  es el valor máximo entre las  $X_i$ . El escalar  $u$  es el valor deseado de correlación a la salida y  $t_i$  es un factor de normalización que se puede calcular por:

$$t_i = \sum_{k,l \in W} \sum_{q=1}^{Q-1} T_i^q(k,l) \cap \left[ \bigcup_{j=1}^M F_j^q(k,l) \right] \quad (10)$$

### 3. RESULTADOS EXPERIMENTALES

En esta sección se presentan algunos resultados experimentales. Se compara el desempeño de los filtros propuestos con el de los filtros lineales MACE. Como medida de desempeño se utiliza la Capacidad de Discriminación (DC, por sus siglas en inglés). La DC se define formalmente como la habilidad de un filtro para distinguir un objeto de una clase entre otros objetos de diferente clase (Yaroslavsky, 1993), y puede expresarse como:

$$DC = 1 - \frac{|C^B(0,0)|^2}{|C^O(0,0)|^2} \quad (11)$$

En este caso  $C^O(0,0)$  es el valor máximo en el plano de correlación sobre el área donde se encuentra el objeto buscado.  $C^B(0,0)$  es el valor máximo en el plano de correlación fuera de área del objeto buscado. El área del objeto a ser reconocido es determinada por su región de soporte. La DC puede alcanzar un valor máximo de uno. Un valor negativo de DC significa que el objeto buscado no fue reconocido.

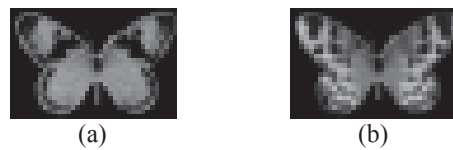


Fig. 1. Objetos de entrenamiento utilizados para el diseño de los filtros: (a) objeto de la clase verdadera y (b) objeto de la clase falsa.

La figura 1 muestra dos objetos similares los cuales fueron utilizados para diseñar cada filtro. La mariposa de la izquierda es el objeto buscado, la otra mariposa es incluida en los filtros como un objeto de la clase falsa. Ambas imágenes son de 24x36 pixeles. Para el diseño de los filtros cada imagen de entrenamiento se dividió verticalmente por la mitad. Con las imágenes obtenidas se formaron el filtro lineal  $h_{\text{MACE}}$ , según la ecuación 2, y el filtro no lineal  $H$ , de acuerdo a la ecuación 5.



Fig. 2. Escena de prueba incluyendo un bosque de fondo.

En la figura 2 se presenta una escena de prueba que contiene las dos mariposas mostradas en 1. La escena

de prueba es de 256x256 píxeles. Todas las imágenes son de 256 niveles de gris. Es importante destacar que el bosque de fondo en la escena de prueba dificulta el reconocimiento aún sin ruido ni oclusión.

Los experimentos realizados consisten en el cálculo de la capacidad de discriminación de los filtros diseñados en diferentes condiciones de ruido y con diferente porcentaje de oclusión del objeto buscado. El ruido utilizado es impulsivo del tipo sal y pimienta con igual probabilidad de ocurrencia de ambos tipos de impulso.



Fig. 3. Escena de prueba con el objeto buscado oculto en un 65%.



Fig. 4. La misma escena que en la figura 3 pero contaminada con 10% de ruido impulsivo.

En la figura 3 se puede observar la escena de prueba con un 65% de oclusión del objeto buscado, mientras el objeto de la clase falsa se encuentra completo. Nótese que el objeto buscado se aprecia con dificultad. En la figura 4 se presenta un ejemplo de la misma escena contaminada con 10% de ruido impulsivo. En los experimentos se calculó la capacidad de discriminación de cada filtro para diferentes porcentajes de oclusión. Para garantizar resultados estadísticamente correctos, para cada caso de prueba, se aplicaron y promediaron 30 realizaciones de ruido aleatorio uniformemente distribuido en la imagen. Debido a la alta cantidad de ruido, es difícil para una persona distinguir ambos objetos. Por otro lado, la naturaleza del ruido impulsivo ocasiona pérdida total de la información de algunos píxeles, dificultando el reconocimiento de forma automática.

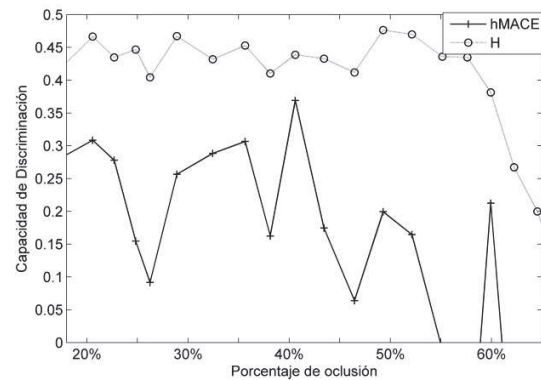


Fig. 5. DC vs porcentaje de oclusión con 5% de ruido impulsivo.

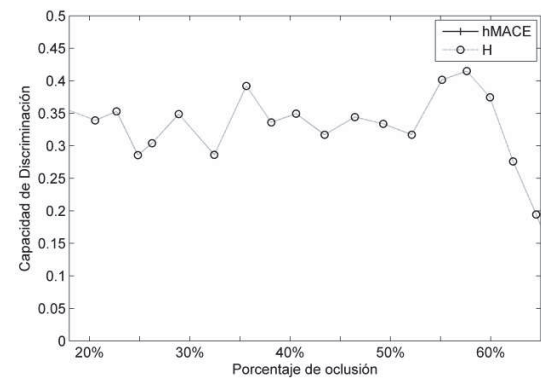


Fig. 6. DC vs porcentaje de oclusión con 10% de ruido impulsivo.

En las figuras 5 y 6 se muestran las gráficas de DC en función del porcentaje de oclusión con 5% y 10% de ruido, respectivamente. Como se puede apreciar, con 5% de ruido el filtro no lineal H presenta un desempeño superior al del filtro hMACE. Por otro lado, como muestra la figura 6, el filtro lineal es incapaz de

reconocer el objeto cuando se incrementa la cantidad de ruido al 10%, aún con mínima oclusión. Por su parte el filtro no lineal propuesto es capaz de reconocer el objeto en esas condiciones hasta con un 65% de oclusión.

#### 4. CONCLUSIONES

En el presente artículo se propuso el diseño de filtros no lineales de correlación para el reconocimiento de objetos parcialmente ocultos. Los filtros propuestos ofrecen una buena capacidad de discriminación y robustez al ruido a costa de un mayor costo computacional. Sin embargo, es posible paralelizar los algoritmos para reducir el tiempo de procesamiento.

#### AGRADECIMIENTOS

El presente trabajo fue desarrollado con apoyo del CONACyT y el Instituto Tecnológico de La Paz.

#### REFERENCIAS

- Doh, Y. H., Kim, J. C., Kim, J. W., Choi, K. H., Kim, S. J. and Alam, M. S., (2004). Distortion-invariant pattern recognition based on a synthetic hit-miss transform. *Opt. Eng.* Vol. 43, pp. 1798-1803.
- Fitch, J. P., Coyle, E. J. and Gallagher, N. C. Jr., (1984). Median filtering by threshold decomposition. *IEEE Trans. Acoust. Speech Signal Process.* Pp. 1183-1188.
- González-Fraga, J. Á., Kober, V. and Álvarez-Borrego, J., (2004). Recognition of partially occluded objects using correlation filters with training. In: *Proceedings of SPIE*, Vol. 5558.
- Hester, C. F. and Casasent, D., (1980). Multivariant technique for multiclass pattern recognition. *Appl. Opt.* Vol.19, pp. 1758-1761.
- Mahalanobis, A., Vijaya-Kumar, B. V. K. and Casasent, D., (1987). Minimum average correlation energy filters. *Appl. Opt.* Vol. 31, pp. 1823-1833.
- Maragos, P., (1989). Morphological correlation and mean absolute error criteria. In: *Proc. conf. IEEE Trans. Acoust. Speech Signal Process.* Pp. 1568-1571.
- Martínez-Díaz, S. and Kober, V., (2008). Nonlinear synthetic discriminant function filters for illumination-invariant pattern recognition. *Opt. Eng.* Vol. 47, pp. 067201-1-9.
- Martínez-Díaz, S., (2012). Parallel architecture for binary images recognition. In: *Proceedings of the World Congress on Engineering.* Vol. II, pp. 856-859.
- VanderLugt, A. B., (1964). Signal detection by complex filtering. *IEEE Trans. Inf. Theory.* Vol. 10, pp. 139-145.
- Wang, Q., Deng, Y. and Liu, S., (2006). Morphological fringe-adjusted joint transform correlation. *Opt. Eng.* Vol. 45, pp. 087002-1-9.
- Yaroslavsky L. P., (1993). The theory of optimal methods for localization of objects in pictures. In: *Progress in Optics*, E. Wolf, Ed., Vol. XXXII, pp. 145–201, Elsevier.