

# Reconocimiento de formas con un correlador óptico aplicado a imágenes desenfocadas: invariancia por medio de la selección y fusión de bandas

A. Vargas y R. Figueroa

*Universidad de La Frontera, Departamento de Ciencias Físicas,  
Av. Francisco Salazar 01145 Temuco - Chile,  
e-mail: avargas@ufro.cl*

J. Campos

*Universidad Autónoma de Barcelona, Departamento de Física,  
Edificio Cc, E-08193 Bellaterra, Barcelona, España,  
e-mail: juan.campos@uab.es*

C. San Martín y J. Marileo

*Universidad de La Frontera, Departamento de Ingeniería Eléctrica,  
Av. Francisco Salazar 01145, Temuco - Chile,  
e-mail: csmarti@ufro.cl*

Recibido el 19 de septiembre de 2005; aceptado el 26 de abril de 2006

In this work, experimental result measuring the performance of a composite filter designed to solve the optical pattern recognition problem of defocus images is presented. We apply a methodology that considers optical decomposition of the input scene in frequency bands. In each band a pattern recognition filter is applied. Finally this information is fused by means of addition or geometric mean. We analyze both the individual information, and the fusion scheme. We also present results when total spectral information of the optical pattern recognition process is taken in to account.

*Keywords:* Pattern recognition; optical pattern recognition; defocused pattern recognition; optical correlators.

En este trabajo se muestran los resultados experimentales correspondientes al rendimiento de un filtro óptico compuesto diseñado para resolver el problema del reconocimiento óptico de imágenes desenfocadas. Aplicamos una metodología que considera la descomposición óptica de una escena de entrada en diferentes bandas de frecuencias. En cada banda se aplica un filtro de reconocimiento. Finalmente, la información es fusionada mediante la adición o la media geométrica. Se analizó tanto la información individual como la fusionada. También presentamos resultados cuando la información espectral total es tomada en cuenta en el proceso de reconocimiento.

*Descriptores:* Reconocimiento de formas; reconocimiento óptico de formas; reconocimiento de forma con desenfoque; correladores ópticos.

PACS: 42.30.S; 42.30; 42.79.C

## 1. Introducción

Cuando la imagen de entrada tiene alguna degradación, la calidad del proceso de reconocimiento de formas empeora. Para mejorar el proceso se diseña un filtro que no sólo resuelve el problema de reconocimiento en sí, sino que también toma en cuenta esta degradación de la imagen de entrada. En la literatura han sido diseñados diversos tipos de filtros y metodologías para resolver diferentes tipos de degradación [1-4]. Una degradación común que se presenta en una imagen, captada por un sistema de adquisición de imágenes, es el desenfoque. En este caso, al utilizar una imagen desenfocada, la calidad del proceso de reconocimiento se deteriora y no hay confiabilidad en los resultados. Así, a medida que el desenfoque de la imagen de entrada aumenta, los resultados son cada vez menos confiables. Para resolver el problema del desenfoque en el proceso del reconocimiento digital de formas, varios autores [5-7] han desarrollado metodologías numéricas basadas en correlaciones cuyos filtros sólo están definidos en regiones de soporte, obtenidas éstas últimas mediante una descomposición de la imagen de entrada en bandas de frecuencias.

El procesado de imágenes con la incorporación de sistemas ópticos otorga innegables ventajas de velocidad y paralelismo [8,10] para realizar operaciones de correlación y/o convolución, y, así potenciar técnicas de reconocimiento óptico de formas. Los sistemas apropiados para realizar esta tarea son los correladores ópticos coherentes. Esta ventaja se incrementa si en el sistema óptico se utilizan pantallas de cristal líquido (*liquid crystal displays*, LCD) para introducir las escenas y los filtros, debido a que sería posible realizar análisis en tiempo real.

La calidad del proceso de reconocimiento de formas puede medirse en base a diferentes criterios, y en general cada filtro es diseñado para optimizar algún criterio [11]. Por ejemplo, el filtro adaptado clásico (*classical matched filter*, CMF) optimiza el criterio de señal ruido (*signal to noise ratio*, SNR). Además, en el diseño de filtros es necesario adaptar el filtro a la señal que se desea detectar. Existe en la literatura otra clase de filtros que permiten adaptar al filtro a un conjunto de señales o imágenes, llamados funciones sintéticas discriminantes (*synthetic discriminant function*, SDF) [12]. En su diseño, es posible optimizar más de un criterio, sin embar-

go, hay que tener presente que cuando un criterio se optimiza hay otro que empeora. En el estudio de diseño de filtros para reconocimiento se ha desarrollado un filtro que permite no sólo considerar un conjunto de imágenes de entrenamiento, sino que además permite optimizar varios criterios mediante un parámetro de balance, este filtro es el llamado filtro de compromiso (*trade-off filter*, TOF) [13].

En este trabajo se presenta la implementación del diseño de un filtro compuesto, para realizar experimentalmente un reconocimiento de formas en un correlador óptico coherente, con imágenes a analizar que presentan desenfoques moderados. En este proceso de reconocimiento la capacidad de discriminación (DC) disminuye a medida que el grado del desenfoque de la imagen aumenta. Este trabajo está basado en nuestro estudio previo de un procedimiento digital para resolver el problema del desenfoque [7]. En dicho estudio no hay consideraciones asociadas a problemas de iluminación y ruido. El método está basado en una descomposición de la información en bandas de frecuencias espaciales. En cada una de las bandas se diseña un filtro con el que se obtiene una correlación, la cual entrega una información parcial, que puede ser usada individualmente o de manera fusionada mediante alguna operación. Cada filtro del banco, es un filtro de compromiso (TOF) diseñado en una región de soporte anular. Este diseño considera la optimización de los criterios de señal ruido (*signal to noise ratio*, SNR) y energía del pico de correlación (*peak correlation energy*, PCE).

Este artículo está estructurado como sigue. En la Sec. 2 se presenta el sistema óptico usado en este experimento, las consideraciones para obtener el conjunto de prueba de las imágenes desenfocadas, el criterio para medir la calidad del proceso: capacidad de discriminación y la metodología empleada. Los resultados experimentales y la discusión se presentan en la Sec. 3 y finalmente las conclusiones en la Sec. 4.

## 2. Sistema experimental

Un correlador óptico es un sistema basado en un conjunto de lentes, con el que se obtiene sucesivos planos de transformadas de Fourier (FT). En este artículo se ha utilizado como sistema experimental un correlador convergente, que tiene la ventaja de permitir ajustar la escala de la transformada de Fourier óptica. Esto es esencial en un correlador óptico debido a que es posible sintonizar la dimensión de la transformada de Fourier a la del filtro digital. En la Fig. 1 se muestra un esquema simplificado del correlador. Se utiliza una fuente puntual de luz láser He-Ne de longitud de onda  $\lambda = 638.8$  nm (1). Mediante la primera lente (2) se obtiene la imagen de la transformada de Fourier óptica de la escena (3) a analizar, en el plano del filtro (4). El filtro permite modificar la información espectral de la escena de entrada. Una segunda lente ubicada en (5), colecta la información del filtrado. En el plano imagen (6) de la escena de entrada (3) a través de la lente (5), se sitúa una cámara CCD (6), donde finalmente se captura la información de la correlación en el dominio espacial.

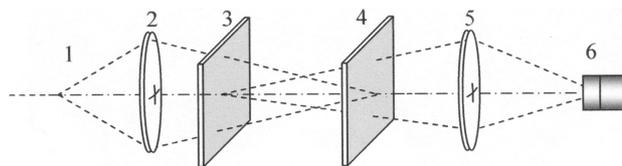


FIGURA 1. Esquema de un correlador óptico convergente: Luz láser (1), lentes convergentes (2) y (5), escena de entrada (3), banco de filtros compuestos (4) y cámara CCD (6).

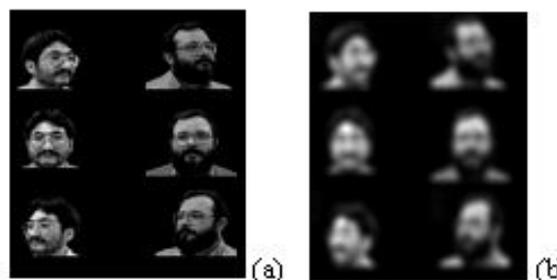


FIGURA 2. Fracción de la escena de entrada original, que contiene dos clases de objetos. Por cada clase hay tres posiciones angulares. (a) Columna izquierda corresponde a objetos a reconocer y columna derecha a rechazar. (b) imagen desenfocada con  $W_{2,0} = 1.3$ .

### 2.1. Escenas de entrada desenfocadas

Para probar el proceso de reconocimiento de formas en el correlador óptico, independiente al desenfoque de la imagen a analizar, se utiliza como imagen de entrada un conjunto de imágenes previamente desenfocadas. La imagen original es mostrada en la Fig. 2a, generada con un software de imágenes comercial. La imagen contiene dos clases de objetos, una clase en el lado izquierdo y la otra en el lado derecho. Cada clase tiene tres posiciones angulares, luego hay seis objetos en total. El conjunto de imágenes desenfocadas son obtenidas modelando el grado de desenfoque bajo la consideración de un sistema de adquisición de imagen con iluminación incoherente [6]. Las imágenes desenfocadas las obtenemos mediante la correlación entre la imagen original y la imagen difraccional de un punto desenfocado. Este último se puede obtener, en el dominio de Fourier, mediante la autocorrelación de la función pupila  $P(u, v)$ , y si el desenfoque es la única aberración presente en el sistema óptico, la función pupila toma la forma  $P(r) = \exp(2\pi i W_{2,0} r^2)$ , y el grado de desenfoque está dado por el parámetro  $W_{2,0}$  [14]. A mayores valores de  $W_{2,0}$  mayor es el desenfoque. En la Fig. 2b se muestra una imagen fuera de foco, con parámetro de desenfoque de 1.3. En esta imagen es posible observar la ausencia de detalles de cada objeto. Para nuestro correlador, se utilizaron escenas de  $256 \times 256$  *pixels*, implementadas en un sustrato de transparencia plástico. Con este conjunto de imágenes desenfocadas se realiza el reconocimiento óptico de formas.

## 2.2. Filtros compuestos para el reconocimiento de formas

Si la única degradación de la imagen de entrada es el desenfoque, la imagen desenfocada en el dominio de Fourier es el producto entre la FT de la imagen sin desenfoque y la función de transferencia óptica del sistema de adquisición (*optical transfer function*, OTF). En este caso la OTF tiene simetría radial. Por otra parte, la capacidad de discriminación (DC) [7] no sólo depende de la similaridad de los objetos a reconocer sino también dependerá de la OTF. Luego, si seleccionamos una región anular delgada con un valor constante de la OTF, en dicha región, la DC es independiente de la OTF y por lo tanto del desenfoque. Entonces nuestro diseño del filtro compuesto, considera un filtro de reconocimiento que toma en cuenta sólo la información de frecuencia en el interior de la región anular.

Teóricamente podemos considerar un anillo con un ancho de banda de un *pixel*, pero en un sistema óptico experimental, sería difícil capturar la información óptica del proceso cuando el ancho de banda es muy delgado. Por ello, incrementamos el ancho de la banda para mejorar los niveles de intensidades de luz. Note que a mayores frecuencias en el dominio de Fourier, la intensidad de la señal decrece con lo que la anchura del anillo debería aumentar.

El filtro de reconocimiento es el filtro de compromiso (TOF), el que considera un balance entre los criterios SNR y PCE. En este caso el parámetro de balance utilizado es de 0.999. El filtro se adapta para reconocer a los objetos de la columna izquierda en la Fig. 2a, y para rechazar a los objetos de la columna derecha. Las regiones de soporte basadas en anillos pasa-banda tiene frecuencias centrales de  $f_i = 37, 67, \text{ y } 106$  para  $i = 1, 2, 3$ , y un ancho variable de  $\Delta f = 26, 34 \text{ y } 44$  *pixeles*, respectivamente. Para cada anillo se diseña un filtro TOF teniendo en cuenta la información correspondiente de los objetos a reconocer y a rechazar. La selección del ancho de banda de la región de soporte está basada en la consideración de tener suficiente iluminación en la cámara CCD del correlador óptico, con la fuente de luz láser que se dispone. Varios experimentos se realizaron para encontrar tal selección de los ancho de banda [15]. Note que a mayores frecuencias centrales decrece la iluminación. También a menores frecuencias centrales la iluminación aumenta. La selección de las frecuencias centrales está basada en consideraciones de no contar con superposición de la información.

## 2.3. Calidad del proceso de reconocimiento

Para medir el proceso del reconocimiento de formas, utilizamos el criterio de la capacidad de discriminación (DC) que mide la habilidad de discriminación del filtro diseñado, en términos de las intensidades de los máximos de correlación del plano de salida. Así, tenemos

$$DC = 1 - \frac{\max \left\{ |c_f^2| \right\}}{\max \left\{ |c_o^2| \right\}}, \quad (1)$$

donde  $c_o$  y  $c_f$  son los valores de los máximos de correlación de los objetos a reconocer y a rechazar, respectivamente. Valores de la DC cercanos a 1 indican una buena discriminación, y cercanos a cero indican una discriminación pobre del objeto a ser reconocido. También, definimos otro criterio basado en la DC: la mejor DC (bDC) y la peor DC (wDC):

$$bDC = 1 - \frac{\min \left\{ |c_f^2| \right\}}{\max \left\{ |c_o^2| \right\}}, \quad wDC = 1 - \frac{\max \left\{ |c_f^2| \right\}}{\min \left\{ |c_o^2| \right\}}, \quad (2)$$

donde *max* y *min* significan el valor más alto y el valor más bajo del máximo de intensidad obtenido en el plano de correlación capturado por la cámara CCD.

## 3. Resultados

En esta sección se presentan los resultados experimentales utilizando un conjunto de imágenes desenfocadas. Estas imágenes de entrada están rotuladas por su parámetro de desenfoque  $W_{2,0} = 0.3; 0.7; 1.3 \text{ y } 1.9$ . Para obtener el rendimiento del filtro compuesto diseñado, primeramente consideramos toda la información espectral del proceso de reconocimiento óptico de formas con imágenes desenfocadas. Posteriormente, consideramos los resultados obtenidos cuando se aplica el banco de filtros compuestos a las escenas de entrada. En este caso, se mide el rendimiento del filtro para cada descomposición. Finalmente, consideramos un análisis del rendimiento del filtro cuando se aplica una posterior fusión de la información descompuesta.

En primer lugar se ha diseñado un filtro compuesto para detectar las imágenes correspondientes a la persona de la derecha en la Fig. 2a y distinguirla de la persona de la izquierda de esta figura. En este filtro se ha tenido en cuenta toda la información espectral de las imágenes. En la Fig. 3a, se muestra la capacidad de discriminación (DC) en función del parámetro de desenfoque  $W_{2,0}$  para el caso en el que se considera toda la información espectral del plano de correlación. Se representan las dos definiciones bDC y wDC (mejor y peor DC) dadas por las Ecs. 2. De esta figura es posible observar, que este filtro realiza una buena discriminación entre las dos clases hasta un grado de desenfoque 0.7 (bDC (●), wDC (■)). Para un grado de desenfoque mayor que 0.7 la capacidad de discriminación disminuye notablemente, e incluso se pueden obtener valores del pico de correlación mayores para las caras que se quieren rechazar. En las Figs 3b, 3c y 3d, se muestran la intensidad de la correlación para imágenes desenfocadas con parámetro de  $W_{2,0} = 0.3, 0.7 \text{ y } 1.3$  respectivamente. En estas figuras podemos observar buenos picos de correlación situados en los objetos a reconocer hasta valores de  $W_{2,0} = 0.7$ .

A continuación se diseñan 3 filtros compuestos para detectar las mismas caras que en el caso anterior, en los que sólo se tienen en cuenta bandas de frecuencia, tal y como se explicó en el apartado 2.2. Consideramos tres regiones de soporte de acuerdo a las tres bandas seleccionadas F1

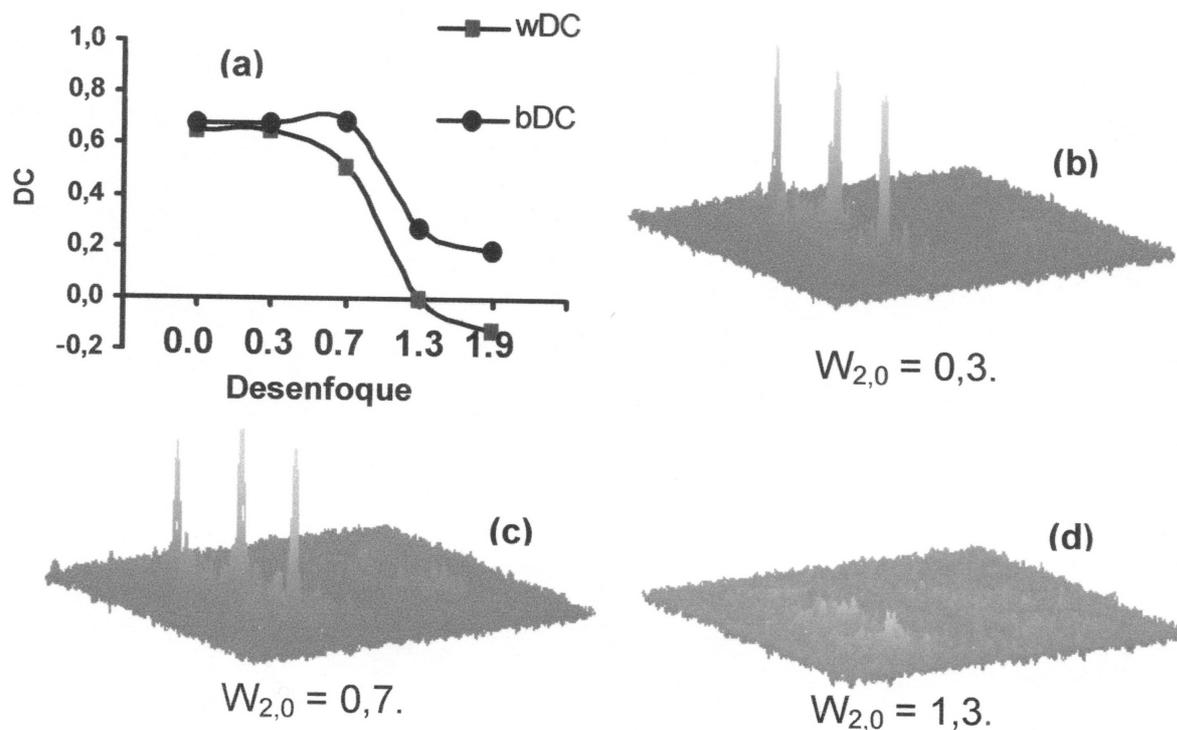


FIGURA 3. Información obtenida de los planos de intensidad de correlación experimentales. El filtro utilizado es un TOF que usa todo el ancho de banda. (a) Valores para wDC y bDC. (b), (c) y (d) intensidad de la correlación experimental para imágenes desenfocadas con  $W_{2,0} = 0,3, 0,7$  y  $1,3$  respectivamente.

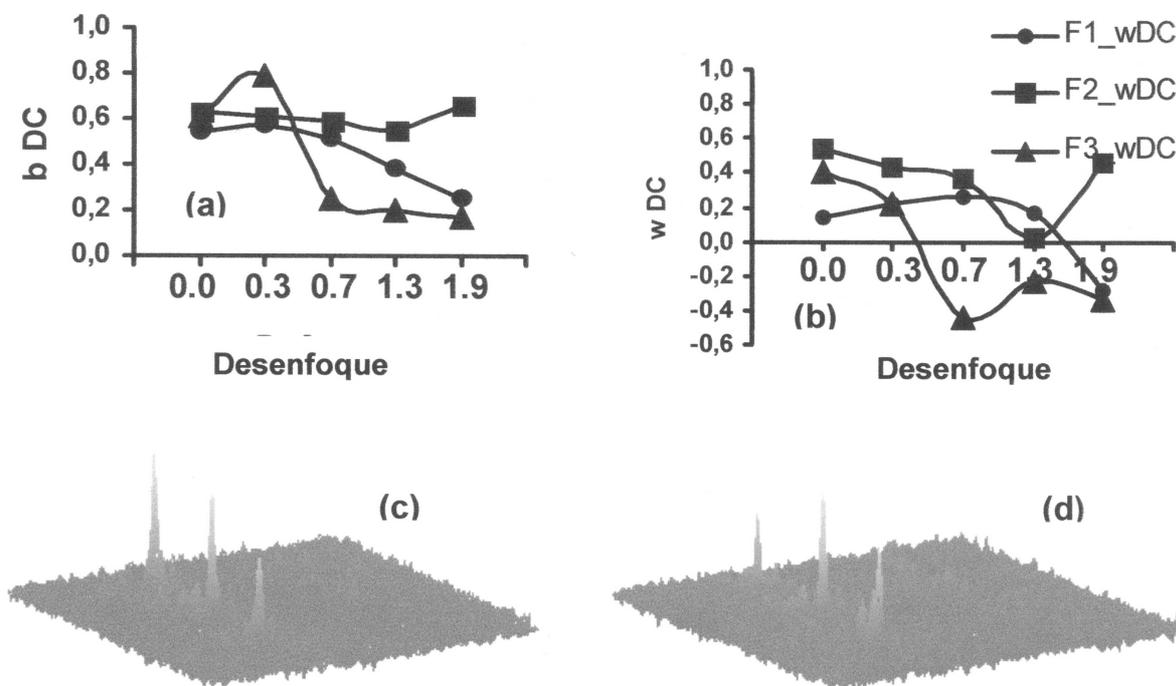


FIGURA 4. Valores para la DC obtenidas con filtros diseñados con regiones de soporte pasabandas versus el grado de desenfoco. (a) y (b) corresponden a la mejor DC y la peor DC respectivamente definidas por las Ecs. 2. (c) y (d) intensidad de la correlación experimentales para la banda F2 con imágenes de entrada con desenfoco de  $W_{2,0} = 0,3$  y  $0,7$  respectivamente.

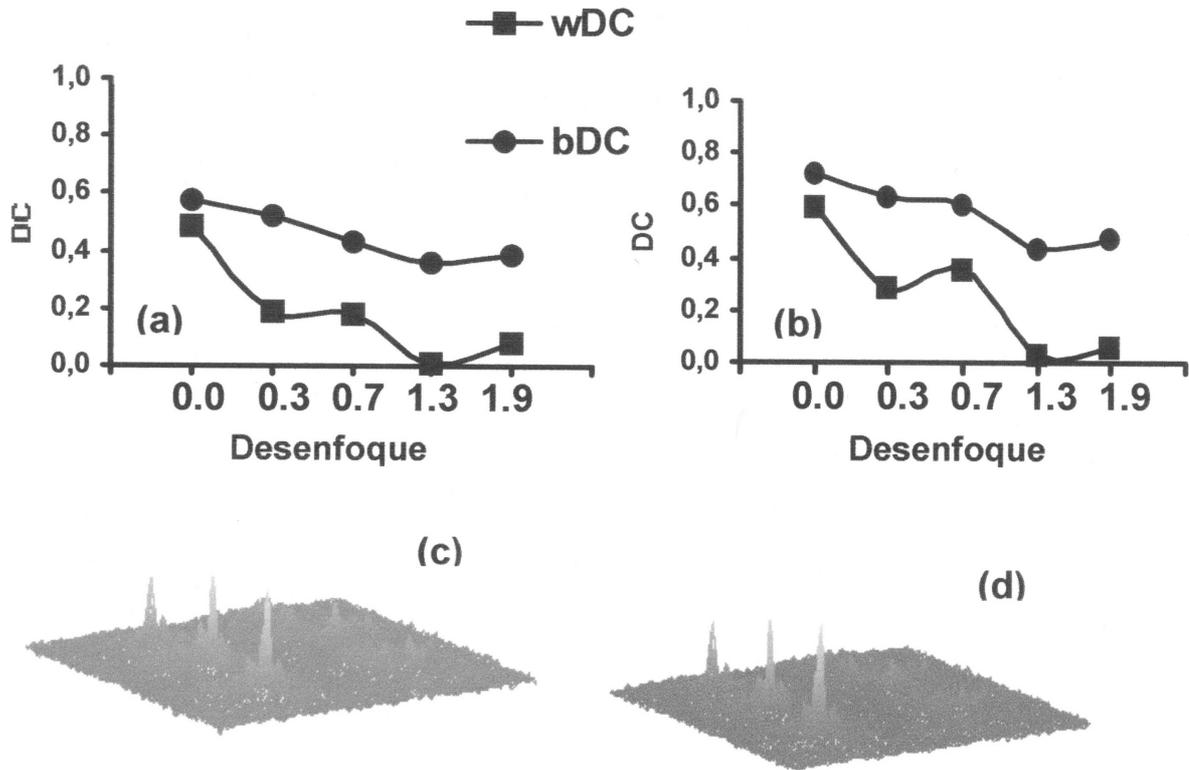


FIGURA 5. Resultados obtenidos mediante la fusión de los canales (a) y (c) mediante la media aritmética, (b) y (d) mediante la media geométrica. (a) y (b) capacidades de discriminación versus el parámetro de desenfoco. (c) y (d) intensidad de las correlaciones obtenidas con un desenfoco  $W_{2,0}=0.7$ .

( $f_1 = 37$ ,  $\Delta f = 26$  *pixeles*), F2 ( $f_2 = 67$ ,  $\Delta f = 34$  *pixeles*) y F3 ( $f_3 = 106$ ,  $\Delta f = 44$  *pixeles*) explicadas en la Sec. 2.2. La región de la banda F1 es la más cercana a la frecuencia cero del plano de Fourier. Con cada uno de estos filtros se realiza una correlación; las capacidades de discriminación obtenidas experimentalmente para cada uno de los filtros se muestran en las Fig. 4a y 4b. Observamos para el filtro compuesto F2 (■), una invariancia al grado de desenfoco de la escena de entrada, para todos los desenfocos estudiados. Para la primera banda F1 (●), la DC tiene una tendencia a disminuir pero no a cero. Para la banda F3 (▲) correspondiente a las altas frecuencias, se obtiene un valor máximo de la DC pero luego comienzan a decrecer debido principalmente a que la señal captada por la cámara CCD es débil y el ruido aumenta. En las Figs. 4c y 4d, se muestran planos de intensidad de correlación obtenidos con la región de soporte correspondiente a la banda F2 para valores del parámetro de desenfoco  $W_{2,0} = 0.3$  y  $W_{2,0} = 0.7$  respectivamente. Cuando se incrementa el grado de desenfoco la señal decrece y el ruido se incrementa.

Cada canal tiene diferente información de las imágenes de entrada, dependiendo de la frecuencia central y de su ancho de banda o región de soporte. Si toda la información se reúne mediante una fusión de los diferentes canales en un único plano de correlación, es posible mejorar los resultados del reconocimiento. En este trabajo, presentamos los resultados correspondientes a dos operaciones de fusión: la suma

o la media geométrica de los canales individuales obtenidos previamente.

En la Fig. 5 se muestran los parámetros bDC (●), wDC (■), para la suma (Fig. 5a) y para la media geométrica (Fig. 5b) de los canales versus el grado de desenfoco. Observamos que bDC tiene una tendencia a mantenerse constante. Se obtienen valores más altos para la fusión mediante la media geométrica. También observamos en la wDC una tendencia similar a la bDC pero hay más inestabilidades en el caso de la media geométrica. En las Figs. 5c y 5d, presentamos planos de intensidad de correlación para las fusiones consideradas. Hemos considerado una imagen de entrada con un parámetro de desenfoco de  $W_{2,0} = 0.7$ . Los gráficos de la izquierda y derecha representan las fusiones suma y media geométrica respectivamente. Se pueden observar picos de autocorrelación más altos que los de correlación cruzada en el caso de la media geométrica (Fig. 5d).

#### 4. Conclusión

Hemos aplicado el correlador óptico convergente para resolver un proceso confiable de reconocimiento de formas con imágenes a analizar desenfocadas. En el diseño de filtros compuestos hemos considerado filtros de compromiso TOF, escogiendo un adecuado balance entre los criterios SNR y PCE. Una DC invariante a desenfoco se obtiene sólo si se utiliza la banda de frecuencias intermedia. Sin embargo, en

las otras bandas estos resultados no son lo suficientemente buenos y no compensan las complicaciones adicionales que tiene la descomposición en bandas, de modo que el uso del filtro completo continúa siendo una buena alternativa. Al considerar la fusión de la información (tratamiento multicanal) los resultados presentan un mejor rendimiento. A medida que el desenfoque aumenta la DC tiene una tendencia a mantenerse constante. En este caso la fusión mediante la media geométrica de los canales presenta una bDC invariante a los desenfoques considerados.

## Agradecimientos

Este trabajo ha sido financiado por el Fondo Nacional de Desarrollo Científico y Tecnológico Fondecyt, Proyectos N°1010532 y N°1040946; Dirección General de Enseñanza Superior del Ministerio de Educación y Cultura Proyecto N°BFM2003-06273-C02-01; y Proyecto de Investigación Conjunta dentro del Programa de Cooperación Científica con Iberoamérica DURSI-CONICYT ACI2003-51.

- 
1. Y.N. Hsu y H.H. Arsenault, *Appl. Opt.* **21** (1982) 4016.
  2. S. Roy, Henry H. Arsenault y Daniel Lefebvre, *Opt. Eng.* **42** (2003) 813.
  3. Xiang Yang Yuan, Chun Kan Tao, Da Yong Yu y Ming Shen Huo, *Opt. Eng.* **41** (2002) 2923.
  4. A. Rahman Alsamman y M.S. Alam., *Opt. Eng.* **41** (2002) 87.
  5. J. Campos, S. Bosch y J. Sallent, *Opt. Commun.* **82** (1991) 370.
  6. A. Vargas, J. Campos y R. Navarro, *Opt. Comm.* **185** (2000) 33.
  7. A. Vargas, J. Campos, C. San Martín y N. Vera, *Optics and Lasers in Engineering* **40** (2003) 67.
  8. R. Navarro, A. Vargas y J. Campos, *J. Opt. A Pure Appl. Opt.* **1** (1999) 116.
  9. A. Attaleb, E. Ahouzi, J. Campos, I. Moreno y M.J. Yzuel, *Opt. Eng.* **41** (2002) 3250.
  10. J.J. Esteve-Taboada, J. García y C. Ferreira, *Opt. Eng.* **41** (2002) 1324.
  11. B.V. Vijaya Kumar y L. Hassebrook, *Appl. Opt.* **29** (1990) 2997.
  12. B.V.K. Vijaya Kumar, *Appl. Opt.* **31** (1992) 4774.
  13. Ph. Refrieger, *Opt. Lett.* **16** (1991) 829.
  14. V.N. Mahajan, *Aberration Theory Made Simple*, Vol TT6. (Spie Optical Engineering Press, USA, 1991).
  15. A. Vargas, C. San Martín, R. Figueroa, J. Campos y M. J. Yzuel, *SPIE Proc.* **4419** (2001) 608.