

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN
SUPERIOR DE ENSENADA



PROGRAMA DE POSGRADO EN CIENCIAS
EN CIENCIAS DE LA COMPUTACIÓN

**Métodos y algoritmos para reconocer objetos parcialmente
ocultos con pequeñas distorsiones geométricas**

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el grado de

DOCTOR EN CIENCIAS

Presenta:

José Ángel González Fraga

Ensenada, Baja California, México, Enero de 2007.

*Al amor de mi vida:
Norma Patricia
y a mis pequeñas princesas:
Valeria y Marianita*

Agradecimientos

Quiero expresar mi agradecimiento:

A mi amada Patty por su comprensión, dedicación, amor incondicional y porque sin el apoyo de ella no hubiera sido posible cumplir esta meta.

A todos mis compañeros, por el apoyo que me proporcionaron: Victor, Everardo, Erika, los profes de Colima, los profes de Ciencias de la UABC, Carlangas, Christian, Julio, Omarcito, Paul, los Ceni2t,

A los profesores de CICESE, por su dedicación en su tarea de la enseñanza: Favela, Hugo, Brizuela, Gilberto, Vitaly, Briseño, Josúe,

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada para el desarrollo de este proyecto.

Ensenada, México
22 de Enero de 2007.

José Ángel González Fraga

Tabla de Contenido

Capítulo	Página
Resumen	II
Abstract	III
Agradecimientos	V
Lista de Figuras	VIII
Lista de Tablas	XII
Tabla de símbolos	XIII
I. Introducción y objetivos	1
I.1. Definición del problema y objetivos	2
I.1.1. Objetivo general	3
I.1.2. Objetivos específicos	3
I.2. La correlación como alternativa para el reconocimiento de patrones .	4
I.3. Antecedentes	7
I.4. Organización de la tesis	11
II. Sistemas lineales y teoría de filtrado	13
II.1. Sistemas lineales e invariantes a desplazamientos	14
II.2. La transformada de Fourier	17
II.2.1. Propiedades de la transformada de Fourier	19
II.3. Análisis de correlación	23
II.4. Correlación entre funciones discretas	25
II.4.1. Teorema del muestreo	26
II.4.2. Transformada discreta de Fourier	27
II.5. Resumen	28
III. Detección y estimación de parámetros en imágenes	29
III.1. Prueba de hipótesis binaria para la detección de señales	30
III.1.1. Filtro espacial de acoplamiento	30
III.2. Estimación de parámetros en imágenes.	34
III.2.1. Modelo aditivo de ruido Gaussiano blanco (AWGN)	35
III.2.2. Estimación de la posición de la señal en el modelo AWGN . .	37
III.3. Resumen	39
IV. Filtros clásicos de correlación	40
IV.1. Criterios para derivar filtros de correlación	41
IV.1.1. Optimización de la relación señal-ruido: filtro de acoplamiento clásico	41
IV.1.2. Optimización de la eficiencia de luz: filtro sólo de fase	42
IV.1.3. Optimización de la estrechura del pico: filtro inverso	44
IV.1.4. Capacidad de discriminación	46
IV.2. Algoritmo para reconocer objetos por medio de la correlación	47

Tabla de Contenido (Continuación)

Capítulo	Página
IV.3. Las invariancias en el reconocimiento de patrones	48
IV.3.1. Invariancia a escala	50
IV.3.2. Invariancia a la rotación	51
IV.4. Derivación de filtros compuestos	52
IV.4.1. Funciones discriminantes sintéticas	52
IV.4.2. SDF de mínima varianza (MVSDF)	56
IV.4.3. Filtro MACE (filtro del mínimo promedio de la energía de correlación)	57
IV.5. Resumen	62
V. Filtros adaptativos para el reconocimiento de patrones	64
V.1. Algoritmo básico para el diseño de un filtro adaptativo ASDF	67
V.2. Resultados experimentales	71
V.2.1. Experimentos con fondos espacialmente homogéneos	81
V.2.2. Incorporación de invariancias al filtro ASDF	86
V.3. Mejora al algoritmo básico del filtro ASDF	94
V.4. Resumen	95
VI. Filtros adaptativos para reconocer objetos parcialmente ocultos	96
VI.1. Filtros adaptativos para reconocer objetos fragmentados y parcial- mente ocultos	98
VI.1.1. Reconocimiento de objetos fragmentados	101
VI.1.2. Reconocimiento de objetos parcialmente ocultos	107
VI.2. Resumen	114
VII. Aplicaciones	115
VII.1.El problema de detección	116
VII.2.El problema de clasificación	124
VII.3.Reconocimiento de fragmentos de diatomeas fósiles	126
VII.4.Resumen	127
VIII.Conclusiones y áreas de oportunidad	130
Bibliografía	133
A. Ruido aditivo	138
A.1. Generación de ruido con distribución Gausiana	139
A.1.1. Ruido correlacionado	139

Lista de Figuras

Figura	Página
1. Esquema de la correlación entre imágenes: imagen de referencia, imagen de prueba y la salida ideal de la correlación.	6
2. Correlación entre dos funciones: $c(x)$ es la correlación entre $f(x)$ y $h(x)$	24
3. Diagrama de bloques de una identificación binaria.	31
4. Filtrado por acoplamiento para la localización de objetos.	39
5. Cálculo de la DC: (a) ejemplo de escena de entrada, (b) salida de correlación. Para calcular la DC según la ecuación (72), C^T es el máximo dentro del área del objeto y C^B es el máximo dentro del área de picos falsos.	46
6. Ejemplo de localización en escenas libres de ruido. (a) Escena de entrada, el objetivo es la mariposa inferior. Respuestas al impulso de los filtros: (b) CMF , (c) POF y (d) IF . Planos de correlación en 3-D proporcionados por: (e) CMF , (f) POF y (g) IF	49
7. Una imagen digital se puede ver como un arreglo bidimensional de M renglones y N columnas, donde cada elemento se le llama pixel. Al ordenar lexicográficamente los renglones se obtiene un vector imagen (columna) con $M \times N$ elementos.	53
8. (a) Imagen de prueba y (b) su región de soporte.	68
9. Diagrama de bloques del algoritmo iterativo para diseñar un filtro $ASDF$	70
10. Imágenes de prueba: (a) Objeto de referencia y (b) fondo de la realidad.	72
11. Respuestas al impulso correspondientes a los filtros $ASDF$ en la salida de la iteración número: (a) uno, (b) dos y (c) veintidós.	73
12. Imágenes de rechazo generadas a partir del fondo y con la misma región de soporte que la del objeto de referencia. El fondo negro de las imágenes de rechazo originales se cambió por blanco para una mejor apreciación.	74
13. Imágenes de prueba: (a) escena formada con el fondo y el objeto de referencia. (b-d) Salidas de correlación obtenidas con los filtros de la Fig. 11 aplicados a la escena (a). El área del rectángulo indica el centro donde se creará una imagen de rechazo a partir del fondo (máxima correlación en el área del fondo) y el área del círculo indica la posición del objeto de referencia en la escena de prueba.	75
14. Desempeño del filtro $ASDF$ durante el proceso de diseño, empleando las imágenes de prueba de la Fig. 10.	76

Lista de Figuras (Continuación)

Figura	Página
15. Reconocimiento de objetos en escenas de prueba con una imagen de fondo de la realidad. (a) Escena de prueba que contiene sólo al objeto de referencia marcado con la flecha, (b) escena de prueba que contiene al objeto de referencia y a cuatro objetos no deseados, (c) posición de los objetos en la escena (b), el objeto de referencia está marcado con la flecha, (d) distribución de la intensidad de correlación correspondiente a la escena de prueba (b) y la cual se obtuvo con el filtro <i>ASDF</i>	80
16. Reconocimiento de objetos en escenas de prueba con fondos generados como procesos estocásticos. (a) Escena de prueba que contiene al objeto de referencia y a cuatro objetos no deseados incrustados en un fondo tipo ruido blanco con desviación estándar de 40 y valor medio de 128, (b) Escena de prueba que contiene al objeto de referencia y a cuatro objetos no deseados incrustados en un fondo tipo ruido correlacionado con desviación estándar de 40, valor medio de 128 y coeficiente de correlación de 0.9. La posición de los objetos es similar a la de la Fig. 15(c).	82
17. Desempeño del filtro <i>ASDF</i> durante el proceso de diseño para el fondo tipo ruido blanco con diferentes desviaciones estándar: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 30$, (d) $\sigma = 40$	83
18. Desempeño del filtro <i>ASDF</i> durante el proceso de diseño para el fondo tipo ruido correlacionado con diferentes desviaciones estándar: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 30$, (d) $\sigma = 40$	84
19. Desempeño de los filtros de correlación cuando la escena de entrada es la mostrada en la Fig. 16(a). El desempeño se midió en términos de la <i>DC</i> contra la desviación estándar del fondo tipo ruido blanco.	85
20. Diferentes versiones del objeto a reconocer. El número debajo de cada objeto indica el grado de rotación que se le aplicó al objeto original.	87
21. Respuestas impulsivas correspondientes a los filtros: (a) <i>ASDF</i> (generado con los objetos 0 y 8 de la Fig. 20) y (b) <i>SDF</i> convencional (generado con los objetos 0, 2, 4, 6, 8, 10, y 12 de la Fig. 20).	88
22. Diferentes versiones del objeto a reconocer. El número debajo de cada objeto indica el factor de escalamiento que se le aplicó al objeto original.	90
23. Objetos modificados, los cuales se utilizaron para diseñar el filtro <i>ASDF</i> con invariancia a la escala.	91
24. Escenas de prueba alteradas por ruido aditivo de media cero y $\sigma = 40$: (a) escena mostrada en la Fig. 15(b) y (b) escena mostrada en la Fig. 16(b)	92

Lista de Figuras (Continuación)

Figura	Página
25. Tolerancia al ruido aditivo blanco de los filtros <i>OF</i> y <i>ASDF</i> . El desempeño de los filtros se probó con las escenas mostradas en la Fig. 24. El desempeño se midió en terminos de la <i>DC</i> contra la desviación estándar del ruido.	93
26. Diagrama de bloques del algoritmo iterativo para diseñar el filtro <i>ASDF</i> para reconocer objetos parcialmente ocultos.	100
27. Imágenes de prueba: (a) objeto de referencia, (b-g) fragmentos del objeto de referencia a ser detectados y (h) mariposa falsa.	102
28. Ampliación del objeto de referencia particionado en seis imágenes independientes.	103
29. Desempeño del filtro <i>ASDF</i> para reconocer objetos fragmentados durante la etapa de diseño.	103
30. Imagen de prueba para el problema de Clasificación: (a) la escena contiene al objeto fragmentado (b) que se muestra en la Fig. 27 y a la mariposa falsa. (b) Imagen tridimensional de la intensidad correlación resultante para la escena de prueba.	104
31. Tolerancia de diferentes filtros hacia el ruido aditivo en términos de falsas alarmas. Se usó la escena de la Fig. 30(a).	105
32. Objetos de prueba: (a) Objeto de referencia, (b-g) objeto de referencia parcialmente oculto por otros objetos similares.	109
33. Objeto de referencia, dividido en cuatro fragmentos (imágenes independientes).	110
34. Fondo utilizado en los experimentos	110
35. Desempeño del filtro Adaptativo SDF en el proceso de diseño.	111
36. Escena de prueba para detectar objetos parcialmente ocultos. (a) objetos de la Fig. 32.(b) incrustados en un fondo de la realidad, (b) plano de correlación en 3D, el cual se obtuvo con el filtro <i>AF</i>	112
37. Tolerancia del filtro <i>ASDF</i> frente a pequeñas distorsiones geométricas: rotación.	112
38. Tolerancia del filtro <i>ASDF</i> frente a pequeñas distorsiones geométricas: escalamiento.	113
39. Imágenes de prueba: (a) Copépodo macho como el objeto a reconocer. (b) Fondo a rechazar.	117
40. Objetos modificados: (a) El copépodo macho de la Fig. 39 como nuevo objeto a reconocer. (b) Objeto de rechazo del fondo, creado durante el proceso iterativo del diseño del filtro <i>ASDF</i>	118
41. Desempeño del filtro <i>ASDF</i> en la etapa de diseño.	119

Lista de Figuras (Continuación)

Figura	Página
42. Imágenes de prueba: (a) Escena con un copépodo macho como el objeto a reconocer. (b) Distribución de la salida de correlación correspondiente a la escena de prueba.	120
43. Imágenes de prueba: (a) T_2 , T_3 , T_4 y T_5 representan diferentes versiones del objeto T1 que es el mismo copépodo de la Fig. 39, ver la Tabla IX para más información. (b) La escena distorsionada con ruido aditivo de naturaleza Gaussiana con media cero y desviación estándar 15.	122
44. Imágenes de prueba: (a) Objetos a reconocer (copépodos macho). (b) Objetos a rechazar (copépodos hembra).	123
45. Imágenes de prueba: (a) Objetos modificados a reconocer (copépodos macho). (b) Objetos modificados a rechazar (copépodos hembra).	123
46. Escena de prueba, los copépodos macho están marcados con la letra M.	124
47. Distribuciones de intensidad de correlación: (a) salida de la escena de la Fig. 46. (b) salida después de aplicar un umbral del 50% del valor del pico más alto.	125
48. Diatomeas fósiles: (a) <i>Stephanodiscus sp</i> y (b) <i>Aspeitia nodulifer</i>	127
49. Imagen de prueba: (a) Escena de prueba, (b) posición de los objetos a reconocer, (c) salida de la correlación con el filtro <i>ASDF</i> y (d) gráfica 3D de la salida de correlación.	128
50. (a) Función de correlación de ruido con parámetros $\rho_V = 0.95$, $\rho_H = 0.95$ y (b) Función de densidad espectral de (a).	142
51. Imágenes afectadas con ruido aditivo de diferentes valores de correlación. El ruido es de media cero y desviación estándar de 20.	142

Lista de Tablas

Tabla	Página
I. Desempeño de varios filtros de correlación en términos de su <i>DC</i> para las escenas de prueba de la Fig. 15.	79
II. Desempeño de varios filtros de correlación en términos de su <i>DC</i> para las escenas de entrada de la Fig. 16.	81
III. Desempeño de filtros de correlación en términos de su <i>DC</i> para objetos rotados en la escena de entrada de la Fig. 15(a)	88
IV. Desempeño de varios filtros de correlación en términos de su <i>DC</i> para objetos escalados en la escena de entrada de la Fig. 15(a).	91
V. Desempeño de los filtros de correlación en términos de las falsas alarmas (clasificación errónea) para los problemas de Detección y Clasificación. Se usaron 50 pruebas para distintas posiciones de los objetos de la Fig. 27.	104
VI. Desempeño de los filtros adaptativos en términos de su <i>DC</i> para los problemas de Detección y Clasificación. Se usaron 50 pruebas estadísticas para distintas posiciones de los objetos en la Fig. 27.	105
VII. Desempeño de los filtros de correlación en términos de falsas alarmas (errores de detección) para las escenas de la Fig. 32	111
VIII. Desempeño de varios filtros de correlación en términos de su <i>DC</i> para las escenas de prueba de la Fig. 32.	113
IX. Desempeño de diferentes filtros para la escena de la Fig. 43(a).	121
X. Desempeño de diferentes filtros para la escena de la Fig. 43(a), alterada por ruido aditivo de media cero y desviación estándar 15.	122

Tabla de símbolos

\mathbf{A}	Matriz real
\mathbf{A}^T	Matriz transpuesta
$\hat{\mathbf{A}}$	Matriz Compleja, en el dominio de Fourier
$\hat{\mathbf{A}}^+$	Matriz transpuesta conjugada de $\hat{\mathbf{A}}$
\mathbf{a}	Vector
\mathbf{a}^T	Vector
$\hat{\mathbf{a}}$	Vector complejo, en el dominio de Fourier
a	Escalar
$s(x, y)$	Escena de entrada
$h(x, y)$	Respuesta al impulso
$t(x, y)$	Imagen de referencia (Objetivo)
$n(x, y)$	Función ruido
$S(u, v)$	Transformada de Fourier de la escena de entrada
$H(u, v)$	Respuesta en frecuencia (Filtro)
$T(u, v)$	Transformada de Fourier de la imagen de referencia
$A^*(u, v)$	Complejo conjugado de $A(u, v)$
$P_n(u, v)$	Densidad espectral de la energía del ruido
$ S(u, v) $	Magnitud de $S(u, v)$
$\phi_S(u, v)$	Fase de $S(u, v)$
x, y	Coordenadas en el dominio del espacio
u, v	Coordenadas en el dominio de la frecuencia
\mathcal{F}, FT	Transformada de Fourier
\mathcal{F}^{-1}, IFT	Transformada inversa de Fourier
FFT	Transformada rápida de Fourier
C^T	Máximo valor de correlación en el área del objeto de referencia
C^B	Máximo valor de correlación en el área del fondo
μ	Valor medio
$E\{\cdot\}$	Esperanza matemática
σ	Desviación estándar
η_H	Eficiencia de Horner (Luz)

Capítulo I

Introducción y objetivos

El principal objetivo del sistema visual humano es reconocer objetos de tal manera que se pueda interactuar con ellos apropiadamente. Aunque los humanos percibimos a los objetos de una forma instantánea y automática, nuestro sistema visual requiere de un proceso complejo para reconocer objetos del mundo real. Uno de los principales obstáculos que enfrenta un sistema visual, sea biológico o artificial, es el hecho de que los objetos oculten parte de ellos mismos o de sus vecinos, es decir, la información que describe a los objetos está frecuentemente incompleta. Es necesario hacer notar que un humano puede experimentar el mismo tipo de confusiones que un sistema por computadora para reconocer un objeto, si es que el contexto de la escena no se entiende claramente. En estos casos, aún el humano requiere entrenamiento para interpretar acertadamente las imágenes del tipo que se le muestran (Sekuler *et al.*, 2000).

La percepción de imágenes artificiales es muy importante en áreas como la robótica, visión e inteligencia artificial. Todas estas áreas tienen un gran interés comercial, por lo que son ampliamente investigadas. Por esta razón, existe un largo historial del reconocimiento de patrones que se puede mencionar desde las primeras décadas del siglo pasado. A pesar de contar con hardware y sistemas más completos y complejos, el reconocimiento de patrones por medios artificiales ha revelado ser muy complicado. La limitada capacidad del hardware que actualmente existe, restringe el tipo de problemas que se pueden solucionar satisfactoriamente. En este sentido, los métodos de reconocimiento de patrones generalmente se ocupan de un número limitado de imágenes en ambientes controlados, y por ende, aún se desconoce si un buen método aplicado a un

caso particular, sería eficaz para todos los casos de la vida (Suetens *et al.*, 1992).

Existen diferentes métodos para realizar la tarea de reconocimiento, se puede ver un resumen de estrategias en Suetens *et al.* (1992), algunas son simplemente heurísticas o herramientas como la programación dinámica, otras tienen bases matemáticas o estadísticas como los métodos basados en la teoría de decisión (clasificador de distancia mínima (Duda *et al.*, 2001), reconocimiento por correlación (Vanderlugt, 1992), los clasificadores óptimos estadísticos como el clasificador de Bayes (Van Trees, 1968), y las redes neuronales (Duda *et al.*, 2001).

I.1. Definición del problema y objetivos

En el ámbito del procesado de imágenes, el reconocimiento de objetos o patrones se define como la tarea de encontrar y etiquetar las partes de una imagen bidimensional, por ejemplo, los objetos que conforman una escena (Gonzalez y Woods, 2002). Un problema importante en esta área, aparece cuando el objeto a reconocer se encuentra parcialmente oculto por otros objetos. Si reconocer objetos es difícil, reconocer objetos parcialmente ocultos es más complicado, esto se debe a que se corrompe la información que lo describe, es decir, se pierde parte de la forma del objeto. En estos casos, lo sensato es aprovechar otras características del objeto como el color, la intensidad y la textura.

En base a lo anterior, este trabajo de investigación trata sobre el diseño de filtros adaptativos de correlación para el reconocimiento de objetos parcialmente ocultos. El problema a tratar en esta investigación se puede definir por los siguientes objetivos de investigación:

I.1.1. Objetivo general

Diseñar nuevos métodos y algoritmos por correlación con invariancias a pequeñas distorsiones geométricas, que sean capaces de identificar y localizar objetos parcialmente ocultos en presencia de ruido (aditivo y disjunto) y que tengan un buen desempeño en términos de su capacidad de discriminación y minimización de falsas alarmas.

Para lograr el objetivo general se han propuesto los siguientes objetivos específicos que forman parte de la metodología empleada en el desarrollo de la investigación:

I.1.2. Objetivos específicos

1. Estudiar los fundamentos teórico-matemáticos de las técnicas por correlación invariantes en posición, rotación y escalamiento.
2. Estudiar los filtros de correlación conocidos, aplicados a los modelos de señal aditivo y disjunto, así como los criterios de calidad clásicos que se emplean para determinar las capacidades de los diversos métodos de correlación.
3. Añadir las invariancias (rotación, escalamiento, etc.) a los diferentes filtros de correlación conocidos y aplicarlos en el reconocimiento de objetos parcialmente ocultos.
4. Proponer, diseñar e implementar nuevos métodos por correlación para identificar objetos parcialmente ocultos basándose en el estudio de los puntos anteriores.
5. Probar y comparar los resultados arrojados por los algoritmos propuestos, para reconocer objetos parcialmente ocultos en imágenes de prueba; basándose en su capacidad de discriminación y minimización de falsas alarmas.
6. Hacer un estudio de las propiedades y del comportamiento de los métodos por correlación que se propondrán para identificar las limitaciones del nivel de oclusión

que nos impiden identificar un objetivo.

7. Probar los algoritmos con imágenes reales de diversa temática (médicas, satelitales, de microscopía, etc.) para identificar objetos parcialmente ocultos.

I.2. La correlación como alternativa para el reconocimiento de patrones

El reconocimiento de patrones por medio de la correlación (CPR¹) es una subrama del reconocimiento estadístico de patrones. Está basada en la selección o diseño de una señal (que también puede ser una imagen) de referencia para después determinar el grado de similitud entre la señal de referencia y el objeto bajo examinación (escena de prueba). El grado de similitud entre las señales es una simple estadística cuyas decisiones se basan en el conocimiento que se tenga sobre el objeto. Por medio de la correlación se podría determinar a que clase pertenece un objeto, o mejor aún, podríamos tener un sistema basado en la correlación que incorpore invariancias para identificar al objeto sin importar desde donde se observe (las variaciones que pudiera presentar un objeto se puede deber a algunos de los siguientes aspectos: desplazamientos laterales, rotaciones, escalamientos, cambios de iluminación y oclusiones). Frecuentemente es deseable discriminar entre diferentes clases de objetos que difieren sutilmente de la clase de interés, es decir, entre objetos parecidos. Finalmente, el objeto se puede encontrar incrustado (o rodeado) por ruido, cuyas características estadísticas pueden ser similares a la clase de interés. Estas consideraciones se pueden tener en cuenta para el diseño de las señales de referencia. Los algoritmos de correlación diseñan o crean señales de referencia tales que su correlación con el objeto produce estadísticas con la información de estas cuestiones (ruido o degradaciones en la señal observada).

¹CPR: Correlation pattern recognition

Una de las fortalezas principales del CPR es la robustez inherente que resulta de evaluar la señal completa en un mismo tiempo, en otras palabras, no se requiere de ningún tipo de preprocesamiento; en contraste con otras técnicas que extraen información en partes para luego comparar las relaciones entre los elementos. El CPR es menos sensitivo a pequeños valores de correspondencia y obstrucciones, debido a que se compara toda la imagen de entrada contra la plantilla (imagen de referencia). Además de que se pueden detectar varios objetos en paralelo dentro de la escena de entrada.

La correlación se puede implementar tanto en el dominio temporal (dominio espacial para las imágenes) como en el dominio de la frecuencia. Como ya se sabe, la propagación de la luz en su aproximación paraxial, se puede modelar como un sistema invariante y aunado a la propiedad intrínseca de las lentes para obtener las FT² en 2 dimensiones, se tiene que este método permite implementar la correlación en tiempo real. Este método para implementar la correlación en el dominio de las frecuencias es la razón por la que se usa la frase de "filtros de correlación". Gracias al recurso del algoritmo de la transformada rápida de Fourier (FFT³) y a los procesadores digitales de altas velocidades, hoy en día la correlación entre imágenes se puede usar efectivamente en implementaciones digitales. En este trabajo de investigación se omiten los sistemas basados en la óptica y se trabaja únicamente con sistemas digitales. El diseño de filtros de correlación y sus aplicaciones se programaron en MATLAB[®] versión 7.0, ya que este software tiene incorporadas diversas rutinas útiles para el tratamiento digital de imágenes, entre las que se destacan las funciones de la FFT.

Para presentar inicialmente el uso del CPR utilizaremos la Fig. 1. En esta figura, tenemos dos imágenes: una imagen que corresponde al patrón que deseamos buscar y una imagen de prueba que contiene varios patrones. En este ejemplo buscamos la

²FT: Fourier Transform

³FFT: Fast Fourier Transform

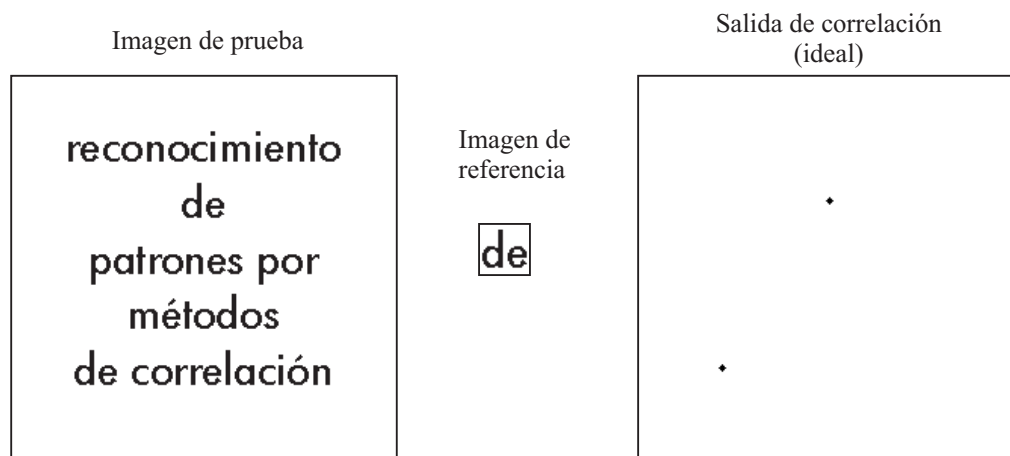


Figura 1: Esquema de la correlación entre imágenes: imagen de referencia, imagen de prueba y la salida ideal de la correlación.

palabra "de" dentro del texto. La salida de la correlación entre las imágenes, debería proporcionar valores grandes en las dos posiciones donde la imagen de referencia sea igual en la imagen de prueba. Después, se deben localizar los objetivos de interés al examinar los picos de correlación correspondientes en la salida, se debe determinar si los picos de correlación tienen un valor suficientemente grande para indicar la presencia del objeto de referencia. En la figura que se analiza, se presenta una salida ideal de la correlación, pero es importante señalar que la palabra "de" se asemeja mucho a la sílaba "do" en "métodos" y podría indicar una falsa presencia de la imagen de referencia. Un objetivo en el diseño de filtros de correlación, es la construcción de imágenes de referencia que nos permitan distinguir al objeto de interés de entre objetos similares que corresponden a otra clase de objetos. Por último, la correlación se puede modelar como un sistema lineal invariante a los desplazamientos (las propiedades de los sistemas lineales se estudiarán en el capítulo II), lo que significa que cualquier desplazamiento en

la imagen de referencia afectará a la salida por la misma cantidad en el desplazamiento. De aquí que, si la imagen de referencia se encuentra centrada en el origen, la posición de los picos en la salida de la correlación indicarán exactamente la posición de los objetos (como en la Fig. 1).

Con la discusión anterior, se justifica el uso de los sistemas basados en la correlación para atacar el problema del reconocimiento de objetos parcialmente ocultos. Implementar la correlación es relativamente fácil, lo complicado es diseñar la imagen de referencia (a la que en ocasiones nos referiremos como filtro) para obtener un reconocimiento confiable de objetos.

I.3. Antecedentes

Aunque son pocos los trabajos que tratan el problema de reconocer objetos parcialmente ocultos, se revisaron algunos donde se emplean diversas técnicas que utilizan la información de los contornos y de las regiones de máxima curvatura para llevar a cabo una clasificación. Se pueden mencionar los siguientes estudios como de mayor relevancia:

Plantillas deformables para el reconocimiento de múltiples objetos ocultos

Mardia *et al.* (1997) construyeron un sistema Bayesiano de reconocimiento al que incorporaron una plantilla deformable. Introdujeron densidades de probabilidad a un modelo estocástico, lo cual permite encontrar versiones deformadas de objetos. Los resultados experimentales, se concentraron en casos donde sólo las fronteras de los objetos son de interés. De este modo, las estructuras de los vértices vecinos fueron fáciles de manipular aun sufriendo alguna deformación. Según los autores, el trabajo se puede extender a casos donde los objetos son más complicados, por ejemplo objetos no compactos y objetos sólidos en lugar de contornos.

Dentro de las desventajas que se pueden apreciar en este trabajo es que no consideraron el reconocimiento en presencia de ruido, trabajaron con imágenes en escala de grises, y sólo consideraron objetos con formas sencillas.

Sistema neuronal para el reconocimiento de objetos parcialmente ocultos en escenas desordenadas

Wiskott y Malsburg (1993), trabajaron en un sistema para la interpretación de escenas compuestas por varios objetos conocidos y con oclusiones mutuas. Se analizaron las escenas para reconocer los objetos presentes y para determinar las relaciones entre su oclusión. Los objetos se representan internamente por un modelo de grafos. Estos se construyen de una manera semiautomática al presentar los objetos contra diferentes fondos, los objetos se reconocen por medio de una liga de acoplamientos dinámicos.

Los experimentos mostraron que su sistema es muy exitoso al analizar escenas confusas (con muchos elementos difíciles de distinguir). Una de las fortalezas de este sistema es que se puede ampliar para nuevos objetos sin necesidad de programar, no necesita ejecutar rutinas de estimación estadísticas como lo requieren muchos sistemas de redes neuronales.

Los autores sólo consideran la información referente a la forma de los objetos a ser reconocidos y la técnica se aplica a imágenes con niveles de gris. A estas debilidades del sistema se aúna el requerimiento de un preprocesamiento (generar la red de grafos para representar a los objetos) y las limitaciones que pudiera tener la arquitectura de la red neuronal.

Reconocimiento de objetos parcialmente ocultos por técnicas de correlación

Campos *et al.* (1994), propusieron varios métodos con filtros de correlación para el reconocimiento de patrones, los cuales se aplicaron para diferenciar objetos ocultos.

En su estudio consideraron los siguientes filtros: el filtro sólo de fase (POF⁴), el filtro inverso (IF⁵) y el filtro de balance óptimo SDF (TOSDF⁶)

Al comparar la capacidad de los tres filtros para el reconocimiento de objetos parcialmente ocultos, los resultados mostraron que al utilizar un filtro modificado *IF*, el pico de correlación siguió siendo lo suficientemente bueno para reconocer un objeto hasta con un 80 % de oclusión. El artículo sólo presenta un análisis preliminar, y como trabajo futuro resalta el análisis de los efectos del ruido, el fondo de la escena, el tamaño de los objetos, y mejorar la capacidad de discriminación.

Algoritmos para correlacionar imágenes encubiertas por ruido con media cero

Khoury *et al.* (1998a,b, 2002a,b), desarrollaron varios algoritmos para optimizar filtros de correlación para reconocer imágenes encubiertas que han sido afectadas por ruido no superpuesto (textura como fondo y con iluminación constante). El encubrimiento es generado al ocultar una parte del objeto y cubriéndolo con un fondo constante no superpuesto (ruido). El estudio lo extendieron al incluir ruido no superpuesto de media cero, solamente trabajan con imágenes binarias y la forma de la apertura del objeto que oculta es conocida.

Reconocimiento de objetos parcialmente ocultos

En González-Fraga (2004) se realizó un estudio de filtros de correlación aplicados al problema de reconocer objetos parcialmente ocultos. Bajo la premisa de que todos los filtros de correlación emplean el contorno de los objetos para llevar a cabo el reconocimiento, se detectaron situaciones donde el contorno casi había desaparecido. En estos

⁴POF: Phase Only Filter

⁵IF: Inverse Filter

⁶TOSDF: Trade-off SDF Filter

casos los filtros clásicos sin entrenamiento resultaron una elección poco atractiva para tratar el problema.

Como resultado del estudio e investigación de filtros de correlación, se propusieron tres métodos para el reconocimiento de objetos parcialmente ocultos utilizando filtros con entrenamiento. Estos métodos mejoran la capacidad de discriminación de los filtros convencionales. En el primer método se aprovecha la característica multiclase de los filtros con entrenamiento, consiste en entrenar el filtro para que reconozca al objeto de referencia y rechace a los objetos que no son de interés. En la segunda propuesta se da un tratamiento por separado al contorno y a la textura del objeto de referencia para aumentar los posibles casos en que se puede reconocer un objeto parcialmente oculto. La última propuesta está orientada a tratar imágenes a color, ésta considera un preprocesamiento para descorrelacionar los canales del color de las imágenes reales. Una vez que se aplica el preprocesamiento se lleva a cabo una correlación independiente por cada canal transformado y de esta forma se consigue aumentar la capacidad de discriminación.

Es necesario mencionar que solo se emplearon filtros invariantes en posición y que aún existe mucho campo abierto para realizar investigación en este tema.

Sistemas basados en la correlación para reconocer objetos biológicos

Villalobos-Flores *et al.* (2001) emplearon un método de correlación invariante para la identificación de 12 especies de diatomeas fósiles. El método hace uso de la transformada de escala y de un filtro sólo de fase. Los resultados experimentales reportaron que el método es capaz de identificar las especies a partir de un pequeño fragmento de los organismos.

Castro-Longoria *et al.* (2001) propusieron un algoritmo invariante, basado en la correlación, para identificar especies de copépodos (pequeños crustáceos marinos que

forman parte del plancton). Se destaca el uso del filtro sólo de fase y la transformada de escala para añadir las invariancias al algoritmo propuesto. En este último trabajo así como en el de Pech-Pacheco y Álvarez-Borrego (1998), se presenta el reconocimiento automático de microorganismos, los cuales se encuentran, en algunos casos, ocultando parte de sí mismos. Cabe señalar que en estos trabajos se trabajó con imágenes limpias (libres de ruido) y se pierde la invariancia al desplazamiento debido a la implementación de la transformada de escala.

I.4. Organización de la tesis

El resto de este documento se ha organizado en la siguiente manera. El capítulo II, trata sobre los sistemas lineales y la teoría de filtrado. Estos son conceptos útiles para el planteamiento y diseño de sistemas basados en la correlación.

En el capítulo III, se introduce a la teoría de la detección y estimación de parámetros en las imágenes. Estos conceptos sirven para demostrar que la correlación es un sistema estadístico óptimo para el reconocimiento y detección de señales en general.

En el capítulo IV se describen los filtros clásicos de correlación, se derivan las familias de los filtros espaciales de correlación y la de los filtros compuestos que se aplican en la detección de objetos en imágenes. También se introducen los conceptos de invariancias para poder clasificar familias de objetos.

El capítulo V expone el diseño de nuevos filtros adaptativos para el reconocimiento de objetos en escenas complicadas y se discuten los resultados obtenidos.

En el capítulo VI, se introduce una mejora en el diseño de los nuevos filtros adaptativos de correlación propuestos. Estos filtros se optimizan para el reconocimiento de objetos parcialmente ocultos y objetos fragmentados. Se exponen algunos ejemplos y se discuten los resultados de los experimentos. Los filtros que se proponen en los capítulos

V y VI representan la contribución de este trabajo para mejorar el reconocimiento de objetos en escenas complicadas y ruidosas.

Posteriormente, se presenta un capítulo de aplicaciones, en donde se utilizan los filtros propuestos para el reconocimiento confiable de objetos biológicos.

En el capítulo VIII se hace una síntesis de las actividades realizadas y se exponen las conclusiones finales y finalmente, en el apéndice A, se presenta un algoritmo para generar ruido con distribución Gausiana.

Capítulo II

Sistemas lineales y teoría de filtrado

Un sistema, de manera informal, se puede caracterizar por una señal de salida en respuesta a una señal de entrada. En un proceso de correlación se involucran dos señales o imágenes. Una imagen de referencia es correlacionada con una imagen de prueba (escena) para detectar y localizar una imagen de referencia (objetivo) dentro de la escena. De esta manera, un sistema de correlación es aquel con una entrada (la escena), una plantilla almacenada o filtro (creado a partir de la imagen de referencia), y una salida (la correlación). Un sistema es lineal si para entradas nuevas, representadas por una suma ponderada de la entradas originales, proporciona una salida que es una suma ponderada de la salida original. De este modo, un sistema basado en la correlación presenta las ventajas de los sistemas lineales e invariantes en el tiempo y por lo tanto se puede describir en términos de su respuesta en frecuencia. Esta y otras propiedades se utilizarán para sintetizar y aplicar filtros de correlación para el reconocimiento de objetos.

En este capítulo se repasan tanto las propiedades de los sistemas lineales como las propiedades de la transformada de Fourier (FT). La FT es una herramienta muy útil en el filtrado de señales en general. El uso de la transformada de Fourier en el análisis de imágenes es muy amplio, por ejemplo, se emplea en detección de bordes al localizar altas frecuencias (cambios bruscos en la escala de grises) en la función imagen, también tiene aplicaciones en la restauración de imágenes alteradas, correlación rápida utilizando el teorema de convolución, caracterización de bordes, compresión de imágenes y en muchas otras aplicaciones.

II.1. Sistemas lineales e invariantes a desplazamientos

La teoría de los sistemas lineales tiene gran importancia en el procesamiento digital de imágenes (Bovik, 2000). Muchas de las técnicas más poderosas para modificar, mejorar, o representar información visual digital se expresan en términos de los conceptos de sistemas lineales. En el procesamiento de imágenes, los filtros lineales se aplican a la imagen en general, para tareas como el mejoramiento de contraste, eliminación de ruido, realce de contornos y también para tareas más específicas sobre los objetos que componen una imagen, como el reconocimiento de patrones.

En este trabajo se usa la variable espacial x , (en contraparte de la variable temporal t que es comunmente usada) para enfatizar nuestro interés en el manejo de imágenes. Una imagen $f(x, y)$ definida en el “mundo real” se considera como una función en dos dimensiones, donde (x, y) son las coordenadas espaciales (en un plano) y a la amplitud de f en cualquier par de coordenadas (x, y) se le denomina *intensidad* o *nivel de gris de la imagen*. Una subclase importante de sistemas son los denominados sistemas lineales invariantes a desplazamientos (LSI¹). Estos sistemas pueden ser caracterizados completamente por la salida del sistema para una entrada en particular, a la cual se denominada entrada puntual en el origen y a la salida resultante se le conoce como respuesta al impulso.

Un sistema \mathbf{T} se puede representar por un proceso de transformación para determinar una función $g(x)$ a partir de una función dada $f(x)$, es decir:

$$\mathbf{T}[f(x)] = g(x). \quad (1)$$

El sistema \mathbf{T} es *lineal si y solo si* para dos entradas arbitrarias $f_1(x)$ y $f_2(x)$ se tienen

¹LSI: Linear shift-invariant system

como salidas $g_1(x)$ y $g_2(x)$ respectivamente, y para dos constantes a y b , se tiene que

$$\mathbf{T}[af_1(x) + bf_2(x)] = a\mathbf{T}[f_1(x)] + b\mathbf{T}[f_2(x)], \quad (2)$$

para cada (x) . Esto se conoce como principio de superposición y de homogeneidad en la teoría de sistemas lineales .

El sistema \mathbf{T} es *invariante a los desplazamientos* si para cada $f(x)$ se obtiene:

$$\mathbf{T}[f(x - x')] = g(x - x'), \quad (3)$$

para cualquier desplazamiento x' . Lo anterior significa que un desplazamiento espacial en la señal de entrada produce la salida esperada, excepto por un desplazamiento idéntico.

La *respuesta al impulso unitario* $\delta(x - x')$ de un sistema \mathbf{T} con entrada $s(x)$ y salida $g(x)$, se denota por $h(x)$, es decir:

$$\mathbf{T}[\delta(x - x')] = h(x; x'), \quad (4)$$

la cual es la respuesta del sistema \mathbf{T} en la posición espacial (x) debido a un impulso localizado en la posición (x') . Cuando el sistema \mathbf{T} es invariante a los desplazamientos, si

$$\mathbf{T}[\delta(x)] = h(x), \quad (5)$$

representa la respuesta a un impulso localizado en el origen espacial, entonces se tiene que

$$\mathbf{T}[\delta(x - x')] = h(x - x'). \quad (6)$$

Lo anterior significa que la respuesta a un impulso aplicado en cualquier posición espacial se puede encontrar a partir de la respuesta al impulso dada por la ecuación (5).

Dicho en otras palabras, $h(x)$ puede caracterizar completamente a un *sistema lineal invariante a desplazamientos* si la respuesta al impulso se conoce. La respuesta $g(x)$ aplicada a una entrada arbitraria $f(x)$ se puede encontrar en términos de $h(x)$ por medio de una expresión que conlleva una convolución lineal (ver Papoulis (1962)), es decir

$$\begin{aligned} g(x) &= \mathbf{T}[f(x)] \\ &= \mathbf{T}\left[\int_{-\infty}^{\infty} f(x')\delta(x-x')dx'\right]. \end{aligned} \quad (7)$$

Si el sistema es lineal, entonces

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} f(x')\mathbf{T}[\delta(x-x')]dx' \\ &= \int_{-\infty}^{\infty} f(x')h(x;x')dx', \end{aligned} \quad (8)$$

y si el sistema es invariante a los desplazamientos, se tiene

$$\begin{aligned} g(x) &= \int_{-\infty}^{\infty} f(x')h(x-x')dx' \\ &= f(x) * h(x). \end{aligned} \quad (9)$$

donde el símbolo $*$ denota la operación de convolución. De esta manera la ecuación (9), representa una convolución entre las funciones $f(x)$ y $h(x)$. Las propiedades de la transformada de Fourier (FT) permiten caracterizar muy bien a un sistema lineal invariante a desplazamientos. La transformada de Fourier $H(u)$ de la respuesta al impulso $h(x)$ de un sistema lineal se le conoce como *respuesta en frecuencia* o *función de transferencia* del sistema o simplemente *función filtro*.

Los filtros lineales en el procesamiento de imágenes se caracterizan en términos de su respuesta en frecuencia, en el procesamiento de imágenes la convolución se lleva a

cabo en el dominio de Fourier empleando los algoritmos de la transformada rápida de Fourier (*FFT*). Cuando una imagen se convoluciona con un filtro espacial pequeño el proceso de convolución es más rápido si se realiza directamente en el dominio espacial. A continuación se analizarán algunos conceptos que serán útiles en este trabajo con relación a la FT.

II.2. La transformada de Fourier

Una manera de analizar las propiedades de una función imagen, es *descomponerla* en funciones más sencillas por medio de una combinación lineal de funciones ortonormales como la transformada de Fourier. La FT bidimensional de la función imagen $f(x, y)$ se define por la doble integral:

$$F(u, v) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) e^{-j2\pi(xu+yv)} dx dy. \quad (10)$$

Para el propósito del procesamiento de imágenes es razonable suponer que la transformada de Fourier es una función periódica que siempre existe. La transformada de Fourier inversa (*IFT*²) se define por:

$$f(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u, v) e^{j2\pi(xu+yv)} du dv. \quad (11)$$

Las variables (x, y) denotan las *coordenadas espaciales* o coordenadas de la imagen, y las coordenadas (u, v) son llamadas *frecuencias espaciales*. La función $f(x, y)$ del lado izquierdo de la ecuación (11) puede ser interpretada como una combinación lineal de funciones base periódicas $e^{j2\pi(xu+yv)}$. Los componentes reales e imaginarios de estas funciones base son las funciones seno y coseno, y la función $F(u, v)$ es una función peso que representa la influencia de las funciones base en $f(x, y)$.

²IFT: Inverse Fourier Transform

En general, se puede ver que los componentes de la transformada de Fourier son cantidades complejas. Como sucede en el análisis de números complejos, algunas veces es conveniente expresar $F(u, v)$ en coordenadas polares:

$$F(u, v) = |F(u, v)| e^{j\phi(u, v)}, \quad (12)$$

donde:

$$|F(u, v)| = \sqrt{R^2(u, v) + I^2(u, v)}, \quad (13)$$

se le llama *magnitud* de la transformada de Fourier, y

$$\phi(u, v) = \arctan \left[\frac{I(u, v)}{R(u, v)} \right], \quad (14)$$

se le denomina *fase* de la transformada. En las ecuaciones (13) y (14) $R(u, v)$ e $I(u, v)$ representan la partes real e imaginaria de $F(u, v)$, respectivamente. Otra cantidad que se utiliza es *el espectro de energía*, definido por el cuadrado de la magnitud:

$$\begin{aligned} P(u, v) &= |F(u, v)|^2 \\ &= F(u, v)F^*(u, v) \\ &= R^2(u, v) + I^2(u, v). \end{aligned} \quad (15)$$

donde el superíndice * denota el complejo conjugado. El término *densidad espectral* también se emplea para referirse al espectro de energía. Es importante señalar que la multiplicación de las transformadas en dos dimensiones se realiza elemento a elemento.

Denotando la transformada de Fourier por el operador \mathcal{F} , las ecuaciones (10) y (11) se pueden abreviar de la forma:

$$F(u, v) = \mathcal{F} \{f(x, y)\}, \quad (16)$$

y

$$f(x, y) = \mathcal{F}^{-1} \{F(u, v)\}. \quad (17)$$

II.2.1. Propiedades de la transformada de Fourier

Las siguientes propiedades de la transformada de Fourier son interesantes desde el punto de vista del procesamiento de imágenes:

Linealidad

La FT es lineal, es decir, para dos constantes arbitrarias complejas a y b se tiene que:

$$\mathcal{F} \{af_1(x, y) + bf_2(x, y)\} = aF_1(u, v) + bF_2(u, v). \quad (18)$$

Traslación

Si $\mathcal{F} \{f(x, y)\} = F(u, v)$ y (x_0, y_0) son constantes reales, entonces

$$\mathcal{F} \{f(x - x_0, y - y_0)\} = F(u, v)e^{-j2\pi(ux_0 + vy_0)}. \quad (19)$$

Esta propiedad de desplazamiento indica que la transformada de Fourier de una función desplazada en el espacio es la misma transformada de la función original (sin desplazar) multiplicada por un factor exponencial con una fase lineal, en otras palabras, una traslación en el espacio cambia el ángulo de fase del espectro.

Similarmente, la propiedad de desplazamiento en el dominio de la frecuencia indica que si $F(u, v)$ se desplaza por las constantes (u_0, v_0) , su transformada inversa se multiplica por $e^{j2\pi(ux_0 + vy_0)}$, es decir:

$$\mathcal{F} \{f(x, y)e^{j2\pi(ux_0 + vy_0)}\} = F(u - u_0, v - v_0), \quad (20)$$

De las propiedades de traslación se deriva que la magnitud de la transformada de

Fourier es invariante a los desplazamientos de entrada. Esta propiedad es importante y la explotan los filtros lineales que se estudian en este trabajo para localizar objetos en una imagen dada.

Simetría Conjugada

Si $f(x, y)$ es una función real, como es el caso de las imágenes, entonces,

$$F(-u, -v) = F^*(u, v), \quad (21)$$

y

$$F(u, v) = F^*(-u, -v). \quad (22)$$

De lo anterior se deriva que

$$|F(u, v)| = |F(-u, -v)|. \quad (23)$$

Una función imagen es siempre real y si también es simétrica o par, es decir $f(x, y) = f(-x, -y)$, entonces el resultado de la transformada de Fourier $F(u, v)$ también es una función real.

Teorema de convolución

La convolución en el dominio espacial entre dos funciones bidimensionales, se define por la doble integral:

$$f(x, y) * g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y') g(x - x', y - y') dx' dy'. \quad (24)$$

De acuerdo con el teorema de convolución, la ecuación anterior se puede expresar por:

$$f(x, y) * g(x, y) = \mathcal{F}^{-1} \{F(u, v)G(u, v)\}, \quad (25)$$

pues este teorema establece que

$$\begin{aligned}\mathcal{F}\{f(x, y) * g(x, y)\} &= F(u, v)G(u, v), \\ \mathcal{F}\{f(x, y)g(x, y)\} &= F(u, v) * G(u, v),\end{aligned}\tag{26}$$

donde $F(u, v)$ y $G(u, v)$ son las transformadas de Fourier de $f(x, y)$ y $g(x, y)$ respectivamente.

Teorema de correlación

Tanto la integral de correlación como la de convolución son de suma importancia en aplicaciones teóricas y prácticas. La correlación para funciones reales se define por:

$$f(x, y) \circ g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y')g(x' - x, y' - y)dx' dy',\tag{27}$$

donde el símbolo \circ denota la operación de correlación. En el caso de funciones complejas, algunos autores (Gonzalez y Woods, 2002) prefieren denotar la correlación por

$$f(x, y) \circ g(x, y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x', y')g^*(x' + x, y' + y)dx' dy',\tag{28}$$

aunque (27) y (28) son equivalentes para funciones reales. El teorema de correlación indica que

$$\begin{aligned}\mathcal{F}\{f(x, y) \circ g(x, y)\} &= F(u, v)G^*(u, v), \\ \mathcal{F}\{f(x, y)g^*(x, y)\} &= F(u, v) \circ G(u, v).\end{aligned}\tag{29}$$

Por lo tanto la correlación se puede calcular por:

$$f(x, y) \circ g(x, y) = \mathcal{F}^{-1}\{F(u, v)G^*(u, v)\},\tag{30}$$

donde el superíndice * indica el complejo conjugado. En realidad, el segundo miembro de la igualdad es $\mathcal{F}^{-1}\{F(u,v)G(-u,-v)\}$ pero como se restringe al caso donde $f(x,y)$ y $g(x,y)$ son funciones reales, entonces se puede notar que $G(-u,-v) = G^*(u,v)$ y que cumplen con la propiedad de simetría conjugada.

Si $f(x,y)$ y $g(x,y)$ son la misma función en la ecuación (27) normalmente se le llama función de *autocorrelación*, si las funciones son diferentes se le llama *correlación cruzada*.

Teorema de Parseval

Este teorema representa la conservación de la energía entre el espacio del objeto y el plano de Fourier. El teorema de Parseval establece que:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x,y)|^2 dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |F(u,v)|^2 du dv. \quad (31)$$

La versión generalizada es la siguiente:

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)h^*(x,y) dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} F(u,v)H^*(u,v) du dv. \quad (32)$$

Desigualdad de Cauchy-Schwartz

La desigualdad de Cauchy-Schwartz establece que

$$\begin{aligned} & \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)g(x,y) dx dy \right|^2 \\ & \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f(x,y)|^2 dx dy \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g(x,y)|^2 dx dy. \end{aligned} \quad (33)$$

Si se representa $f(x,y) = f_1(x,y)\sqrt{p(x,y)}$ y $g(x,y) = g_1(x,y)\sqrt{p(x,y)}$, con una función peso real positiva, es decir $p(x,y) \geq 0$, se consigue una forma más general de

la desigualdad:

$$\begin{aligned} & \left| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_1(x, y) g_1(x, y) p(x, y) dx dy \right|^2 \\ & \leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f_1(x, y)|^2 p(x, y) dx dy \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |g_1(x, y)|^2 p(x, y) dx dy. \end{aligned} \quad (34)$$

y la igualdad se logra *si y solo si* $f_1(x, y) = \alpha g_1^*(x, y)$, donde α es una constante arbitraria.

II.3. Análisis de correlación

El análisis de correlación es un método ampliamente usado para detectar cualquier similitud entre dos señales que varían en el tiempo. Sin embargo, se debe tener cuidado porque se puede obtener una correlación arbitrariamente grande al multiplicar una de las funciones por una constante determinada. Esto significa que la correlación es también una medida de energía hablando de las funciones que son comparadas, y éstas se deben normalizar de alguna manera para obtener un criterio similar.

De acuerdo con el teorema de Correlación dado por la ecuación (29), la correlación se puede calcular tanto en el dominio espacial (efectuando operaciones directas sobre los valores de los píxeles en una imagen) o bien en el dominio de las frecuencias (multiplicando las transformadas de Fourier de las funciones imagen involucradas).

Para ejemplificar el funcionamiento de la correlación se presenta la Fig. 2. Recuérdese que en un sistema lineal, a la función $h(x)$ se le denomina respuesta al impulso. Dada una función $f(x)$ como la entrada a dicho sistema, entonces la salida de correlación dada por $c(x)$ muestra que el valor máximo se localiza en la coordenada $x = 80$, que es precisamente donde coinciden las dos funciones. Este principio se explota en el desarrollo de esta investigación para determinar la presencia y posición de un objeto en una imagen

determinada. Estos resultados se obtuvieron al utilizar las funciones de la transformada rápida de Fourier que vienen incorporadas en las librerías de MATLAB[®].

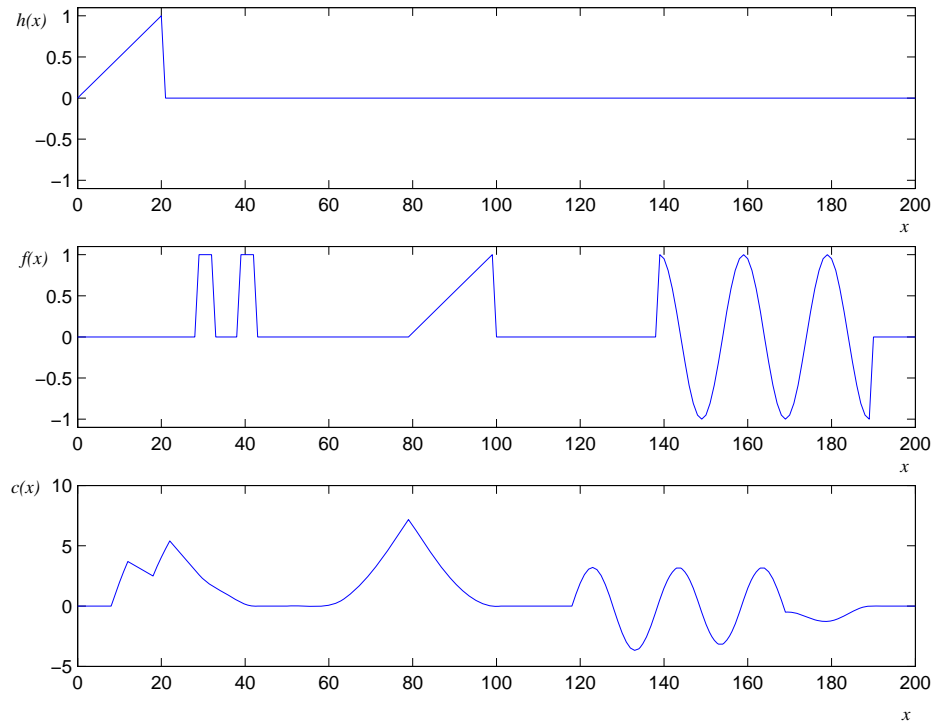


Figura 2: Correlación entre dos funciones: $c(x)$ es la correlación entre $f(x)$ y $h(x)$.

Como se pudo apreciar, tanto la convolución como la correlación de dos señales se pueden obtener al aplicar la transformada inversa de Fourier al producto de las transformadas de las dos señales. La diferencia entre estas operaciones radica en el simple conjugado de la segunda señal. Además, se tiene que para funciones reales, la correlación está relacionada con la convolución de la siguiente manera :

$$\begin{aligned}
f(x) * h(-x) &= \int_{-\infty}^{\infty} f(x')h[-(x-x')] dx' \\
&= \int_{-\infty}^{\infty} f(x')h(x'-x) dx' \\
&= f(x) \circ h(x).
\end{aligned} \tag{35}$$

Lo que significa que para funciones reales y pares, la convolución y la correlación son equivalentes.

II.4. Correlación entre funciones discretas

Hasta este momento solo se han considerado funciones continuas para representar la información contenida en las imágenes, y aunque esto proporciona mayor formalismo, son inadecuadas para la mayoría de los propósitos prácticos. Antes de continuar se define una imagen digital.

Una *imagen digital* es una imagen $f(x)$ que se ha discretizado tanto en las coordenadas espaciales como en el brillo o intensidad. Una imagen digital puede considerarse como una matriz cuyos índices de fila y columna (m, n) identifican un punto de la imagen y el valor del correspondiente elemento de la matriz indica el nivel de gris en ese punto. Los elementos de una distribución digital de este tipo se denominan píxeles.

La forma en que influye la discretización de señales en el procesamiento de Fourier es especialmente importante y de aquí surge la necesidad de un análisis de estos efectos. Las señales discretas se representan por un arreglo de valores $\{f(x_k)\}$, los cuales se toman de la función $f(x)$ en cada punto $x_k = k\Delta x$, donde Δx es llamado *intervalo de muestreo*.

Nuestra atención se enfoca a las señales de *banda limitada*, es decir, aquellas funciones cuyas transformadas de Fourier se anulan en ciertas frecuencias. Las imágenes

naturales se pueden considerar con amplia generalidad como miembros de esta clase especial de funciones. La intuitiva idea de que cuanto más pequeño es el intervalo de muestreo más exacta es la representación de la señal, es válida hasta cierto límite. Usando técnicas de interpolación, un cierto intervalo de muestreo es suficiente para reconstruir con exactitud una señal y una Δx más pequeña no implica una mejor descripción de $f(x)$. Las funciones de banda limitada se definen como aquellas cuya transformada de Fourier tiene una región de soporte finita, donde la región de soporte es el subconjunto del dominio donde la función tiene un valor diferente de cero. Este resultado se resume en el siguiente teorema de muestreo de Whittaker-Shannon.

II.4.1. Teorema del muestreo

Sea $f(x)$ una función de banda limitada, cuya transformada de Fourier $F(u)$ es diferente de cero para valores de (u) fuera del intervalo $[-W, W]$. La condición necesaria y suficiente para que $f(x)$ se pueda recuperar a partir de un conjunto de muestras $\{f(k\Delta x)\}$ es:

$$\Delta x \leq \frac{1}{2W} \equiv \frac{1}{v_n}, \quad (36)$$

donde $v_n = 2W$ es conocida como la frecuencia Nyquist.

La demostración de este teorema se puede encontrar en Papoulis (1962) y en Wozencraft y Jacobs (1965). La posibilidad de recuperar una señal $f(x)$ está basada en el hecho de que cuando se muestrea una señal en un dominio, se obtienen réplicas de la señal original en el otro dominio con un periodo Δu dado por

$$\Delta u = \frac{1}{\Delta x}, \quad (37)$$

donde Δx es el intervalo de muestreo, considerando que una función discreta tiene una transformada de Fourier periódica y una función periódica tiene una FT discreta.

Si la FT de $f(x)$ es cero fuera del intervalo $[-W, W]$, la condición para que las réplicas periódicas de la señal no se traslapen está dada por la ecuación (36). Finalmente, si uno de los términos periódicos se puede aislar del resto, la función original se puede reconstruir simplemente al aplicar la transformada de Fourier inversa.

El teorema de Whittaker-Shannon establece el límite menor para el cual una función se puede digitalizar sin pérdida de información durante el proceso de muestreo: la tasa de muestreo debe ser dos veces la frecuencia más grande contenida en la señal. Las equivocaciones de no observar este principio llevan a errores conocidos bajo el nombre genérico de *traslape* (en inglés se usa el término *aliasing*) o pérdida de información.

II.4.2. Transformada discreta de Fourier

Como ya se mencionó, emplear funciones continuas es una forma elegante y sencilla de entender estos conceptos, pero para aplicaciones reales donde se utiliza una computadora digital, se requiere manejar funciones discretas (imágenes digitales), entonces se emplea la transformada discreta de Fourier (DFT^3) cuyo par de transformadas está dado por:

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(xu/M + yv/N)}, \quad (38)$$

para $u = 0, 1, \dots, M - 1$ y $v = 0, 1, \dots, N - 1$, y

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(xu/M + yv/N)}, \quad (39)$$

para $x = 0, 1, \dots, M - 1$ y $y = 0, 1, \dots, N - 1$. Donde $M \times N$ indica las dimensiones de la imagen (número de píxeles).

En la práctica, para calcular este par de transformadas se utiliza el algoritmo de *la transformada rápida de Fourier (FFT)*, el cual reduce el número de multiplicaciones y

³DFT: Discrete Fourier Transform

sumas complejas de un orden N^2 a $N \log_2 N$. Lo que representa un ahorro considerable de tiempo cuando N es relativamente grande.

Todos los comentarios de las secciones anteriores acerca de la convolución, correlación y la transformada de Fourier en el dominio de funciones continuas, también son válidos para funciones discretas (imágenes digitales). Las integrales se cambian por sumas en las ecuaciones respectivas.

II.5. Resumen

En este capítulo se hizo un repaso de los conceptos básicos de señales y sistemas. Se definió la transformada de Fourier y sus propiedades para caracterizar a los sistemas lineales e invariantes a desplazamientos. Se dió una breve introducción al análisis por correlación, que es el principio que se explota para la detección de objetos dentro de imágenes digitales. En el siguiente capítulo se abordan con detalle algunos filtros clásicos de correlación para el reconocimiento de objetos en general.

Capítulo III

DetECCIÓN Y ESTIMACIÓN DE PARÁMETROS EN IMÁGENES

Existen diferentes tipos de modelo de señal para diferentes tipos de ruido (aditivo, multiplicativo, no trasladado, etc.), que pueden contaminar a las imágenes. En este capítulo nos enfocaremos exclusivamente al modelo de señal aditivo (señal + ruido), el cual consiste de una señal conocida que ha sido alterada por ruido aditivo. Cuando una señal es alterada por ruido, que es aleatorio debido al proceso de captura de la señal, frecuentemente es importante extraer o restaurar la señal original de su versión ruidosa. A este proceso se le conoce como restauración. Por otro lado, nuestro interés se enfoca en la clasificación de una señal como una versión ruidosa de las posibles versiones de la señal original. Es posible que en la señal de entrada (o escena) solo se encuentre ruido, entonces el objetivo es determinar si se encuentra la señal deseada o no dentro del escenario de detección. A esta tarea se le conoce como detección. Una generalización del concepto de detección es la estimación, en donde el objetivo es estimar un parámetro (que también puede ser un valor dentro de un intervalo en lugar de un conjunto discreto de valores) de una señal ruidosa.

En este capítulo se derivan los filtros espaciales de acoplamiento como resultado de analizar los conceptos de detección y estimación de parámetros en las imágenes. Primero se estudia el caso de detectar una señal dentro de dos clases de señales. Posteriormente se introduce la noción de estimación de parámetros para la derivación de filtros de correlación.

III.1. Prueba de hipótesis binaria para la detección de señales

Aunque el uso de los sistemas basados en la correlación para detectar señales se remonta a principios de los 40's al resolver un problema en teoría de comunicaciones (Wozencraft y Jacobs, 1965), no fue sino hasta mediados de los 60's que Vanderlugt diseñó uno en el plano de las frecuencias y se volvieron muy populares los filtros espaciales de acoplamiento para el reconocimiento de patrones en óptica (*OPR*¹). Gracias al algoritmo de la transformada rápida de Fourier los sistemas basados en la correlación también son populares en implementaciones digitales. Las simulaciones que se muestran a lo largo de este trabajo se programaron en rutinas MATLAB[®]. En esta sección, se mostrará la importancia de los sistemas de correlación en el reconocimiento de objetos.

III.1.1. Filtro espacial de acoplamiento

Considérese un escenario de detección de señales en el cual la señal recibida $s(x)$ se obtiene de una de dos posibles maneras. Una posibilidad (a la que se hará referencia como hipótesis H_0) solo contiene ruido, denotado por $n(x)$. En otras palabras, H_0 denota la ausencia de la señal de referencia. La otra posibilidad (la hipótesis H_1) supone que $s(x)$ es la suma de una señal conocida $t(x)$ contaminada con ruido aditivo $n(x)$.

$$\begin{aligned} H_0 : s(x) &= n(x) \\ H_1 : s(x) &= t(x) + n(x). \end{aligned} \tag{40}$$

Entonces, dada la señal observada $s(x)$ y la información estadística conocida acerca del ruido y la señal de referencia $t(x)$, el objetivo de la prueba es decidir entre las dos hipótesis. El ruido se supone estacionario, de media cero y con una densidad espectral de

¹OPR: Optical Pattern Recognition

energía $P_n(u)$. El problema de decidir entre H_0 y H_1 se aborda incorporando $s(x)$ a un filtro lineal e invariante en el tiempo con función de transferencia $H(u)$ y posteriormente comparando el pico de salida con un umbral T . Si el pico excede T , se selecciona H_1 . De lo contrario se elige H_0 . Ahora la pregunta es: ¿Cómo elegir el filtro $H(u)$? Esta arquitectura básica se resume en la Fig. 3.

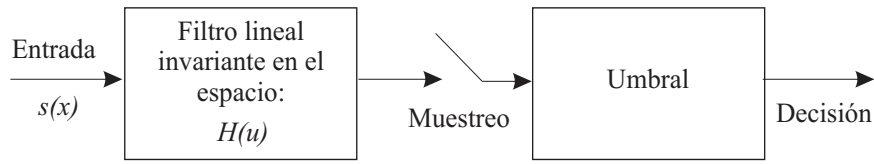


Figura 3: Diagrama de bloques de una identificación binaria.

Antes de elegir el filtro $H(u)$, debemos establecer un sistema métrico adecuado para caracterizar sus virtudes. En el escenario de identificación, se desea que el valor del pico de prueba sea muy grande para H_1 y muy pequeño para H_0 . Sin embargo, debido a la naturaleza aleatoria de $n(x)$, pueden ocurrir errores y ocasionalmente el pico puede exceder T para H_0 o estar por debajo de T para H_1 . Una buena medida para caracterizar el efecto del ruido en el pico de salida es la relación señal-ruido (SNR^2) definida por:

$$SNR = \frac{|E\{\eta|H_1\}|^2}{var\{\eta\}}, \quad (41)$$

donde η es el valor del pico en la salida del filtro y $E\{\cdot\}$ y $var\{\cdot\}$ indican el valor esperado y la varianza respectivamente.

El numerador en la ecuación (41) se refiere al valor promedio del pico dada la hipótesis H_1 , el cual se desea lo más grande posible. El denominador en la ecuación

²SNR: Signal to Noise Ratio

(41) se refiere a la varianza de η (la misma para H_1 y H_0), que se desea minimizar. Cuando el ruido $n(x)$ es Gaussiano, la SNR en (41) esta monótonamente relacionada a la probabilidad de error en el sentido de que al maximizar la SNR se conducirá a una probabilidad de error mínima. De esta manera el objetivo consiste en seleccionar $H(u)$ de tal forma que maximice la SNR en la ecuación (41).

Considerando que el ruido $n(x)$ posee media cero, $E\{\eta|H_1\}$ es insignificante comparado con la salida del pico cuando $t(x)$ pasa a través del filtro $H(u)$ (Vijaya-Kumar, 1994). Este pico está ubicado generalmente en el origen de la salida de la correlación, de tal forma que

$$E\{\eta|H_1\} = \int_{-\infty}^{\infty} T(u)H(u) du. \quad (42)$$

La varianza de η se debe solamente al ruido de entrada $n(x)$ del filtro. La varianza es insignificante comparada con la energía total en la salida del filtro (Vijaya-Kumar, 1994). La densidad espectral de la energía del ruido en la salida del filtro está dada por $P_n(u) |H(u)|^2$:

$$var\{\eta\} = \int_{-\infty}^{\infty} P_n(u) |H(u)|^2 du. \quad (43)$$

Sustituyendo las ecuaciones (42) y (43) en la ecuación (41), se obtiene una expresión para SNR en términos de $H(u)$:

$$SNR = \frac{\left| \int_{-\infty}^{\infty} T(u)H(u) du \right|^2}{\int_{-\infty}^{\infty} P_n(u) |H(u)|^2 du}. \quad (44)$$

Con la ayuda de la desigualdad de Cauchy-Schwarz se encuentra la $H(u)$ que maximiza esta SNR :

$$H(u) = \alpha \frac{T^*(u)}{P_n(u)}. \quad (45)$$

donde α es una constante arbitraria compleja. Los detalles del procedimiento para

encontrar la SNR_{max} se describen en la sección IV.1.1.

Considérese el caso especial de ruido blanco cuya densidad espectral de energía $P_n(u)$ es una constante N_0 . Entonces el filtro óptimo está dado por

$$H(u) = \alpha T^*(u). \quad (46)$$

Aplicando la transformada de Fourier inversa y suponiendo arbitrariamente que $\alpha = 1$ y que $s(x)$ es una función real, se obtiene

$$h(x) = t(-x), \quad (47)$$

donde $t(-x)$ es la IFT de $T^*(u)$. Así, la salida $c(x)$ de este filtro está dada por la convolución de la señal recibida $s(x)$ con la respuesta al impulso $h(x)$ del filtro, es decir,

$$\begin{aligned} c(x) &= s(x) * h(x) \\ &= s(x) * t(-x) \\ &= \int_{-\infty}^{\infty} s(x')t(x' - x)dx', \end{aligned} \quad (48)$$

donde el símbolo $*$ denota la operación de la convolución. De modo que la salida $c(x)$ del filtro óptimo es la correlación cruzada entre la señal recibida $s(x)$ con la señal de referencia $t(x)$. De esta manera estos sistemas de correlación son óptimos desde el punto de vista de la SNR cuando el ruido de entrada es blanco. Además, la respuesta al impulso $h(x)$ del filtro es una replica invertida de la señal de referencia $t(x)$, esta es la razón por la que recibe el nombre de *filtro de acoplamiento* o CMF^3 por sus siglas en inglés.

³CMF: Clasical Matched Filter

III.2. Estimación de parámetros en imágenes.

La estimación de parámetros es uno de los problemas fundamentales con muchas aplicaciones en el procesamiento de señales en general y en particular en el procesamiento de imágenes. El problema se formula de la siguiente manera: Dadas las muestras observadas de una señal de entrada (imagen) $\{s_k, k = 0, 1, \dots, N - 1\}$ que implícitamente dependen del parámetro θ , el problema consiste en determinar este parámetro con la mejor precisión posible. Se supone que la manera en la que el parámetro determina la señal depende de la exactitud de algunos factores aleatorios implicados en el proceso de captura de la señal, como puede ser ruido aleatorio. El parámetro a estimar puede tomar valores arbitrarios. Se puede suponer que las probabilidades $P(\theta)$ de estos valores se conocen de antemano, si los parámetros toman valores en un intervalo continuo se utiliza la densidad de probabilidad $p(\theta)$ para estimarlos. A este tipo de probabilidades se les denomina probabilidades a priori. Debido a la naturaleza de los factores aleatorios involucrados en el problema, la calidad de la estimación de los parámetros solo puede ser evaluada en términos estadísticos. Esta evaluación depende de la probabilidad de error (para los parámetros que toman un conjunto de valores finitos y discretos) o, para los parámetros con valores en un intervalo continuo se usa la densidad de probabilidad de la estimación del error $p(\theta - \hat{\theta})$ donde $\hat{\theta}$ es el parámetro estimado. Se sabe de la teoría de decisión y de la teoría de estimación de parámetros que el estimador que provee un mínimo de probabilidad de error es aquel que tiene una probabilidad a posteriori $P(\theta | \{s_k\})$ máxima (Duda *et al.*, 2001), es decir, la probabilidad de que el parámetro tenga un valor particular θ dada la señal $\{s_k\}$, lo cual se expresa:

$$\hat{\theta}_{opt} = \arg \max_{\theta} P(\theta | \{s_k\}), \quad (49)$$

donde $\arg \max$ solicita el valor o valores del argumento θ que maximizan la expresión

$P(\theta | \{s_k\})$.

De aquí en adelante se hace referencia a la Probabilidad a posteriori máxima como estimación *MAP*⁴. Por lo tanto, dada la señal observada $\{s_k\}$, el dispositivo óptimo de la estimación del parámetro debe ser capaz de medir las probabilidades a posteriori de los valores del parámetro y de determinar el valor para el cual esta probabilidad es máxima. Lo anterior significa que se debe especificar exactamente la manera en que los factores aleatorios dentro de la señal observada determinan la señal de entrada.

III.2.1. Modelo aditivo de ruido Gaussiano blanco (AWGN)

Uno de los modelos más aceptados en la observación de señales, es el modelo aditivo de ruido Gaussiano blanco (*AWGN*⁵), donde el ruido se considera independiente de la señal, este modelo se describe por

$$\{s_k\} = \{t_k(\theta) + n_k\}, \quad (50)$$

donde $t_k(\theta)$ son las muestras de una señal hipotéticamente “verdadera”, la cual es determinísticamente dependiente del parámetro θ , y $\{n_k\}$ son los valores aleatorios del ruido estadísticamente independientes de $t_k(\theta)$. Por otro lado, el ruido $\{n_k\}$ se considera con media cero y desviación estándar σ , por lo tanto tiene una densidad de probabilidad:

$$p_n(n) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{n^2}{2\sigma^2}\right]. \quad (51)$$

Dado el parámetro θ a partir de la ecuación (50) se puede encontrar que la densidad de probabilidad de las muestras de la señal de entrada está dada por:

$$P(\{s_k\} | \theta) = P(\{n_k = s_k - t_k(\theta)\}), \quad (52)$$

⁴MAP: Maximum a posteriori probability

⁵AWGN: Additive White Gaussian Noise

y la probabilidad a posteriori $P(\theta|\{s_k\})$ se puede derivar de la regla de Bayes:

$$P(\theta|\{s_k\}) = \frac{P(\{s_k\}|\theta)p(\theta)}{p(\{s_k\})}, \quad (53)$$

nótese que $p(\{s_k\})$ no depende de θ , por lo tanto la probabilidad a posteriori se puede expresar:

$$P(\theta|\{s_k\}) \cong P(\{s_k\}|\theta)p(\theta) = P(n_k = s_k - t_k(\theta))p(\theta). \quad (54)$$

Al sustituir la ecuación (51) en (54):

$$\begin{aligned} P(\theta|\{s_k\}) &\cong \prod_{k=0}^{N-1} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2} |s_k - t_k(\theta)|^2\right) p(\theta) \\ &= \frac{1}{(2\pi\sigma^2)^{(N-1)/2}} \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^{N-1} |s_k - t_k(\theta)|^2\right) p(\theta), \end{aligned} \quad (55)$$

el factor $\frac{1}{(2\pi\sigma^2)^{(N-1)/2}}$ es independiente de k por lo que se puede descartar sin pérdida de optimización, entonces la ecuación anterior se puede expresar por

$$\begin{aligned} P(\theta|\{s_k\}) &\cong \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^{N-1} |s_k - t_k(\theta)|^2\right) p(\theta) \\ &= \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^{N-1} |s_k - t_k(\theta)|^2 + \ln(p(\theta))\right) \end{aligned} \quad (56)$$

$$= \exp\left(-\frac{1}{2\sigma^2} \sum_{k=0}^{N-1} s_k^2 + \frac{1}{\sigma^2} \sum_{k=0}^{N-1} s_k t_k(\theta) - \frac{1}{2\sigma^2} \sum_{k=0}^{N-1} t_k^2(\theta)\right) p(\theta), \quad (57)$$

como el término s_k no depende del parámetro θ también se puede eliminar la primer suma dentro del exponencial sin pérdida de optimización:

$$P(\theta|\{s_k\}) \cong \exp\left(\frac{1}{\sigma^2} \sum_{k=0}^{N-1} s_k t_k(\theta) - \frac{1}{2\sigma^2} \sum_{k=0}^{N-1} t_k^2(\theta) + \ln(p(\theta))\right), \quad (58)$$

de aquí que el estimador *MAP* se obtiene al maximizar el logaritmo de $P(\theta|\{s_k\})$, es decir:

$$\begin{aligned}
\hat{\theta}_{MAP}^{opt} &= \arg \max_{\theta} \{ \ln P(\theta | \{s_k\}) \} \\
&= \arg \max_{\theta} \left\{ \frac{1}{\sigma^2} \sum_{k=0}^{N-1} s_k t_k(\theta) - \frac{1}{2\sigma^2} \sum_{k=0}^{N-1} t_k^2(\theta) + \ln(p(\theta)) \right\}. \quad (59)
\end{aligned}$$

Si la probabilidad a priori del parámetro $p(\theta)$ se considera uniformemente distribuida, entonces el estimador *MAP* se convierte en la estimación por máxima verosimilitud (o estimador *ML*⁶) es decir,

$$\hat{\theta}_{ML}^{opt} = \arg \max_{\theta} \left\{ \frac{1}{\sigma^2} \sum_{k=0}^{N-1} s_k t_k(\theta) - \frac{1}{2\sigma^2} \sum_{k=0}^{N-1} t_k^2(\theta) \right\} \quad (60)$$

El estimador *ML* también se usa cuando se desconoce la distribución a priori del parámetro.

Las ecuaciones (59) y (60) definen la estructura del dispositivo estimador de un parámetro, lo que implica que este dispositivo debe consistir de una unidad para calcular la correlación entre la señal de entrada $\{s_k\}$ y la señal de referencia $\{t_k(\theta)\}$ para diferentes valores del parámetro, seguida por una unidad para generar una estimación al encontrar el valor más grande en la salida de la correlación.

III.2.2. Estimación de la posición de la señal en el modelo AWGN

Un caso importante en la estimación de parámetros de imágenes se encuentra en la localización de un objeto. El objetivo de la localización consiste en determinar las coordenadas de un objeto contenido dentro de una imagen. Sean (x_0, y_0) las coordenadas desconocidas de un objetivo, definidas por la observación de sus muestras $\{t_k(x_0, y_0)\}$, las cuales están mezcladas con ruido aditivo Gaussiano blanco e independiente como en

⁶ML: Maximum likelihood

el modelo *AWGN*. Primero, nótese que la energía de la señal del objetivo, representada por el segundo término en el exponencial de la ecuación (59) ($\sum t_k^2(\theta)$) no depende de las coordenadas del objeto, entonces a partir de la ecuación (59) se puede encontrar que el estimador *MAP* de coordenadas se define por la ecuación:

$$(\hat{x}_0, \hat{y}_0)_{MAP}^{opt} = \arg \max_{(x_0, y_0)} \left\{ \frac{1}{\sigma^2} \sum_{k=0}^{N-1} s_k t_k(x_0, y_0) + \ln(p(x_0, y_0)) \right\} \quad (61)$$

donde $p(x_0, y_0)$ es la densidad de distribución a priori de las coordenadas (Yaroslavsky y Eden, 1996).

Aplicando el teorema del muestreo a la ecuación anterior, las señales discretas $\{s_k\}$ y $\{t_k(x_0, y_0)\}$ se pueden convertir a sus correspondientes señales continuas $s(x, y)$ y $t(x - x_0, y - y_0)$, de aquí que

$$(\hat{x}_0, \hat{y}_0)_{MAP}^{opt} = \arg \max_{(x_0, y_0)} \left\{ \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} s(x, y) t(x - x_0, y - y_0) dx dy + N_0 \ln(p(x_0, y_0)) \right\} \quad (62)$$

donde N_0 es la densidad espectral del ruido.

Como las coordenadas del objeto son parámetros desplazados en el espacio, la función de correlación entre la señal de entrada y la señal de referencia necesaria para encontrar el estimador óptimo de las coordenadas (\hat{x}_0, \hat{y}_0) se puede implementar por un filtro lineal con una respuesta al impulso dada por la señal de referencia $t(x, y)$. Esta implementación es denominada *filtrado por acoplamiento* y se esquematiza en la Fig. 4. La respuesta en frecuencia $H_{opt}(u, v)$ del filtro de acoplamiento se puede calcular como el complejo conjugado de la transformada de Fourier de la imagen de referencia, es decir:

$$H_{opt}(u, v) = T^*(u, v). \quad (63)$$

En la sección III.1.1 se derivó este filtro por medio de pruebas de hipótesis. En el

capítulo IV se mostrará que el filtro CMF maximiza la relación SNR.

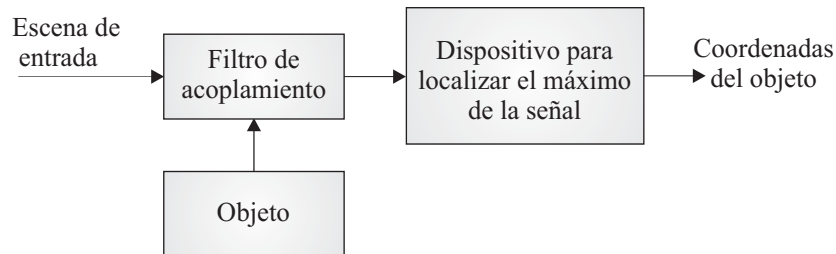


Figura 4: Filtrado por acoplamiento para la localización de objetos.

III.3. Resumen

En este capítulo se introdujeron los principales conceptos que se emplearán en los siguientes capítulos. Tales conceptos incluyen el modelo de señal, tipo de ruido, relación señal-ruido, etc. Pero sobre todo la importancia de los sistemas de correlación como sistemas óptimos para detectar y localizar objetos dentro de una escena dada. Por medio de la prueba de hipótesis se mostró la derivación de un filtro clásico en el reconocimiento de objetos en óptica, el CMF. La estimación de ciertos parámetros en una imagen, como las coordenadas de un objeto, se puede realizar por medio de la máxima verosimilitud al incluir la información a priori que se tiene acerca del parámetro que se desea estimar. Con la ayuda de este estimador estadístico se pueden derivar los filtros espaciales de correlación, demostrando que la posición del objeto está dado por la ubicación del pico de correlación.

Capítulo IV

Filtros clásicos de correlación

El objetivo de este capítulo es proporcionar las ideas básicas del uso de la correlación como una herramienta para el reconocimiento de patrones. La correlación se puede pensar como la salida de un filtro de acoplamiento (un filtro lineal e invariante a desplazamientos cuya respuesta al impulso corresponde a la versión invertida del objeto de referencia) y que en el capítulo anterior se demostró que es óptimo para detectar señales conocidas contaminadas por ruido aditivo blanco. A pesar de lo anterior, en un escenario de detección se pueden obtener valores grandes de correlación para objetos parecidos, lo que nos llevaría a falsas detecciones o clasificaciones. Otro problema que puede surgir, es cuando los objetos presentan distorsiones geométricas como desplazamientos laterales, cambios de escala, etc. Por ejemplo, la salida del CMF se deteriora rápidamente cuando se presentan distorsiones en el objeto de referencia. Por todo lo expuesto anteriormente y a lo largo de varias décadas se han desarrollado diferentes versiones de filtros de correlación que permiten superar ciertas deficiencias en los filtros espaciales.

En este capítulo se derivan algunos filtros de correlación en base a algún criterio de calidad. En la literatura se pueden encontrar un sinnúmero de filtros de correlación, sin embargo, los discutidos en este trabajo de investigación se seleccionaron por ser óptimos desde algún punto de vista, por ejemplo el filtro de acoplamiento CMF que se deriva al maximizar la relación señal-ruido también se puede derivar desde el punto de vista estadístico al minimizar el error de localización del pico de correlación (estimación de máxima verosimilitud). Por otro lado, todos los filtros que se discuten en este capítulo

mantienen la característica de ser invariantes a los desplazamientos laterales gracias a las propiedades de la transformada de Fourier.

IV.1. Criterios para derivar filtros de correlación

La correlación se puede implementar tanto para el procesamiento óptico como digital. Desde el desarrollo del filtro de acoplamiento clásico propuesto por VanderLugt (VanderLugt, 1964), varios filtros de correlación se han propuesto para el reconocimiento de patrones. Primero, a partir del hecho de que la fase de la transformada de Fourier contiene información más significativa que el módulo, Horner y Gianino (1984) sugirieron el *filtro sólo de fase*, el cual produce picos de correlación más altos y más agudos que el filtro de acoplamiento. Otro tipo de filtro espacial, que provee de una función delta o impulso como salida de localización (pico) en el plano de correlación, es el *filtro inverso*, también utilizado en la tarea de restauración de imágenes (recuperación de imágenes degradadas) (Gonzalez y Woods, 2002). A continuación se describe la derivación de los filtros *CMF*, *POF* e *IF* al optimizar alguna medida de desempeño.

IV.1.1. Optimización de la relación señal-ruido: filtro de acoplamiento clásico

Se ha mencionado que la relación señal-ruido en presencia de ruido blanco se optimiza por un filtro que no es otra cosa que el complejo conjugado de la transformada de Fourier del patrón a ser identificado. Este resultado se puede generalizar para ruido

correlacionado utilizando la desigualdad de Cauchy-Schwarz:

$$\begin{aligned}
SNR_{H(u,v)} &= \alpha \frac{\left| \iint T(u,v)H(u,v)dudv \right|^2}{\iint P_n(u,v)|H(u,v)|^2dudv} \\
&= \alpha \frac{\left| \iint \left[\frac{T(u,v)}{\sqrt{P_n(u,v)}} \right] \left[\sqrt{P_n(u,v)}H(u,v) \right] dudv \right|^2}{\iint P_n(u,v)|H(u,v)|^2dudv} \\
&\leq \alpha \frac{\iint \frac{|T(u,v)|^2}{P_n(u,v)}dudv \iint P_n(u,v)|H(u,v)|^2dudv}{\iint P_n(u,v)|H(u,v)|^2dudv} \\
&= \alpha \iint \frac{|T(u,v)|^2}{P_n(u,v)}dudv, \tag{64}
\end{aligned}$$

donde los límites de integración van de $-\infty$ a ∞ , y se consigue la igualdad *si y solo si*

$$H(u,v) = \beta \frac{T^*(u,v)}{P_n(u,v)} \equiv H_{CMF}(u,v) \tag{65}$$

con la constante compleja arbitraria β y la función de densidad espectral del ruido $P_n(u,v)$.

Este filtro se conoce como *CMF* o filtro de acoplamiento clásico y consiste de una versión blanqueada del complejo conjugado de la transformada de Fourier del objeto, multiplicado por un insignificante factor que es inversamente proporcional a la densidad espectral de energía del ruido de entrada.

IV.1.2. Optimización de la eficiencia de luz: filtro sólo de fase

Mientras los filtros de correlación están inspirados por el hecho de que los filtros de acoplamiento optimizan la relación *SNR*, en muchos casos este parámetro no es de

mayor interés. La eficiencia de luz se obtiene por la expresión:

$$\eta_H = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |c(x, y)|^2 dx dy}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |s(x, y)|^2 dx dy}, \quad (66)$$

y representa la razón de la intensidad de la luz en el plano de correlación $|c(x, y)|$ y la luz total en la entrada dada por $|s(x, y)|^2$. La eficiencia de luz, en términos del filtro se puede expresar como sigue:

$$\eta_H = \frac{\int \int |S(u, v)H(u, v)|^2 dudv}{\int \int |S(u, v)|^2 dudv} \quad (67)$$

la cual indica la relación entre la luz total en la salida y la luz total en la entrada.

Aunque esta medida no es de tanta importancia en la correlación digital, si lo es en la correlación óptica. Se utiliza para asegurar que el plano de correlación gane la mayor parte de la luz de entrada y que los detectores en el plano puedan responder de manera más rápida y precisa (Vijaya-Kumar y Hassebrook, 1990). Debido a que $H(u, v)$ es siempre menor o igual a 1 en los sistemas ópticos, es fácil ver que η_H también es menor o igual a 1. Para el caso del filtro convencional *POF*, cuya función de transmitancia tiene un módulo $|H(u, v)| = 1$, se consigue el valor máximo para η_H , es decir, $\eta_H = 1$.

El filtro sólo de fase o *POF* se expresa por:

$$H(u, v) = \frac{T^*(u, v)}{|T(u, v)|} = e^{-j\phi(u, v)} \quad (68)$$

donde $\phi(u, v)$ es la fase de la transformada de Fourier del objetivo.

En particular, el filtro sólo de fase es, de entre todos los filtros con módulo unitario, el óptimo en la relación señal-ruido (*SNR*) y en la relación pico-energía de correlación

(PCE^1).

Cercanamente relacionado con la PCE se encuentra la capacidad de discriminación (DC^2), definida como la habilidad de distinguir entre patrones similares, la cual es consecuentemente superior en un POF que un CMF . Esto se puede apreciar intuitivamente si consideramos el filtro sólo de fase como una versión de realce de frecuencias de un filtro de acoplamiento. Es fácil ver que el POF es un CMF dividido por su módulo. El módulo de las imágenes comunes concentra la mayoría de su energía en las bajas frecuencias y de esta manera el denominador en el POF atenúa las bajas frecuencias y realza las altas. Finalmente, como las componentes de las altas frecuencias de la transformada de Fourier están relacionadas con los bordes y la forma de un objeto, es decir con sus elementos característicos, el POF llega a ser más sensible a las diferencias entre imágenes, (ver Fig. 6).

De cualquier modo, la relación señal-ruido obtenida con filtros sólo de fase es pequeña debido a su naturaleza pasa todo.

IV.1.3. Optimización de la estrechura del pico: filtro inverso

Otro atributo deseable de un filtro es que pueda producir picos de correlación lo más agudo posible. Esto se puede caracterizar por medio de la relación pico-energía de correlación (PCE). Aunque el POF proporciona un pico más estrecho que el CMF , la

¹PCE: Peak-to-Correlation Energy

²DC: Discrimination Capability

estrechez del pico medida por la relación PCE , es optimizada por el filtro inverso (IF):

$$\begin{aligned}
PCE_{H(u,v)} &= \alpha \frac{\left| \iint T(u,v)H(u,v)dudv \right|^2}{\iint |T(u,v)|^2 |H(u,v)|^2 dudv} \\
&= \alpha \frac{\left| \iint \left[\frac{T(u,v)}{|T(u,v)|} \right] [|T(u,v)| H(u,v)] dudv \right|^2}{\iint |T(u,v)|^2 |H(u,v)|^2 dudv} \\
&\leq \alpha \frac{\iint |T(u,v)|^2 |H(u,v)|^2 dudv \iint dudv}{\iint |T(u,v)|^2 |H(u,v)|^2 dudv} \\
&= \alpha \iint dudv = 1
\end{aligned} \tag{69}$$

donde los límites de integración van de $-\infty$ a ∞ y se consigue la igualdad *si y solo si*

$$H(u,v) = \beta \frac{T^*(u,v)}{|T(u,v)|^2} \equiv H_{IF}(u,v), \tag{70}$$

donde β es una constante arbitraria.

El resultado anterior indica que el filtro inverso proporciona una razón entre el pico y la energía de correlación igual a uno, y consecuentemente, concentra toda la energía que pasa a través del filtro en el pico de correlación. Esto se puede explicar fácilmente en términos físicos al encontrar en

$$T(u,v)H(u,v) = T(u,v)\beta \frac{T^*(u,v)}{|T(u,v)|^2} = \beta, \tag{71}$$

es decir, β indica la magnitud del impulso, en este caso la magnitud del pico de correlación.

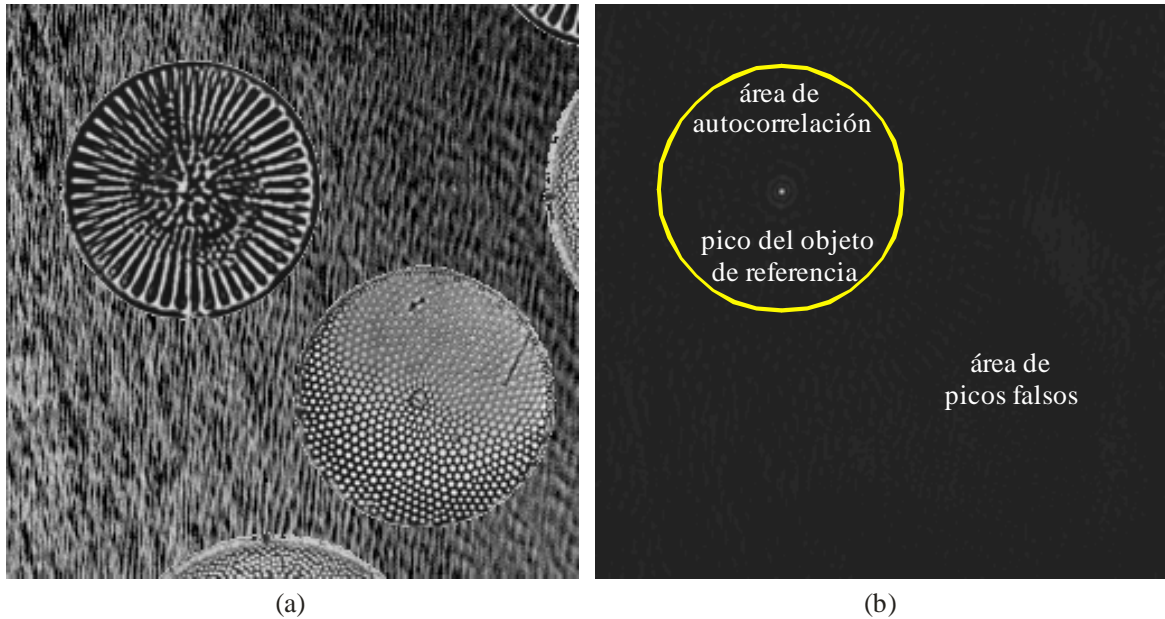


Figura 5: Cálculo de la DC: (a) ejemplo de escena de entrada, (b) salida de correlación. Para calcular la DC según la ecuación (72), C^T es el máximo dentro del área del objeto y C^B es el máximo dentro del área de picos falsos.

IV.1.4. Capacidad de discriminación

La capacidad de discriminación o DC , es uno de los criterios de desempeño más importantes en el reconocimiento de patrones ya que proporciona un valor que indica que tan bueno es un filtro para detectar y discriminar entre diferentes clases de objetos. Si un objeto de referencia se encuentra incrustado en un fondo, el cual contiene objetos falsos, entonces la DC se puede expresar como:

$$DC = 1 - \frac{|C^B(0,0)|^2}{|C^T(0,0)|^2}, \quad (72)$$

donde C^B es el máximo en el plano de correlación sobre el área del fondo a rechazar (u objetos no deseados) y C^T es el máximo en el plano de correlación sobre el área en

que se encuentra el objeto de referencia. Tal como se indica en la Fig. 5, el área de la posición del objeto de referencia se determina en la vecindad cercana a la posición actual del objeto. El área del fondo es complementaria al área donde se encuentra el objeto de referencia. Es preciso aclarar que el máximo valor que se puede obtener con la DC es la unidad, mientras que valores menores a cero indican que el filtro falló en el reconocimiento. Analizando la ecuación (72) se puede observar que un valor negativo se obtiene sólo si la intensidad del pico de correlación cruzada $|C^B(0,0)|^2$, es mayor que la intensidad del pico de autocorrelación $|C^T(0,0)|^2$.

Los filtros adaptativos que se estudiarán en los siguientes capítulos, representan la aportación del ejercicio investigativo doctoral y se han generado al optimizar este criterio de calidad.

IV.2. Algoritmo para reconocer objetos por medio de la correlación

Antes de continuar con la teoría del diseño de filtros de correlación, se explicará brevemente como se lleva a cabo el reconocimiento de objetos por medio de estos filtros. En la Fig. 6(a) se presenta un escenario de detección libre de ruido. Se presentan los negativos de las imágenes originales para una mejor apreciación de las respuestas al impulso de los filtros. El problema de reconocimiento consiste en detectar a la mariposa inferior dentro de esa escena. En los incisos del b-d se muestran las respuestas al impulso correspondientes a los filtros CMF , POF e IF , respectivamente y las cuales se han desarrollado conforme a los algoritmos descritos con anterioridad. En los incisos del e-g se muestran las intensidades de las salidas de correlación en 3-D. Como se podrá apreciar, los picos en la salida del filtro CMF son de mayor extensión que aquellos que generan los filtros POF e IF . Asimismo, los picos de correlación resultantes de aplicar el filtro

inverso son los más estrechos.

Una vez que se tiene el filtro $H(u, v)$, el algoritmo que se sigue para determinar si un objeto se encuentra presente o no dentro de una escena dada, es el siguiente:

1. Calcular la transformada de Fourier de la escena de entrada $s(x, y)$, para obtener $S(u, v) = \mathcal{F}[s(x, y)]$.
2. Efectuar la operación punto a punto del producto de $S(u, v)$ por la función filtro $H(u, v)$.
3. Obtener el plano de correlación al aplicar la *IFT* al resultado en (2), es decir $c(x, y) = \mathcal{F}^{-1}[S(u, v)H^*(u, v)]$.
4. Localizar los picos en la salida de la correlación $c(x, y)$. Recuérdese que la posición de los picos indican la posición de los objetos.
5. Si el pico correspondiente al objeto de referencia es mayor que un umbral ϵ , entonces el filtro tuvo éxito, de lo contrario falló el reconocimiento.

IV.3. Las invariancias en el reconocimiento de patrones

Aunque cada filtro optimiza algún criterio de calidad, las salidas se ven rápidamente afectadas cuando los objetos presentan distorsiones geométricas. Dentro del reconocimiento de patrones existe un gran interés en el diseño de sistemas invariantes, esto es, sistemas que sean capaces de detectar la presencia de un objeto de referencia en la escena de entrada, a pesar de que éste se encuentre desplazado lateralmente, con distinto tamaño, girado o incluso con distinta luminosidad.

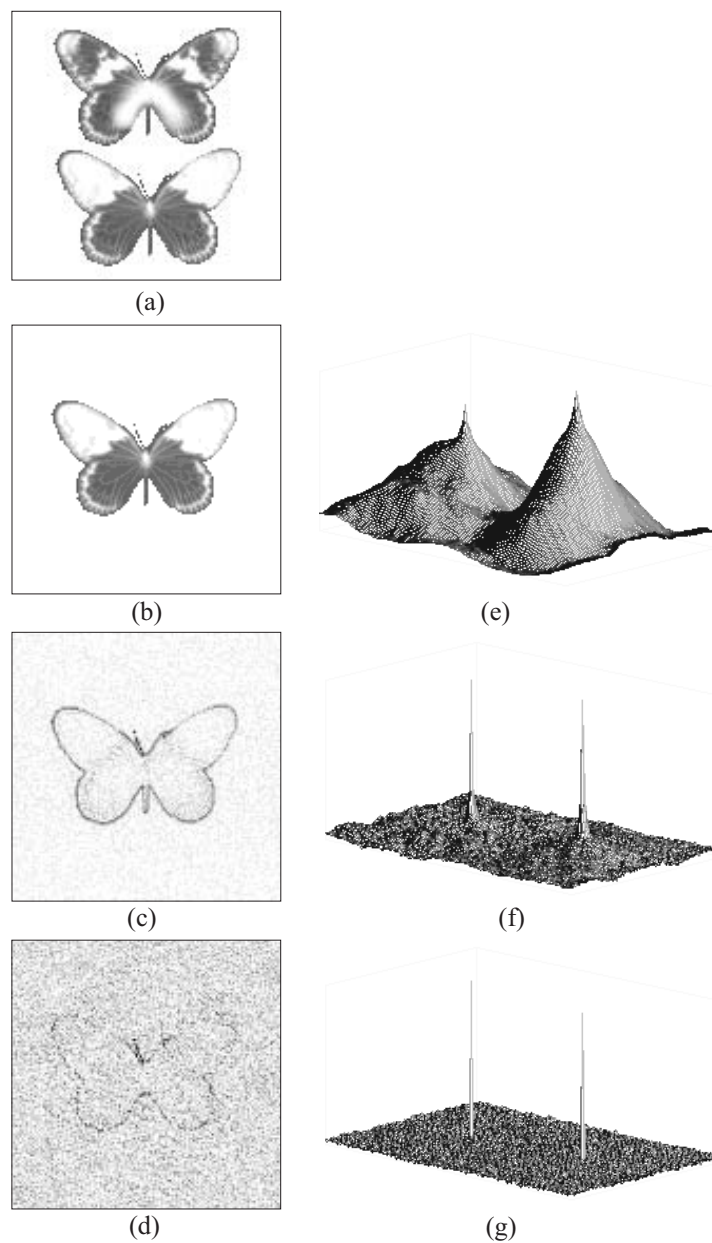


Figura 6: Ejemplo de localización en escenas libres de ruido. (a) Escena de entrada, el objetivo es la mariposa inferior. Respuestas al impulso de los filtros: (b) *CMF*, (c) *POF* y (d) *IF*. Planos de correlación en 3-D proporcionados por: (e) *CMF*, (f) *POF* y (g) *IF*.

La invariancia a los desplazamientos laterales del objeto de referencia sobre la escena de entrada implícitamente se encuentra presente en los filtros de correlación descritos anteriormente. Pero para añadir otro tipo de invariancias hay que recurrir a otros métodos. Existen dos mecanismos para incluir invariancias en los filtros. El primero consiste en usar transformaciones geométricas por medio de formulaciones matemáticas como la transformada de Mellin que se emplea para añadir la invariancia al escalamiento o la expansión en armónicos circulares para añadir invariancia a las rotaciones. Sin embargo, no todas las distorsiones geométricas se pueden modelar matemáticamente, de aquí que el segundo mecanismo consiste en emplear un conjunto de imágenes representativas de las posibles distorsiones que pudiera presentar un objeto para formar un filtro compuesto.

A continuación se expone brevemente en que consisten las invariancias por medio de formulaciones matemáticas.

IV.3.1. Invariancia a escala

Podría pensarse que un cambio de escala equivale a una traslación radial en coordenadas polares, de la misma forma que una rotación es una traslación a lo largo de una coordenada angular. Sin embargo la rotación es una operación lineal, esto es, dos sucesivas rotaciones son equivalentes a una rotación de ángulo igual a la suma de los ángulos de las rotaciones sucesivas, mientras que escalar no es una operación lineal (Arsenault *et al.*, 1989). Hasta la fecha el mejor método para obtener invariancia a escala incluye el uso de momentos o la descomposición en armónicos radiales de *Mellin* (Casasent y Psaltis, 1976b,c,a) o la *Transformada de escala*.

IV.3.2. Invariancia a la rotación

Matemáticamente, rotar un objeto equivale a un desplazamiento a lo largo de la coordenada angular en una geometría polar. Esto sugiere que la rotación se puede añadir a un filtro de acoplamiento por medio de una transformación de coordenadas, tal como se realiza en Villalobos-Flores *et al.* (2001) y en Castro-Longoria *et al.* (2001).

Otra forma de adicionar la invariancia, es buscar alguna parte del objeto que sea invariante a cualquier rotación, esta es una idea que sugiere una descomposición en coordenadas polares y nos lleva a la descomposición de armónicos circulares (Arsenault *et al.*, 1989). La clave para incorporar invariancias a rotaciones a los filtros de acoplamiento, es utilizar un sólo armónico circular. Como la correlación es una operación lineal, la correlación de una componente con todo el objeto, equivale a la suma de las correlaciones individuales de dicha componente con todas las demás. La correlación de la componente seleccionada con cualquier otra componente da una salida de amplitud baja, pero la correlación de una componente con si misma, da un valor absoluto alto que es también invariante a la rotación. La rotación solo cambia la fase en el pico de correlación, pero no su intensidad. La razón por la que la correlación típica no es invariante a la rotación se debe a que el filtro de acoplamiento utiliza todo el objeto, es decir, todos sus componentes armónicos circulares.

El desempeño de un filtro que emplee la expansión en armónicos circulares se ve afectado por la selección del centro de expansión del objeto (también llamado centro propio o de desarrollo). Existen varios métodos para encontrar el centro propio correcto del objeto (Hsu y Arsenault, 1982; Garcia-Martinez *et al.*, 1995).

IV.4. Derivación de filtros compuestos

Tal como se mencionó anteriormente, los filtros de acoplamiento sufren de severa sensibilidad a las distorsiones de la señal de entrada. Puesto que las distorsiones como los cambios de escala, rotación y diferencia en iluminación están siempre presentes en una aplicación real de reconocimiento de patrones, el filtro $H(u, v)$ se debe diseñar para producir valores similares de salida tanto con imágenes distorsionadas como con las que no lo están. Las funciones discriminantes sintéticas ($SDF's$ ³) representan un método alternativo para el diseño de filtros invariantes a tales distorsiones (Vijaya-Kumar, 1992a).

IV.4.1. Funciones discriminantes sintéticas

El desempeño de los filtros de correlación decrece rápidamente cuando los objetos presentan distorsiones, como rotaciones, cambios de escala, deformaciones, etc. En la tarea del reconocimiento de objetos, una de las técnicas para añadir invariancias a las distorsiones está basada en el uso de las funciones discriminantes sintéticas (SDF), estos filtros emplean un conjunto de imágenes de entrenamiento (patrones) que son lo suficientemente representativas de las posibles distorsiones que un objeto pudiera presentar.

Básicamente, un filtro SDF es una combinación lineal de filtros de acoplamiento para diferentes patrones (Casasent, 1984).

Los coeficientes en la combinación lineal del filtro SDF deben satisfacer un conjunto de condiciones en la salida del filtro. La principal desventaja del filtro SDF es la aparición de picos falsos que pertenecen a otros objetos en la salida del filtro, dando como resultado que el filtro tenga una capacidad de discriminación pobre.

³SDF: Synthetic Discriminant Function

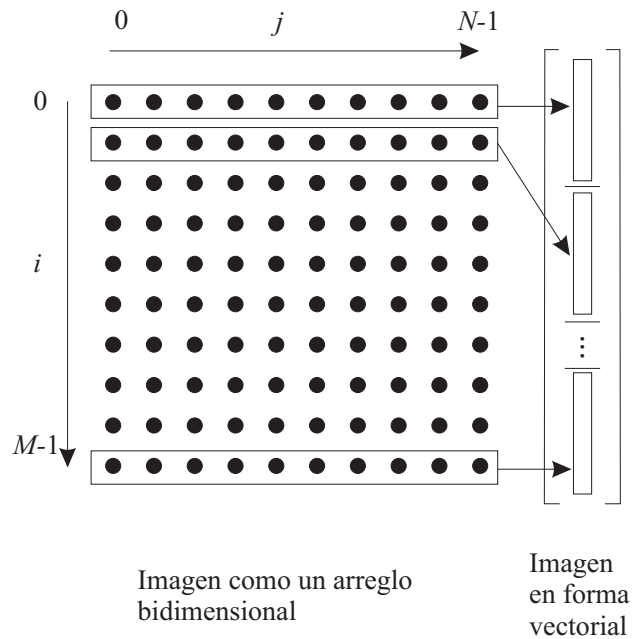


Figura 7: Una imagen digital se puede ver como un arreglo bidimensional de M renglones y N columnas, donde cada elemento se le llama pixel. Al ordenar lexicográficamente los renglones se obtiene un vector imagen (columna) con $M \times N$ elementos.

Problema de reconocimiento intraclase

Sea $\{t_i(x, y); i = 1, 2, \dots, N\}$ un conjunto de imágenes de entrenamiento (linealmente independientes), cada una con d pixeles. La respuesta al impulso del filtro SDF se puede expresar como:

$$h(x, y) = \sum_{i=1}^N a_i t_i(x, y), \quad (73)$$

donde $\{a_i; i = 1, 2, \dots, N\}$ son los coeficientes que deben satisfacer las condiciones en el origen:

$$t_i \circ h = u_i, \quad (74)$$

donde $\{u_i; i = 1, 2, \dots, N\}$ son los valores preespecificados para el origen en la salida de la correlación de cada imagen de entrenamiento.

Sea la matriz \mathbf{R} con N columnas y d renglones, donde la i -ésima columna es la imagen $t_i(x, y)$ en su forma vector (ver Fig. 7). Sean \mathbf{a} y \mathbf{u} los vectores columna de $\{a_i\}$ y $\{u_i\}$, respectivamente. Las ecuaciones (73) y (74) se pueden escribir en forma matriz-vector (usando vector imagen (columna) con $M \times N$ elementos.):

$$\mathbf{h} = \mathbf{R}\mathbf{a}, \quad (75)$$

$$\mathbf{u} = \mathbf{R}^T \mathbf{h}, \quad (76)$$

donde \mathbf{R}^T denota a la matriz transpuesta de \mathbf{R} .

Al sustituir la ecuación (75) en la ecuación (76) se obtiene:

$$\mathbf{u} = (\mathbf{R}^T \mathbf{R})\mathbf{a}. \quad (77)$$

El (i, j) -ésimo elemento de la matriz $\mathbf{S} = (\mathbf{R}^T \mathbf{R})$ es el valor en el origen de las correlaciones cruzadas entre las imágenes de entrenamiento $t_i(x, y)$ y $t_j(x, y)$. Si la matriz \mathbf{S} es no singular, la solución al sistema de ecuaciones esta dado por

$$\mathbf{a} = (\mathbf{R}^T \mathbf{R})^{-1} \mathbf{u}, \quad (78)$$

y el filtro en forma vectorial es

$$\mathbf{h}_{SDF} = \mathbf{R}(\mathbf{R}^T \mathbf{R})^{-1} \mathbf{u}. \quad (79)$$

Al fijar todos los elementos de \mathbf{u} con un mismo valor, el filtro SDF se puede usar para detectar patrones distorsionados que pertenecen a una misma clase de objetos. En las implementaciones que se realizaron se fijaron los elementos de \mathbf{u} en la unidad, es decir

$$\mathbf{u} = [\mathbf{11}\dots\mathbf{1}]^T. \quad (80)$$

Esta configuración del filtro es conocida como el filtro con igual salida de correlación (*ECP-SDF* por sus siglas en inglés, (Vijaya-Kumar, 1992b)).

Problema de reconocimiento multiclase

Supóngase que se tienen diferentes versiones distorsionadas del objeto de referencia y varias clases de objetos que se desean rechazar. Por simplicidad se considera el problema de reconocimiento entre dos clases de objetos. De aquí que, se desea diseñar un filtro para reconocer las imágenes de entrenamiento de una clase, denominada la clase verdadera, y rechazar las imágenes de entrenamiento de otra clase, denominada la clase falsa.

Supóngase además, que tenemos M imágenes de entrenamiento de la clase falsa es decir $\{p_i(x, y); i = 1, 2, \dots, M\}$. De acuerdo al método del filtro *SDF*, la imagen compuesta $h(x, y)$ es una combinación lineal de todas las imágenes de entrenamiento $\{t_1(x, y), \dots, t_N(x, y), p_1(x, y), \dots, p_M(x, y)\}$. Tanto el problema de reconocimiento intra-clase como el multiclase (es decir, la discriminación entre los objetos de la clase verdadera contra los objetos de la clase falsa) pueden resolverse por medio de los filtros *SDF*. Entonces, podemos definir la salida del filtro como $\{u_i = 1; i = 1, 2, \dots, N\}$ para los objetos de la clase verdadera. y $\{u_i = 0, i = N + 1, N + 2, \dots, N + M\}$ para los objetos de la clase falsa, es decir

$$\mathbf{u} = [\mathbf{11}\dots\mathbf{100}\dots\mathbf{0}]^T. \quad (81)$$

usando el filtro dado por la ecuación (79) para el reconocimiento de patrones, se espera que el pico central de correlación tenga un valor cercano a uno para los objetos no entrenados de la clase verdadera y un valor cercano a cero para los objetos correspondientes a la clase falsa. Evidentemente, el método descrito se puede extender para discriminar cualquier número de clases. Es preciso hacer notar, que este sencillo procedimiento, para generar filtros, es una deficiencia para tomar el control absoluto sobre plano de correlación. Lo que conduce a que picos falsos puedan aparecer en cualquier lugar sobre

el plano de correlación, puesto que el valor predeterminado en la salida de la correlación solo puede controlar un solo punto (pico).

Una vez que se obtiene el filtro vector \mathbf{h}_{SDF} , es necesario llevarlo a la forma matricial en $2-D$ para obtener $h(x, y)$. Una vez que se tiene el filtro, el reconocimiento de objetos se realiza mediante la siguiente operación punto a punto:

$$c(x, y) = \mathcal{F}^{-1} \{ \mathcal{F} [s(x, y)] \cdot \mathcal{F} [h(x, y)] \}, \quad (82)$$

donde \mathcal{F} y \mathcal{F}^{-1} denotan al par de Transformadas de Fourier, $s(x, y)$ es la escena de prueba, $c(x, y)$ la salida de la correlación y (x, y) son las coordenadas en el dominio espacial.

IV.4.2. SDF de mínima varianza (MVSDF)

Supóngase que la imagen de entrada es una de las imágenes de entrenamiento alterada por ruido aditivo de media cero, el cual se representa por el vector \mathbf{n} . Entonces el filtro SDF produce una salida de correlación en el origen que está formada por dos partes. La primera parte debido a la imagen de entrenamiento es c_i . La segunda, debida al ruido n en la entrada, es una variable aleatoria con media cero y varianza σ^2 :

$$\begin{aligned} \sigma^2 &= E\{(\mathbf{h}^T \mathbf{n})^2\} \\ &= E\{\mathbf{h}^T \mathbf{n} \mathbf{n}^T \mathbf{h}\} \\ &= \mathbf{h}^T \mathbf{C} \mathbf{h}, \end{aligned} \quad (83)$$

donde \mathbf{C} es la matriz de covarianza para el ruido \mathbf{n} . Se desea minimizar esta varianza de salida. El SDF de mínima varianza ($MVSDF$) está diseñado para minimizar σ^2 en (83) y al mismo tiempo satisfacer las condiciones en la ecuación (74). Por lo tanto, la

solución dada por Vijaya-Kumar (1986) es:

$$\mathbf{h}_{MV SDF} = \mathbf{C}^{-1} \mathbf{R} (\mathbf{R}^T \mathbf{C}^{-1} \mathbf{R})^{-1} \mathbf{u}. \quad (84)$$

De nueva cuenta, $\mathbf{h}_{MV SDF}$ en la ecuación (84) satisface las condiciones en (74). En general, *MV SDF* no ha sido muy popular por dos razones principales. La primera es que invertir la matriz de covarianza \mathbf{C} con $d \times d$ elementos es difícil en general, excepto en pocos casos especiales. La segunda razón es que el *MV SDF* no controla completamente el plano de correlación, lo que lleva a un problema de grandes lóbulos laterales (picos fuera del área de localización de los objetos). Este problema se aligera usando el filtro del mínimo promedio de la energía de correlación (*MACE*⁴), el cual se discute a continuación.

IV.4.3. Filtro MACE (filtro del mínimo promedio de la energía de correlación)

Una de las maneras de superar el problema de grandes lóbulos laterales en los filtros *SDF*'s es minimizar los lóbulos laterales y al mismo tiempo satisfacer las condiciones de la ecuación (74). Sin embargo, no es posible expresar el nivel de los lóbulos como una función del filtro $H(u, v)$ utilizando una expresión aproximada y por otro lado una optimización tampoco es posible. Una alternativa más tratable es minimizar la correlación promedio del plano de energía E mientras se cumplen las condiciones impuestas en la ecuación (74).

Descripción del filtro de la mínima energía promedio de correlación

Se desea diseñar un filtro de correlación que garantice picos de correlación agudos, que mantenga las condiciones impuestas en los valores de los picos y mantenga la

⁴MACE: Minimum Average Correlation Energy

invarianza en la posición. Para llevar a cabo una buena detección es necesario reducir los niveles de la función de correlación en todos los puntos excepto en el origen del plano de correlación. Específicamente, el valor de la correlación en el origen se debe definir. De tal manera que la definición del problema es: *minimizar la energía de la función de correlación y satisfacer las condiciones en el origen.*

Primero se establecen las condiciones. Se denota a la función de correlación de la i -ésima imagen de la secuencia $t_i(x)$ con el filtro $h(x)$ por:

$$g_i(x) = t_i(x) \circ h(x), \quad (85)$$

y se denota la transformada de Fourier discreta de la función de correlación por $G_i(k)$. La función de transferencia se denota por $H(k)$ y la transformada de Fourier de la imagen $t_i(x)$ por $R_i(k)$.

Utilizando notación vectorial, la condición en la amplitud del pico de la correlación es:

$$g_i(0) = \hat{\mathbf{t}}_i^+ \hat{\mathbf{h}} = u_i, \quad (86)$$

para toda imagen de entrenamiento $\{t_i; i = 1, 2, \dots, N\}$, y donde u_i es el valor especificado por el usuario en la i -ésima función de correlación en el origen y también el i -ésimo elemento del vector de control \mathbf{u} . El símbolo $+$ denota al complejo conjugado transpuesto. El vector columna $\hat{\mathbf{t}}_i$ representa la i -ésima imagen de entrenamiento en el dominio de Fourier, la cual se obtiene al ordenar lexicográficamente los renglones de la imagen en el dominio de la frecuencia (como se indicó en la Fig. 7) y $\hat{\mathbf{h}}$ es el filtro en forma vectorial en el dominio de la frecuencia.

Entonces tenemos que para toda i , este filtro también debe minimizar la energía del

plano de correlación:

$$\begin{aligned}
 E_i &= \sum_{i=1}^d |g_i(x)|^2 \\
 &= \frac{1}{d} \sum_{i=1}^d |G_i(k)|^2 \\
 &= \frac{1}{d} \sum_{i=1}^d |H(k)|^2 |R_i(k)|^2,
 \end{aligned} \tag{87}$$

qdonde d indica el número de pixeles dentro de la imagen. Utilizando la forma vectorial, la energía del plano se puede escribir como

$$E_i = \hat{\mathbf{h}}^+ \hat{\mathbf{D}}_i \hat{\mathbf{h}}, \tag{88}$$

donde el superíndice $+$ indica el transpuesto conjugado de un vector complejo y $\hat{\mathbf{D}}_i$ es una matriz diagonal de tamaño $d \times d$ cuyos elementos diagonales son el cuadrado de la magnitud de los elementos de $\hat{\mathbf{t}}_i$ asociados, es decir,

$$\hat{\mathbf{D}}_i(k, k) = |\hat{\mathbf{t}}_i|^2, \tag{89}$$

Entonces el problema es encontrar el vector $\hat{\mathbf{h}}$ que minimice $\hat{\mathbf{h}}^+ \hat{\mathbf{D}}_i \hat{\mathbf{h}}$ para todo i y que al mismo tiempo cumpla con las condiciones impuestas en el pico de correlación. Estas condiciones se pueden escribir para todas las imágenes de entrenamiento en forma vectorial como:

$$\hat{\mathbf{R}}^+ \hat{\mathbf{h}} = \mathbf{u} \tag{90}$$

donde la matriz $\hat{\mathbf{R}}$ se compone por los vectores columna $\hat{\mathbf{t}}_i$. La solución a este problema no existe debido a que no es posible minimizar las condiciones de todas las E_i

simultáneamente. Sin embargo, es posible minimizar el valor promedio de E_i , es decir

$$\begin{aligned}
 E_{AV} &= \frac{1}{N} \sum_{i=1}^N E_i \\
 &= \frac{1}{N} \sum_{i=1}^N \hat{\mathbf{h}}^+ \hat{\mathbf{D}}_i \hat{\mathbf{h}} \\
 &= \frac{1}{N} \hat{\mathbf{h}}^+ \left[\sum_{i=1}^N \hat{\mathbf{D}}_i \right] \hat{\mathbf{h}} \\
 &= \frac{1}{N} \hat{\mathbf{h}}^+ \hat{\mathbf{D}} \hat{\mathbf{h}}, \tag{91}
 \end{aligned}$$

esta es la razón por la que se le denomina *filtro de la mínima energía promedio de correlación* (Mahalanobis *et al.*, 1987).

Ahora el enunciado del problema se reduce a minimizar $\hat{\mathbf{h}}^+ \hat{\mathbf{D}} \hat{\mathbf{h}}$ sujeto a la condición lineal $\hat{\mathbf{R}}^+ \hat{\mathbf{h}} = \mathbf{u}$. La solución a este problema se puede encontrar utilizando el método de multiplicadores de Lagrange. Este método reduce el problema restringido de N variables en uno sin restricciones de $N + 1$ variables cuyas ecuaciones pueden ser resueltas.

La función que se desea minimizar se forma por:

$$\Delta = \hat{\mathbf{h}}^+ \hat{\mathbf{D}} \hat{\mathbf{h}} - 2\lambda_1(\hat{\mathbf{h}}^+ \hat{\mathbf{t}}_1 - u_1) - \dots - 2\lambda_N(\hat{\mathbf{h}}^+ \hat{\mathbf{t}}_N - u_N), \tag{92}$$

donde $\lambda_1, \dots, \lambda_N$ son los parámetros que nos permiten satisfacer la condición de minimización y se han multiplicado por el factor 2. Ya que se tiene una ecuación cuadrática, al calcular el gradiente de Δ con respecto de $\hat{\mathbf{h}}$ se obtiene:

$$\Delta_{\hat{\mathbf{h}}} = 2\hat{\mathbf{D}}\hat{\mathbf{h}} - 2\lambda_1\hat{\mathbf{t}}_1 - \dots - 2\lambda_N\hat{\mathbf{t}}_N \tag{93}$$

Al igualar la expresión a $\mathbf{0}$ (el vector cero) se puede escribir

$$\Delta_{\hat{\mathbf{h}}} = \hat{\mathbf{D}}\hat{\mathbf{h}} - \lambda_1\hat{\mathbf{t}}_1 - \dots - \lambda_N\hat{\mathbf{t}}_N, \tag{94}$$

donde se puede ver que el vector $\hat{\mathbf{h}}$ que minimiza (92) debe satisfacer

$$\hat{\mathbf{D}}\hat{\mathbf{h}} = \lambda_1\hat{\mathbf{t}}_1 + \dots + \lambda_N\hat{\mathbf{t}}_N \quad (95)$$

$$= \sum_{i=1}^N \lambda_i \hat{\mathbf{t}}_i, \quad (96)$$

de aquí que los coeficientes λ_i se deben seleccionar de tal forma que cumplen con las condiciones impuestas en la ecuación (90). Cuando la matriz $\hat{\mathbf{D}}$ es invertible se puede reescribir $\hat{\mathbf{h}}$ de la forma

$$\hat{\mathbf{h}} = \hat{\mathbf{D}}^{-1} \left[\sum_{i=1}^N \lambda_i \hat{\mathbf{x}}_i \right] \quad (97)$$

$$= \sum_{i=1}^N \lambda_i \hat{\mathbf{D}}^{-1} \hat{\mathbf{x}}_i. \quad (98)$$

En términos del vector $\mathbf{L} = [\lambda_1, \lambda_2, \dots, \lambda_N]^T$,

$$\hat{\mathbf{h}} = \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}} \mathbf{L}. \quad (99)$$

Sustituyendo (99) en la ecuación (90), la ecuación condicional se convierte en

$$\hat{\mathbf{R}}^+ \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}} \mathbf{L} = \mathbf{u}, \quad (100)$$

la cual se resuelve para \mathbf{L} como

$$\mathbf{L} = (\hat{\mathbf{R}}^+ \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}})^{-1} \mathbf{u}. \quad (101)$$

Finalmente, al sustituir \mathbf{L} de la ecuación (101) en la ecuación (99), se obtiene la expresión para $\hat{\mathbf{h}}$ como

$$\hat{\mathbf{h}}_{MACE} = \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}} (\hat{\mathbf{R}}^+ \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}})^{-1} \mathbf{u}. \quad (102)$$

El vector $\hat{\mathbf{h}}$ en la ecuación (102) satisface simultáneamente $\hat{\mathbf{R}}^+\hat{\mathbf{h}} = \mathbf{u}$, y minimiza $\hat{\mathbf{h}}^+\hat{\mathbf{D}}\hat{\mathbf{h}}$. Como $\hat{\mathbf{D}}$ es una matriz diagonal, su inversión es trivial, entonces la única matriz que se necesita invertir es $\hat{\mathbf{R}}^+\hat{\mathbf{D}}^{-1}\hat{\mathbf{R}}$. Esta facilidad para calcular la correlación con lóbulos laterales bajos han hecho al filtro *MACE* muy popular.

En la construcción del filtro *MACE* en la ecuación (102) no se considera la tolerancia al ruido. Una manera de añadirle tolerancia al ruido es intercambiar la energía E y la varianza σ^2 satisfaciendo al mismo tiempo las condiciones en la ecuación (90). Una manera alternativa es alterar el contenido de la matriz diagonal \mathbf{D} de tal forma que \mathbf{D}^{-1} no contenga grandes valores de frecuencias donde la señal es baja y el ruido es alto. Una característica de las salidas de correlación de los filtros *MACE* es su naturaleza aguda. Esta propiedad conduce a realzar las altas frecuencias en el filtro, causando graves distorsiones y sensibilidad al ruido.

IV.5. Resumen

Hasta este momento se tiene una idea clara de que existen diferentes tipos de filtros y que cada uno de ellos se ha diseñado para mejorar alguna condición en el reconocimiento de objetos en imágenes. Se vió que al optimizar algunas medidas de desempeño también se pueden derivar algunos filtros como el *CMF* resultado de maximizar la *SNR*, el *POF* resultado de maximizar la eficiencia de Horner y el filtro inverso que se deriva al maximizar la *PCE*. El filtro inverso aunque proporciona un pico estrecho es de poco valor práctico ya que amplifica las altas frecuencias (cambios bruscos en la escala de grises dentro de una imagen), y por lo tanto el ruido de entrada.

En este capítulo se dió a conocer la familia de filtros *SDF* los cuales requieren de un conjunto de imágenes de entrenamiento para ser generados, en esta familia se encuentra el filtro *SDF* convencional, el filtro *MACE* y el filtro *MVSDF*. Estos filtros compuestos

surgen de la necesidad de incorporar invariancias en el reconocimiento de patrones, sobre todo cuando es complicado describir matemáticamente las posibles distorsiones que pudiera presentar un objeto (por ejemplo, distorsiones geométricas en 3-D, oclusiones, cambios de iluminación, etc.). Además son de gran utilidad cuando se quiere discernir entre diversas clases de objetos.

Capítulo V

Filtros adaptativos para el reconocimiento de patrones

Desde la introducción del filtro de acoplamiento clásico (VanderLugt, 1964), se han propuesto diferentes tipos de filtros basados en la correlación para el reconocimiento de patrones. La manera tradicional de diseñar un filtro de correlación consiste en optimizar algún criterio de calidad: eficiencia de luz, relación señal-ruido, etc. En Vijaya-Kumar y Hassebrook (1990) se sintetizan las diferentes medidas de desempeño que se han propuesto para medir la calidad de los filtros de correlación. Algunas medidas de desempeño, en esencia, se pueden mejorar utilizando un enfoque adaptativo durante el diseño del filtro. Como el lector podrá intuir, el enfoque adaptativo pretende construir un filtro que se adapte a una escena en particular. De acuerdo a este concepto, nuestra atención se centra en el diseño de un filtro cuya característica de desempeño sea buena para una escena dada, es decir, con un conjunto de patrones fijos o un fondo fijo a rechazar, en lugar de un filtro que se haya diseñado con parámetros de desempeño promedio sobre un conjunto de imágenes.

Como se mencionó con anterioridad, la capacidad de discriminación o DC es uno de los criterios de desempeño más importantes en el reconocimiento de patrones ya que indica qué tan bueno es un filtro para discriminar entre diferentes clases de objetos. Yaroslavsky (1993) efectuó un análisis teórico de métodos por correlación, y sugirió un filtro de correlación con una probabilidad mínima de errores de localización anómalos

(falsas alarmas) al que llamó el filtro óptimo (OF^1). Una característica importante del filtro OF es su adaptabilidad a la escena en diferentes aplicaciones de reconocimiento de patrones o detección de objetos, ya que en su respuesta en frecuencia se considera el espectro de energía de los objetos no deseados en la escena observada o bien el fondo a rechazar. La desventaja del filtro OF en una implementación óptica, es su extremada baja eficiencia de luz. El filtro con la máxima eficiencia de luz es el filtro sólo de fase (POF) propuesto por Horner y Gianino (1984). La desventaja que presenta el filtro POF es una pobre capacidad de discriminación cuando el objeto de referencia tiene un bajo contraste y se encuentra incrustado en una escena con un fondo complicado (Yaroslavsky, 1993). En la literatura se pueden encontrar diferentes esfuerzos por mejorar las características del OF , por ejemplo, en Kober *et al.* (1994) se presenta una aproximación del filtro OF por medio de filtros sólo de fase con cuantización, es decir, se diseñó un filtro con una eficiencia de luz alta y una capacidad de discriminación muy cercana a la del OF . Cuando el objeto de referencia se encuentra incrustado en un fondo tipo ruido (generado como un proceso estocástico), el diseño del filtro óptimo también se puede obtener como en Javidi y Wang (1994). Otro fructífero método para sintetizar filtros adaptativos, es por medio de un enmascaramiento de ceros de las componentes espectrales del filtro de correlación, lo que resulta en una mejora de la capacidad para discriminar objetos similares (Kober y Campos, 1996; Ahouzi *et al.*, 1994).

Las funciones discriminantes sintéticas (SDF) constituyen un atractivo método para el reconocimiento de patrones con invariancias a las distorsiones geométricas (Casasent, 1984). Básicamente, los filtros SDF utilizan un conjunto de imágenes de entrenamiento para sintetizar una plantilla que produce salidas centrales de correlación con un valor preespecificado en respuesta a las imágenes de entrenamiento. La principal deficiencia del uso de los filtros SDF es la aparición de lóbulos laterales (picos falsos de correlación)

¹OF: Optimal Filter

que se deben a la falta de control sobre todo el plano de correlación, dando esto como resultado que el filtro *SDF* posea una baja capacidad de discriminación. Una solución parcial a este problema se sugirió por Mahalanobis *et al.* (1987), quienes proponen controlar todo el plano al minimizar la energía de correlación promedio de las imágenes de entrenamiento y de esta manera suprimir la presencia de picos extraños y al mismo tiempo proporcionar un pico bien definido para la detección fácil del objeto de referencia. Aunque estos filtros efectúan un control indirecto sobre las falsas alarmas, son más sensibles a las variaciones entre clases que otros filtros compuestos (Billet y Singher, 2002).

En este capítulo, se proponen algunos algoritmos para diseñar nuevos filtros adaptativos de correlación con una capacidad de discriminación determinada (Gonzalez-Fraga *et al.*, 2006, 2005a). Estos filtros se adaptan a una escena de entrada de prueba que se construye a partir del objeto a reconocer, de los objetos falsos, y de un fondo que se desea rechazar. Estos filtros novedosos pueden suprimir los lóbulos laterales (picos falsos) del fondo dado, así como los correspondientes a los objetos falsos. En otras palabras, los filtros sugeridos realizan un control directo sobre todo el plano de correlación. Las implementaciones digitales de los filtros están basadas en el filtro *SDF* convencional, por lo que se denominó *ASDF*². Gracias a las propiedades de los filtros compuestos, a nuestro filtro *ASDF* se le pueden incorporar invariancias, dando como resultado un filtro más robusto que puede ser utilizado en problemas del mundo real. En la sección V.2 se presentan y se discuten algunos resultados obtenidos con estos filtros en varias simulaciones. El desempeño de los filtros adaptativos en las escenas se compara con varios filtros por correlación en términos de la capacidad de discriminación y de la robustez al ruido.

²ASDF: Adaptive SDF filter

V.1. Algoritmo básico para el diseño de un filtro adaptativo ASDF

Un primer problema de reconocimiento de patrones a resolver, se puede definir de la siguiente manera: se desea un filtro que garantice un pico de correlación con un valor grande, correspondiente al objeto de referencia y que suprima los picos falsos de correlación correspondientes al fondo de la escena. En otras palabras, para conseguir un buen reconocimiento del objeto es necesario reducir los niveles de la función de correlación en todos los puntos, excepto en el origen del plano de correlación, donde la condición o restricción sobre el valor del pico se debe conocer. De aquí que, dados un objeto de referencia, objetos falsos y un fondo a rechazar, el problema anterior se puede resolver con la ayuda de un algoritmo iterativo. El algoritmo en cada iteración suprime el pico de correlación más alto correspondiente al fondo o a los objetos no deseados, y por consiguiente incrementa la capacidad de discriminación de manera monótona hasta llegar a un valor determinado. Por otro lado, recuerdése que la estimación precisa de la posición del objeto de referencia se puede llevar a cabo por medio de los filtros de correlación al localizar los picos en la salida (Kober y Campos, 1996).

Puesto que se tiene interés en un filtro que sea capaz de identificar a un objeto dentro de una escena complicada y ruidosa, y que además el filtro presente una buena capacidad de discriminación. Con la ayuda de los filtros adaptativos *SDF* se puede obtener un valor predeterminado de *DC*. En la actualidad, los filtros convencionales proporcionan desempeños pobres bajo estas condiciones. El algoritmo para el diseño del filtro requiere tener conocimiento de la imagen de fondo. De tal manera que el filtro se adapta bien al fondo conocido y conseguirá detectar al objeto de referencia (con una posición desconocida) dentro de una escena formada con el fondo. La imagen fondo se puede describir estocásticamente (ver apéndice A para detalles sobre la generación de

ruido) o bien determinísticamente, como en el caso de una fotografía. El fondo puede contener también objetos falsos con coordenadas desconocidas.

Antes de continuar, es preciso definir la región de soporte de una imagen de entrada. Matemáticamente, la función de soporte se define como

$$w(x, y) = \begin{cases} 1 & \text{al interior del objeto de referencia} \\ 0 & \text{fuera del objeto de referencia} \end{cases} \quad (103)$$

De esta manera, si la mariposa de la Fig. 8(a) es el objeto de referencia, su función de soporte es la mostrada en la Fig. 8(b), en donde la zona blanca representa valores unitarios y la zona negra valores en cero.

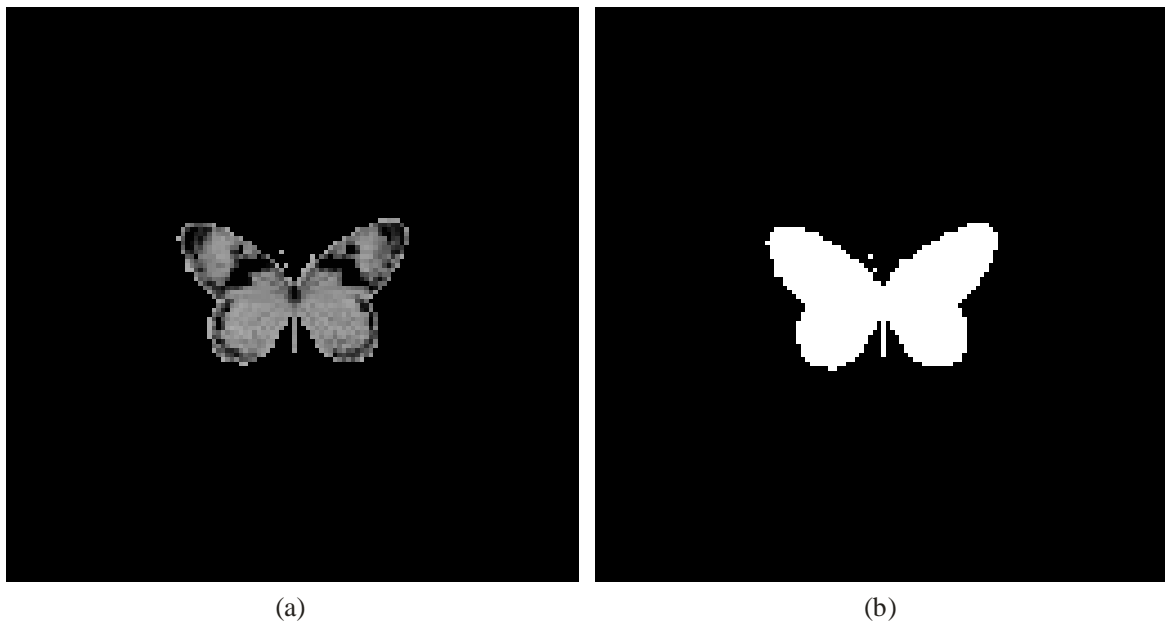


Figura 8: (a) Imagen de prueba y (b) su región de soporte.

Ahora continuemos con el diseño del algoritmo. El primer paso dentro del algoritmo

propuesto, consiste en realizar una correlación entre la imagen fondo (fotografía determinística o una realización de un proceso estocástico) y un filtro *SDF* convencional, el cual es inicialmente entrenado con el objeto de referencia. En seguida, el punto máximo en la salida de la correlación se fija como el origen para formar un objeto nuevo a partir del fondo a ser rechazado. Este objeto tiene una región de soporte igual a la del objeto de referencia, en otras palabras, ambos objetos tienen la misma forma pero diferente contenido. El objeto creado se agrega a la clase de objetos falsos. Ahora tenemos el problema de reconocimiento para dos clases como en la sección IV.4.1, y se diseña un nuevo filtro *SDF*; es decir, la clase verdadera contiene únicamente al objeto de referencia y la clase falsa consiste de los objetos falsos creados. Este procedimiento se repite de manera iterativa hasta alcanzar un valor determinado de *DC*. Finalmente, nótese que si otros objetos a ser rechazados se conocen, estos pueden incluirse directamente en la clase falsa de entrenamiento para el diseño del *filtro adaptativo SDF* (o *ASDF*). En la Fig. 9 se muestra el diagrama de bloques del procedimiento.

En forma detallada, el algoritmo básico propuesto consiste de los siguientes pasos:

1. Diseñar un filtro *SDF* convencional, el cual inicialmente está entrenado con el objeto de referencia y al que llamaremos filtro *ASDF*.
2. Calcular la correlación entre el fondo y el filtro *ASDF*.
3. Calcular la *DC* de acuerdo a la ecuación (72), pero considerando C^B como el máximo valor en la salida de la correlación con el fondo y C^T como el máximo valor en la salida de la correlación con la imagen de referencia.
4. Si el valor de la *DC* es mayor o igual al valor deseado, entonces el procedimiento del diseño termina, de lo contrario, ir al paso 5.
5. Crear un nuevo objeto de rechazo a partir del fondo. El centro del nuevo objeto

se especifica en la posición del pico de correlación más alto. Este objeto se incluye en la clase de objetos falsos.

6. Diseñar un nuevo filtro *ASDF* usando el problema de reconocimiento para dos clases que se describe en la sección IV.4.1. La clase verdadera contiene al objeto de referencia y la clase falsa consiste de los objetos falsos creados en el paso 5. Ir al paso 2.

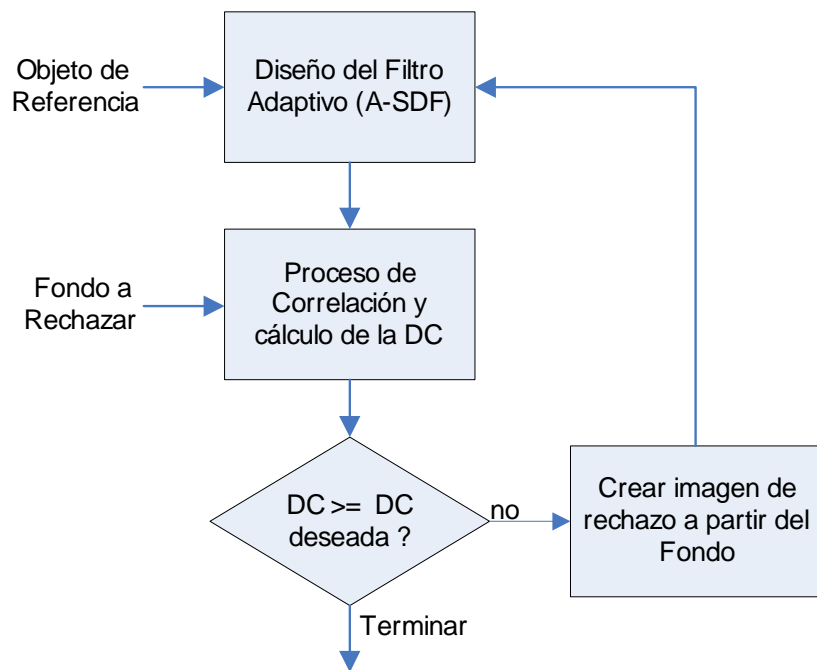


Figura 9: Diagrama de bloques del algoritmo iterativo para diseñar un filtro ASDF

En cada iteración el algoritmo selecciona de entre todos los picos falsos aquel que suprimirá en el siguiente paso para asegurar un incremento monótono en la función DC contra el índice de la iteración en el diseño del filtro. La salida del algoritmo presentado, es el filtro *ASDF* creado en la última iteración. En la etapa del reconocimiento se

espera que el desempeño del filtro sea similar a la que se obtuvo en la etapa de diseño. Una gran cantidad de simulaciones por computadora mostraron que para escenas de entrada complicadas (con fondos como fotografías de la realidad y fondos estocásticos), el número de iteraciones necesarias para alcanzar un valor de DC mayor a 0.9 fue de 10.

Es importante resaltar que otros algoritmos iterativos de optimización que se utilizan con frecuencia para diseñar filtros compuestos, tales como el algoritmo de recocido simulado o algoritmos genéticos, requieren de una complejidad computacional significativamente mayor (Billet y Singher, 2002).

V.2. Resultados experimentales

En esta sección se presentan los resultados que se obtuvieron en las diferentes simulaciones efectuadas con el filtro $ASDF$ básico. Se emplearon diversos fondos para medir la efectividad del $ASDF$. Los resultados se miden en términos de la DC y se comparan contra aquellos que arrojan los filtros convencionales CMF , SDF , POF y OF . Para efectos de calcular la DC , el área de la posición del objeto de referencia se definió con un círculo centrado en el origen del objeto y que aproximadamente representa el 8% de su área total. El objeto de referencia que se utilizó en los experimentos es el vehículo que se muestra en la Fig. 10(a), en donde el área blanca debe tener el valor de cero para que cumpla con las definiciones formales de señal de entrada. El tamaño de las imágenes utilizadas es de 256×256 píxeles. El rango dinámico de las señales imagen se encuentra en el intervalo $[0 - 255]$, en donde el color negro representa al valor cero y el color blanco al valor 255. La media y la desviación estándar sobre el área del objeto de referencia son 91 y 21, respectivamente. El tamaño del objeto es de aproximadamente

32×18 píxeles. Para el primer experimento que se considera en esta sección, se empleó un fondo de la realidad, espacialmente inhomogéneo. Se muestra en la Fig. 10(b). La media y la desviación estándar de este fondo son 116 y 40 respectivamente.

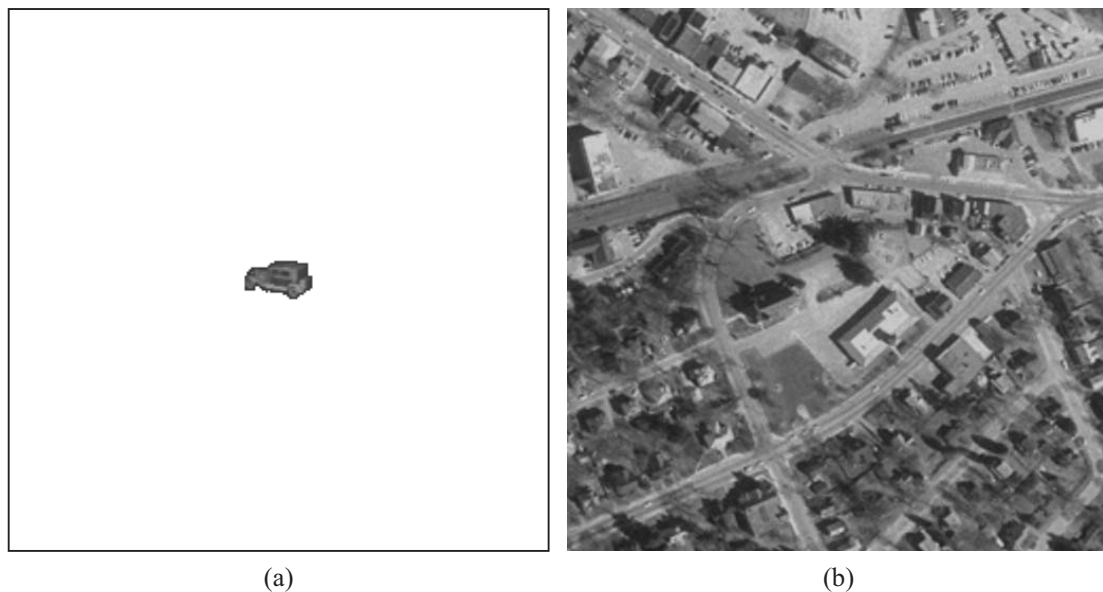


Figura 10: Imágenes de prueba: (a) Objeto de referencia y (b) fondo de la realidad.

Usando como entrada las imágenes de la Fig. 10 en el algoritmo descrito en la sección anterior, se diseñó un filtro *ASDF*. Para apreciar la mejora del filtro por medio del algoritmo propuesto, analicemos los siguientes resultados. En la Fig. 11 se muestran las respuestas al impulso con respecto a los filtros que resultaron después de las iteraciones uno, dos y veintidós respectivamente. En realidad se obtiene una respuesta al impulso con valores reales, los cuales han sido escalados en el rango $[0-255]$ para su visualización. El filtro adaptativo generado en la primera iteración equivale a un filtro CMF, ya que solo se usó al objeto de referencia en su diseño. Las dos primeras imágenes de rechazo

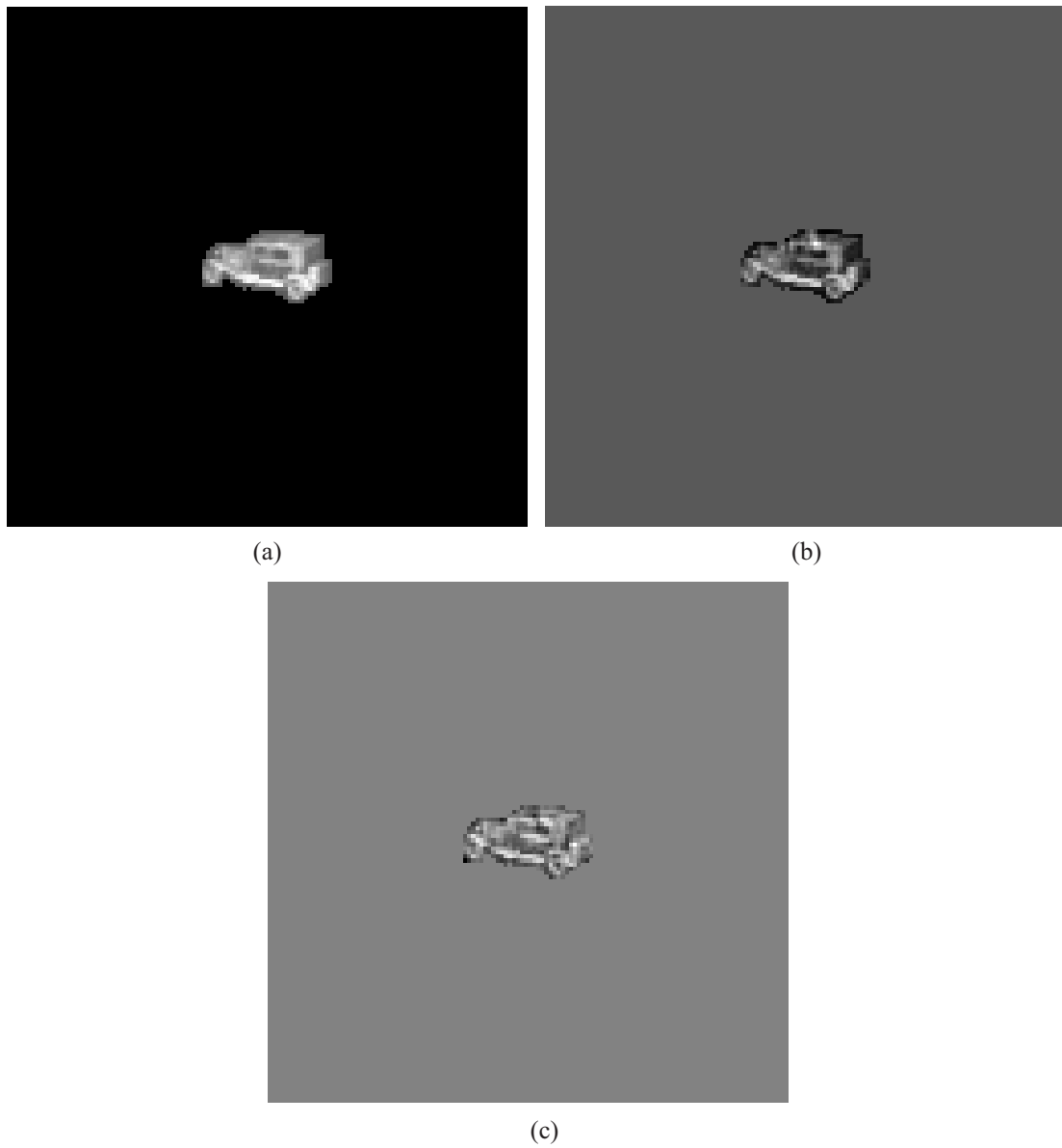


Figura 11: Respuestas al impulso correspondientes a los filtros *ASDF* en la salida de la iteración número: (a) uno, (b) dos y (c) veintidós.

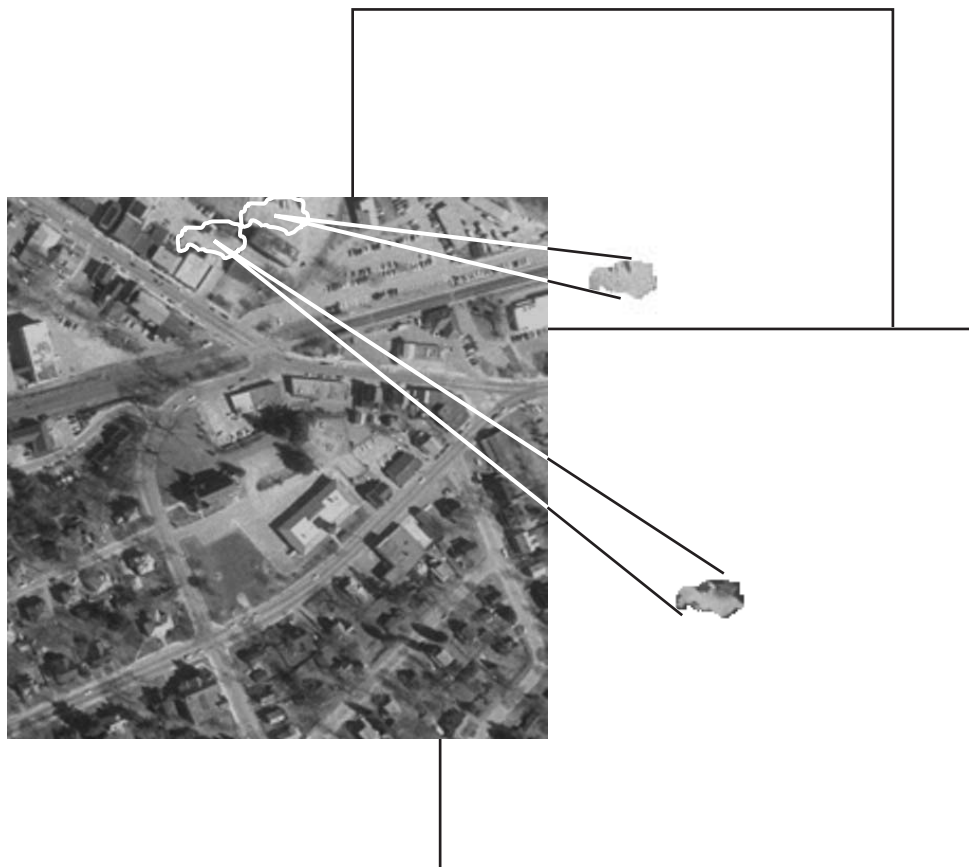


Figura 12: Imágenes de rechazo generadas a partir del fondo y con la misma región de soporte que la del objeto de referencia. El fondo negro de las imágenes de rechazo originales se cambió por blanco para una mejor apreciación.

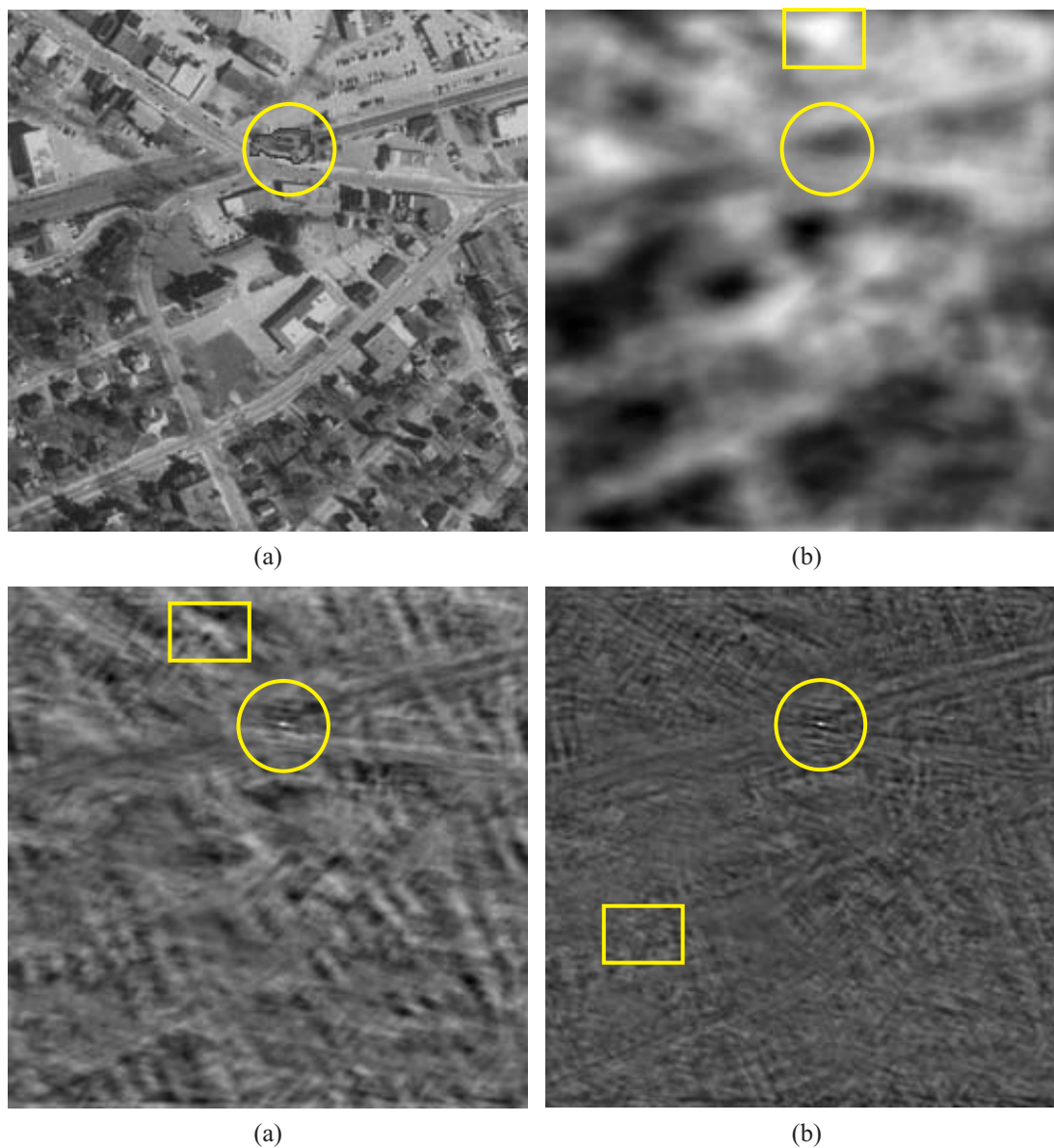


Figura 13: Imágenes de prueba: (a) escena formada con el fondo y el objeto de referencia. (b-d) Salidas de correlación obtenidas con los filtros de la Fig. 11 aplicados a la escena (a). El área del rectángulo indica el centro donde se creará una imagen de rechazo a partir del fondo (máxima correlación en el área del fondo) y el área del círculo indica la posición del objeto de referencia en la escena de prueba.

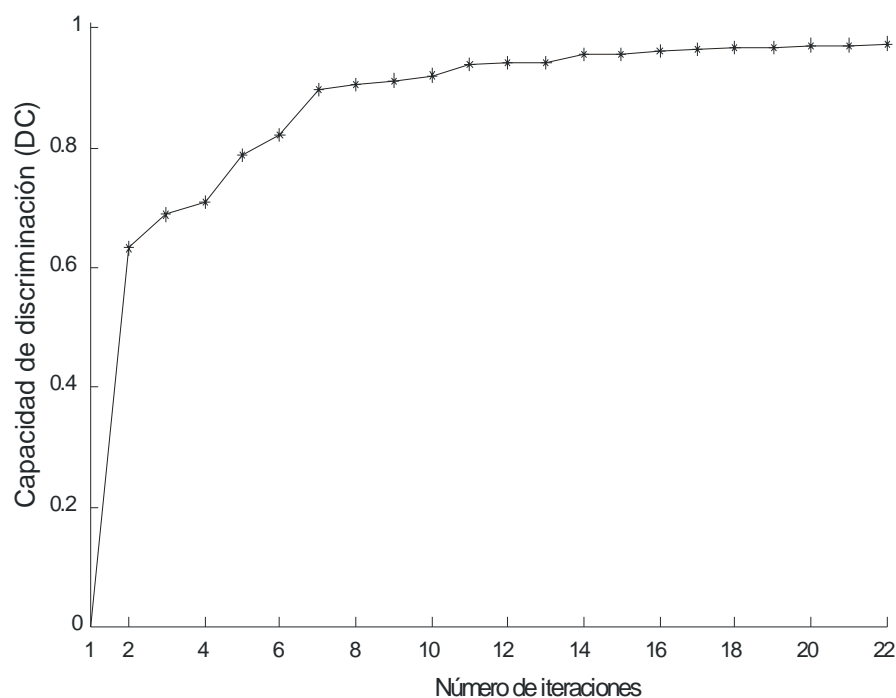


Figura 14: Desempeño del filtro ASDF durante el proceso de diseño, empleando las imágenes de prueba de la Fig. 10.

creadas por el algoritmo a partir del fondo se muestran en la Fig. 12, las cuales se centran para que coincidan con el objeto de referencia. Para probar el desempeño de los filtros en un problema de detección, el objeto de referencia se colocó en una posición arbitraria sobre el fondo, como en la Fig. 13(a). En los incisos (b-d) de la Fig. 13 se muestran las salidas de correlación correspondientes a la escena de prueba, al emplear respectivamente los filtros de la Fig. 11. En los planos de correlación de la Fig. 13, el área al interior del rectángulo indica la posición del pico de correlación cruzada, es decir la correlación entre el filtro y el fondo. La posición de la correlación cruzada corresponde al centro donde se creó la imagen de rechazo que fue utilizada en la siguiente iteración

del algoritmo. El área al interior del círculo corresponde a la posición del objeto en la escena de prueba. La correlación entre el filtro diseñado después de la primera iteración del algoritmo y la escena de prueba proporcionó una $DC = -0.242$. Al usar el filtro diseñado después de la segunda iteración proporciona una $DC = 0.631$ en la salida. La Fig. 13(c) representa a la salida de correlación obtenida con el filtro ASDF que se obtuvo después de la iteración 22, en donde la $DC = 0.975$ es el valor que se espera obtener en las pruebas.

La gráfica de la Fig. 14 muestra la evolución del comportamiento en el desempeño del filtro adaptativo durante el proceso de diseño, en términos de la DC . Antes de la primer iteración el valor de la DC es negativa ($DC = -0.24$). Ya que valores negativos indican que el filtro falló al intentar reconocer al objeto de referencia, se acordó que en todas las gráficas los valores negativos fueran sustituidos por ceros. Obtener una DC cercana a la unidad, como en el caso de nuestro filtro ($DC = 0.975$), significa que se ha logrado un gran control del plano de correlación para la escena de entrada (fondo) que se construirá a partir del fondo y del objeto de referencia.

Como se podrá apreciar en los planos de correlación resultantes (Fig. 13), el objetivo de la optimización del filtro es eliminar los picos de correlación del fondo, o dicho en otras palabras controlar por completo el plano de correlación, y de esta manera obtener un pico bien definido para el objeto de referencia. En teoría, si se utilizara las 65536 imágenes de rechazo correspondientes a todos los píxeles contenidos en la imagen del fondo de entrenamiento (ya que el tamaño de la imagen es de 256×256 píxeles), la salida de la correlación correspondería a un área completamente negra (ceros) con un solo punto blanco, correspondiente a la posición del objeto a reconocer.

Para poder comparar contra otro filtro adaptativo, consideremos la función de transferencia del filtro óptimo de Yaroslavsky (1993), la cual se puede aproximar en el dominio de Fourier por:

$$H_{OF}(u, v) = \frac{T^*(u, v)}{|T(u, v)|^2 + |B(u, v)|^2}, \quad (104)$$

donde $T(u, v)$ y $B(u, v)$ representan las FT del objeto de referenciade y de la escena de entrada, respectivamente. Una característica notable de la aproximación del OF es su adaptabilidad a la escena debido a que su respuesta a la frecuencia considera una aproximación del espectro de energía del fondo a rechazar. Se ha demostrado (Javidi y Wang, 1994; Kober y Campos, 1996) que para el modelo de señal disjunto o no superpuesto, de un objeto a ser detectado y un fondo tipo ruido, se debe formar un nuevo objeto de referencia como la suma del objeto y la función inversa de soporte del objeto ponderada por el valor medio del fondo, es decir, en el dominio de Fourier el espectro del nuevo objeto está dado por $\hat{T}(u, v) = T(u, v) + \mu_B W(u, v)$. Donde μ_B es el valor de la media del fondo, $W(u, v)$ es la transformada de Fourier de la función inversa de soporte del objeto, definida como ceros al interior del área del objeto y unos fuera del área (Kober y Campos, 1996). La función de transferencia del OF en la ecuación (104) utiliza el espectro del nuevo objeto de referencia.

El primer experimento se enfocó en el problema de detección, para lo cual se construyeron escenas con el objeto de referencia y en diferentes posiciones, similares a la mostrada en la Fig. 15(a). Los resultados de las pruebas se presentan en la línea 1 de la Tabla I, en donde se compara con los desempeños de los filtros CMF , SDF , POF y OF en términos de la DC . Como esperabamos, el filtro propuesto, al cual nos referimos como $ASDF$ proporciona el mejor desempeño. Es importante señalar que el desempeño de un filtro puede variar al colocar el objeto en diferentes posiciones dentro del fondo, esto se debe a la inhomogeneidad del fondo o bien de la información que circunda al objeto

de referencia. Por lo tanto, se utilizaron pruebas estadísticas en nuestros experimentos para diferentes posiciones en donde se midió la DC con un intervalo de confianza. El intervalo de confianza al 95 % de la DC resultó ser igual a 0.9752 ± 0.0003 . Nótese que los filtros CMF y POF fallan en el reconocimiento del objeto.

En un segundo experimento, se añadieron cuatro objetos no deseados (objetos falsos) sobre el fondo con la finalidad de confundir al filtro y medir su respuesta en un problema de clasificación. Una muestra de este tipo de escenas se presenta en la Fig. 15(b), donde la posición de los objetos en la escena de prueba se puede apreciar mejor en la Fig. 15(c). El desempeño de los filtros en términos de la DC para este tipo de escenas se muestra en la línea 2 de la Tabla I. En este caso, el filtro adaptativo propuesto también proporciona el mejor desempeño. Para garantizar resultados correctos, se efectuaron 30 pruebas estadísticas del experimento para distintas posiciones de los objetos, y con un 95 % de confianza la DC resultó ser igual a 0.9434 ± 0.0072 . La Fig. 15(d) muestra la distribución de la intensidad de la salida de la correlación que se obtuvo con el filtro $ASDF$ para la última escena de prueba. Los filtros CMF y POF fueron incapaces de reconocer al objeto ya que presentaron grandes picos en el área del fondo.

Tabla I: Desempeño de varios filtros de correlación en términos de su DC para las escenas de prueba de la Fig. 15.

Escena	MSF	POF	OF	ASDF
a	-0.24	-0.63	0.83	0.98
b	-0.53	-0.39	0.54	0.94

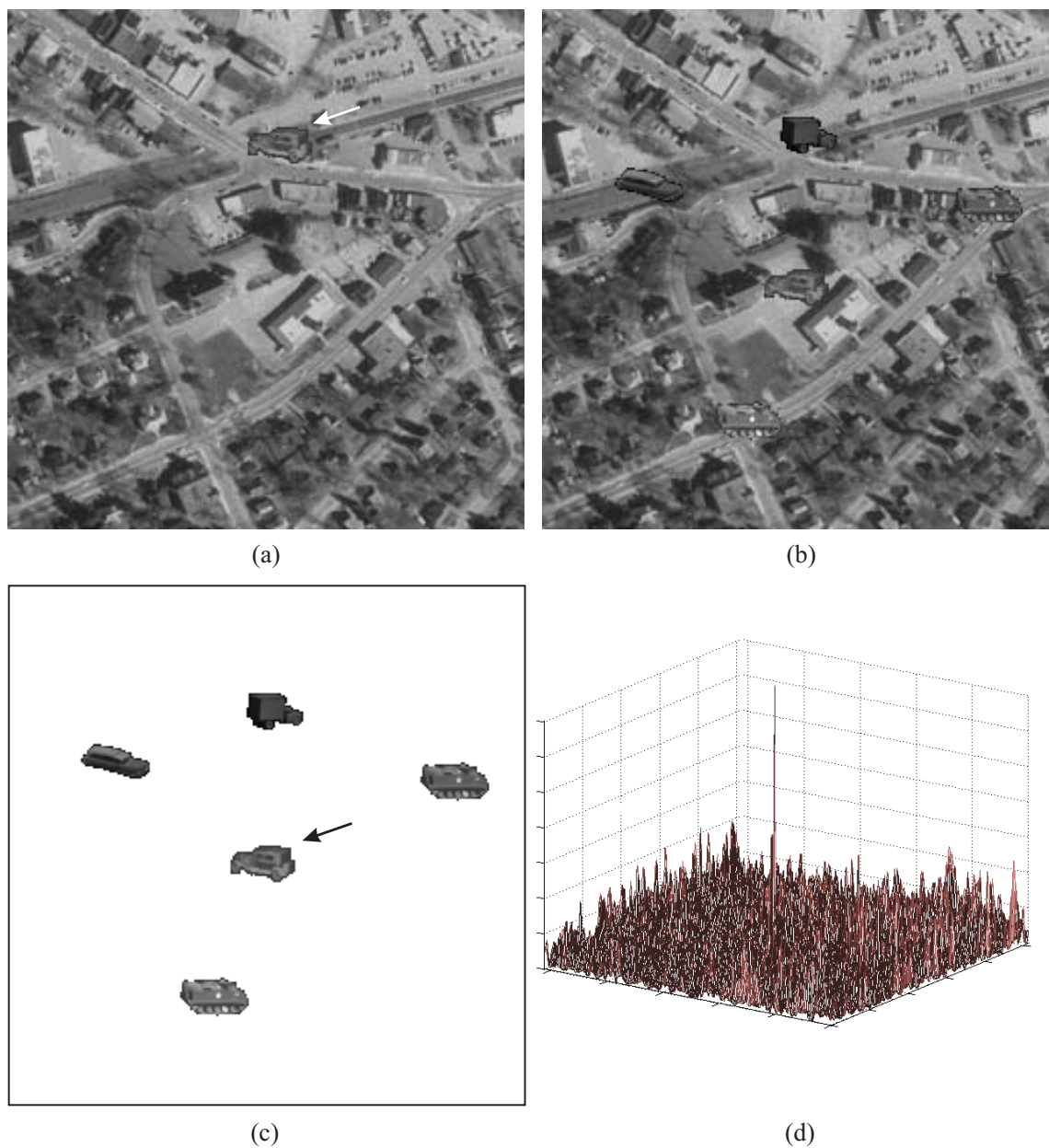


Figura 15: Reconocimiento de objetos en escenas de prueba con una imagen de fondo de la realidad. (a) Escena de prueba que contiene sólo al objeto de referencia marcado con la flecha, (b) escena de prueba que contiene al objeto de referencia y a cuatro objetos no deseados, (c) posición de los objetos en la escena (b), el objeto de referencia está marcado con la flecha, (d) distribución de la intensidad de correlación correspondiente a la escena de prueba (b) y la cual se obtuvo con el filtro ASDF.

Tabla II: Desempeño de varios filtros de correlación en términos de su DC para las escenas de entrada de la Fig. 16.

Escena	MSF	POF	OF	ASDF
a	-0.12	-0.18	0.30	0.91
b	-0.92	-0.87	0.35	0.93

V.2.1. Experimentos con fondos espacialmente homogéneos

También se analizó el desempeño de los filtros de correlación para el objeto sobre fondos que han sido generados como realizaciones de ruido espacialmente homogéneo. Se consideraron dos tipos de ruido con distribución Gaussiana, estos son realizaciones de procesos estacionarios de ruido blanco (Fig. 16(a)) y de ruido correlacionado (Fig. 16(b)). El valor medio del fondo siempre fué $\mu_B = 128$ y el valor de la desviación estándar se varió en $\sigma = 10, 20, 30$ y 40 . Para cada tipo de ruido (blanco y correlacionado) y para cada valor diferente de desviación estándar se generó un filtro de acuerdo al algoritmo básico de diseño del *ASDF*. Para el ruido correlacionado, el coeficiente de correlacion siempre se tomó como $\rho = 0.9$. Las Fig. 17 y 18 muestran el desempeño de los filtros adaptativos en el proceso de diseño en términos de la DC contra el índice de la iteración para un conjunto de valores de desviación estándar para los modelos de ruido blanco y correlacionado. Se puede observar que el desempeño del filtro adaptativo para estos dos modelos es similar. Nuevamente, un valor determinado de DC se puede alcanzar en ambos casos con pocas iteraciones.

Las simulaciones por computadora se dirigieron de una manera similar a las que se realizaron con el fondo de la realidad. En primera instancia se probaron los filtros de correlación con el objeto de referencia colocado arbitrariamente en la imagen fondo. Después, se insertaron cuatro objetos no deseados en la escena (ver Fig. 16). En la Fig.

16 se muestran algunos ejemplos de las escenas de prueba de un fondo con desviación estándar 40.

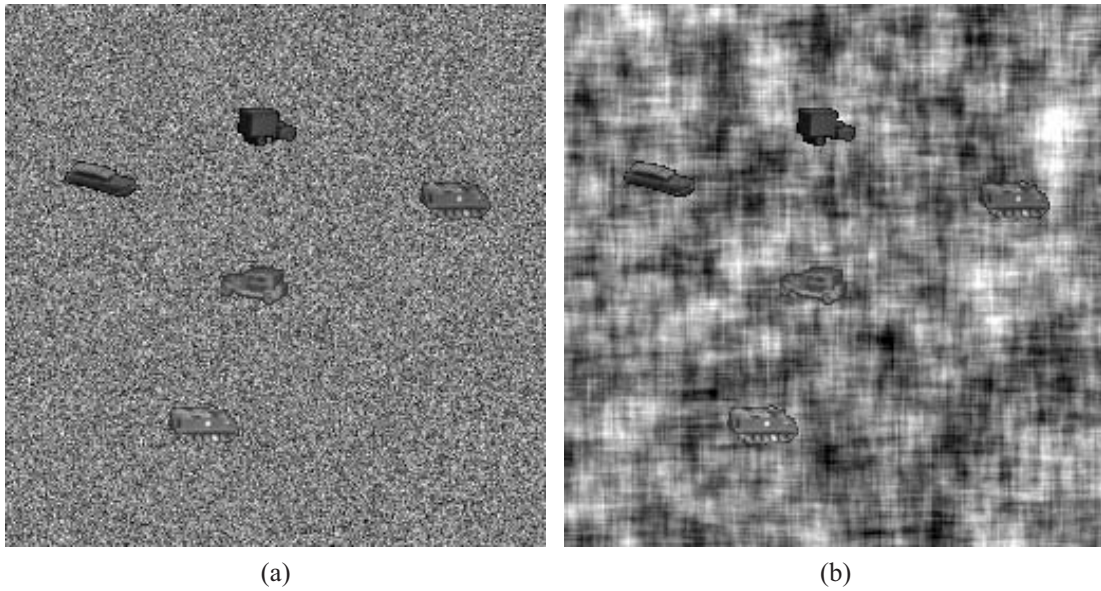


Figura 16: Reconocimiento de objetos en escenas de prueba con fondos generados como procesos estocásticos. (a) Escena de prueba que contiene al objeto de referencia y a cuatro objetos no deseados incrustados en un fondo tipo ruido blanco con desviación estándar de 40 y valor medio de 128, (b) Escena de prueba que contiene al objeto de referencia y a cuatro objetos no deseados incrustados en un fondo tipo ruido correlacionado con desviación estándar de 40, valor medio de 128 y coeficiente de correlación de 0.9. La posición de los objetos es similar a la de la Fig. 15(c).

El desempeño de varios filtros de correlación con respecto a la DC para las escenas de prueba en la Fig. 16(a) contra la desviación estándar del fondo tipo ruido blanco se muestra en la Fig. 19. Es claro que el algoritmo propuesto es capaz de adaptar bien el filtro diseñado a las variaciones del fondo tipo ruido, mientras que el desempeño del OF decae rápidamente cuando el valor de la desviación estándar del fondo se incrementa.

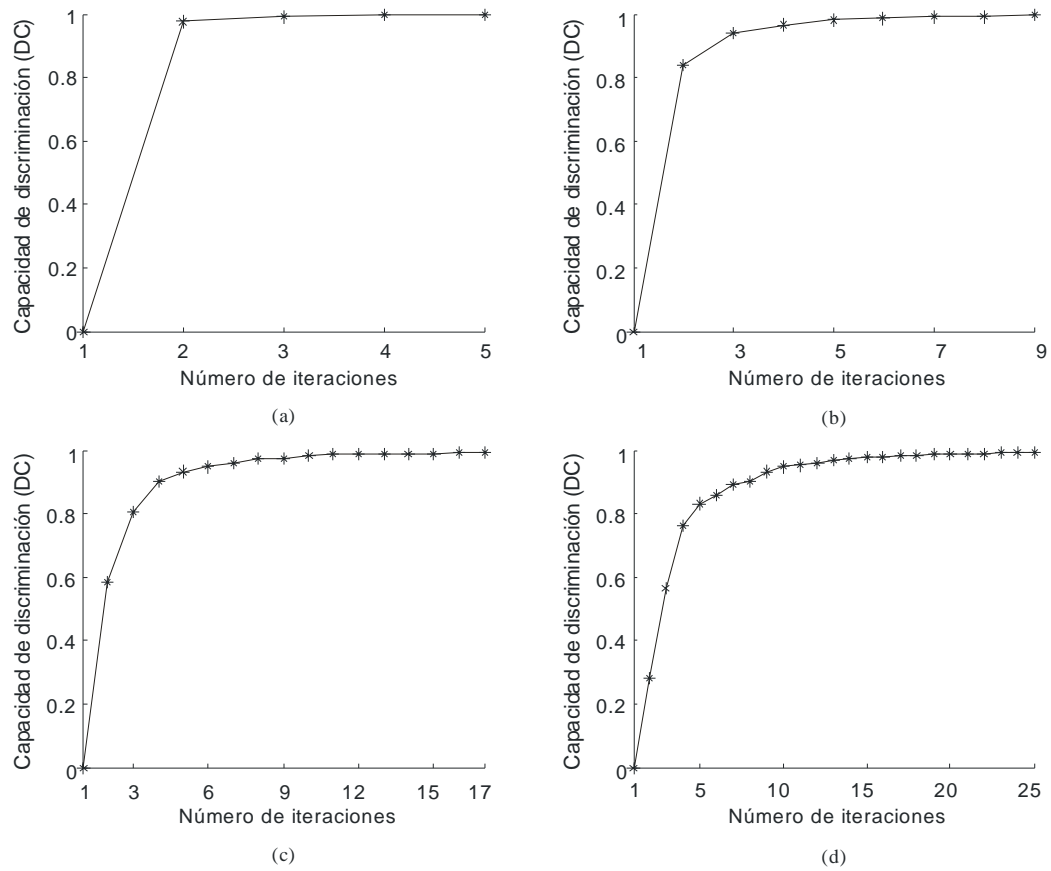


Figura 17: Desempeño del filtro *ASDF* durante el proceso de diseño para el fondo tipo ruido blanco con diferentes desviaciones estándar: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 30$, (d) $\sigma = 40$.

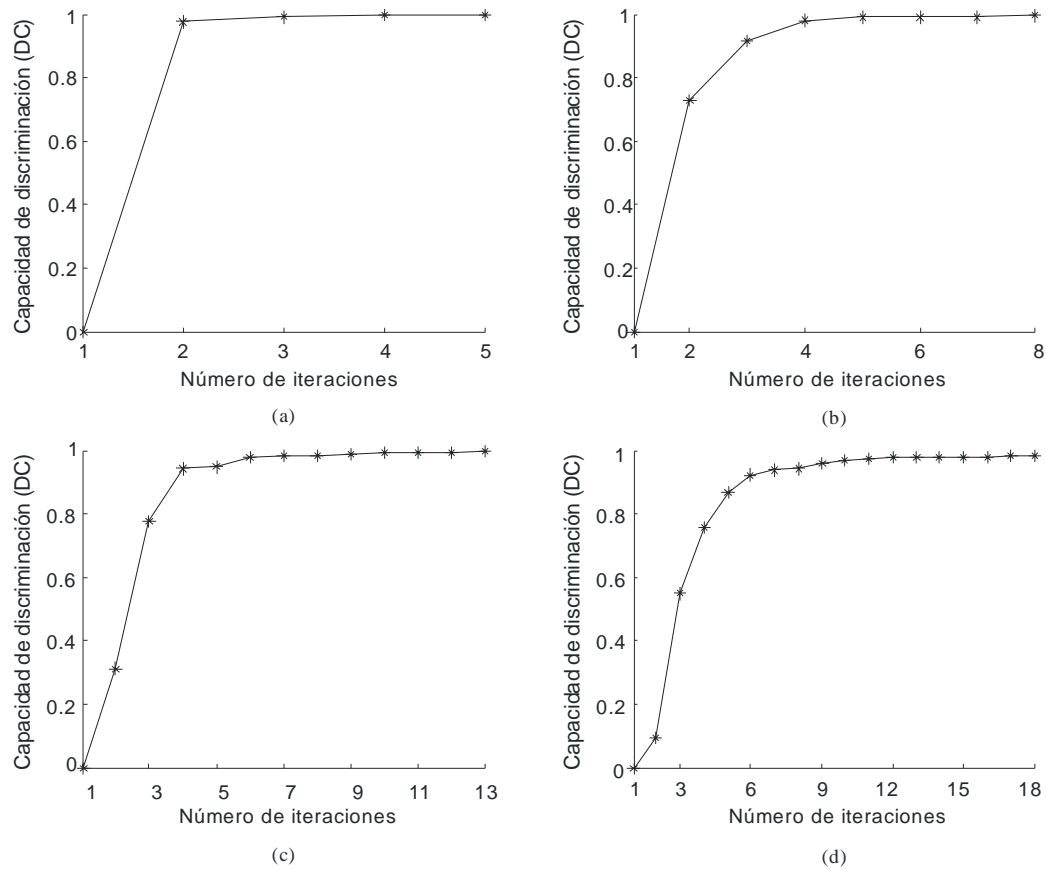


Figura 18: Desempeño del filtro *ASDF* durante el proceso de diseño para el fondo tipo ruido correlacionado con diferentes desviaciones estándar: (a) $\sigma = 10$, (b) $\sigma = 20$, (c) $\sigma = 30$, (d) $\sigma = 40$.

El desempeño de los filtros para el fondo tipo ruido correlacionado es similar. Los resultados numéricos del desempeño de los filtros de correlación para las escenas de prueba mostradas en la Fig. 16 se proporciona en la Tabla II (la desviación estándar del ruido en ambos casos es de 40). Se realizaron 60 experimentos estadísticos para diferentes posiciones de los objetos y para diferentes realizaciones del fondo tipo ruido. Por ejemplo, para el ruido blanco con desviación estándar 40 se generó un filtro, el cual

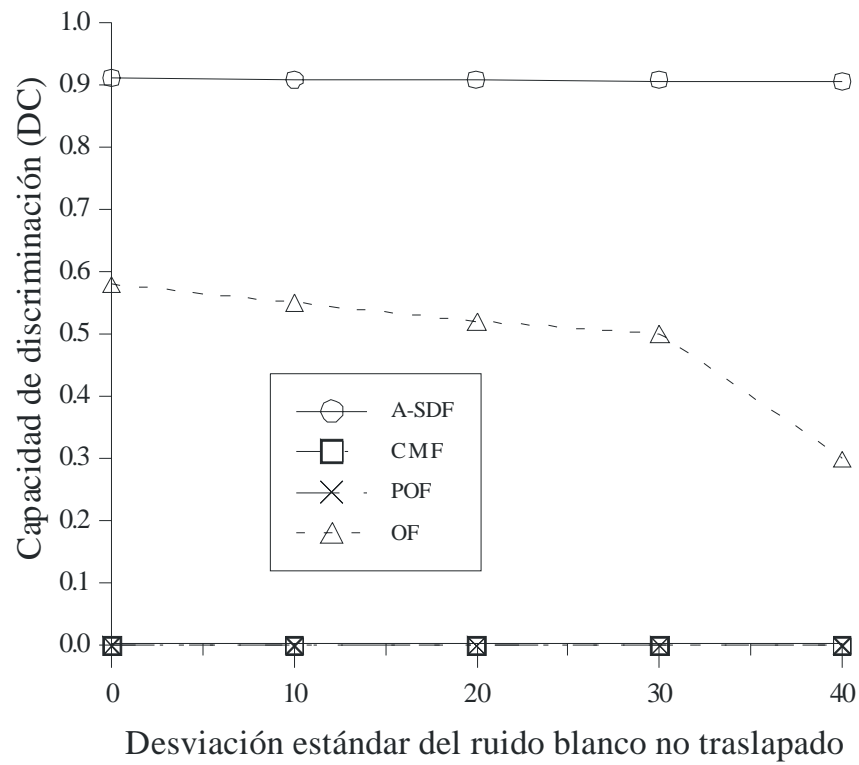


Figura 19: Desempeño de los filtros de correlación cuando la escena de entrada es la mostrada en la Fig. 16(a). El desempeño se midió en términos de la DC contra la desviación estándar del fondo tipo ruido blanco.

se probó con varias realizaciones de ruido blanco pero con el mismo valor de desviación estándar.

Para las escenas (a) y (b), la *DC* con un intervalo de confianza del 95 % resultó ser de 0.914 ± 0.0062 y 0.932 ± 0.0058 respectivamente. Se puede apreciar que el filtro propuesto proporciona el mejor desempeño en términos de la *DC*. El desempeño del *OF* es pobre debido a la incorporación de los cuatro objetos no deseados en la escena de prueba. Lo que nos conduce a un modelo espacialmente inhomogéneo del fondo. Así que el filtro no es efectivo. Los filtros *CMF* y *POF* fallaron en todos los experimentos al intentar detectar al objeto de referencia.

V.2.2. Incorporación de invariancias al filtro ASDF

En esta sección se describe la tolerancia de los filtros de correlación para pequeñas distorsiones geométricas en la imagen de referencia. Varios métodos se han propuesto para mejorar el reconocimiento de patrones en la presencia de tales distorsiones, estos métodos se pueden clasificar ampliamente en dos grupos. La primera clase concierne formalmente con las distorsiones en dos dimensiones de escalamiento y rotación. Tales métodos incluyen transformadas variantes al espacio y funciones armónicas circulares. La segunda clase de métodos utilizan imágenes de entrenamiento que son lo suficientemente descriptivas y representativas de las distorsiones esperadas. El método propuesto está basado en la segunda técnica. En nuestros experimentos se utilizó la escena de entrada mostrada en la Fig. 15 con un objeto incrustado y distorsionado geoméricamente. Se efectuaron 60 experimentos estadísticos para cada prueba con diferentes posiciones del objeto distorsionado. En los experimentos previos, la clase verdadera solo contenía a un solo objeto. Cuando incorporamos invariancias, la clase verdadera contendrá más de un objeto para representar las posibles versiones en que se podría encontrar el objetivo a identificar.

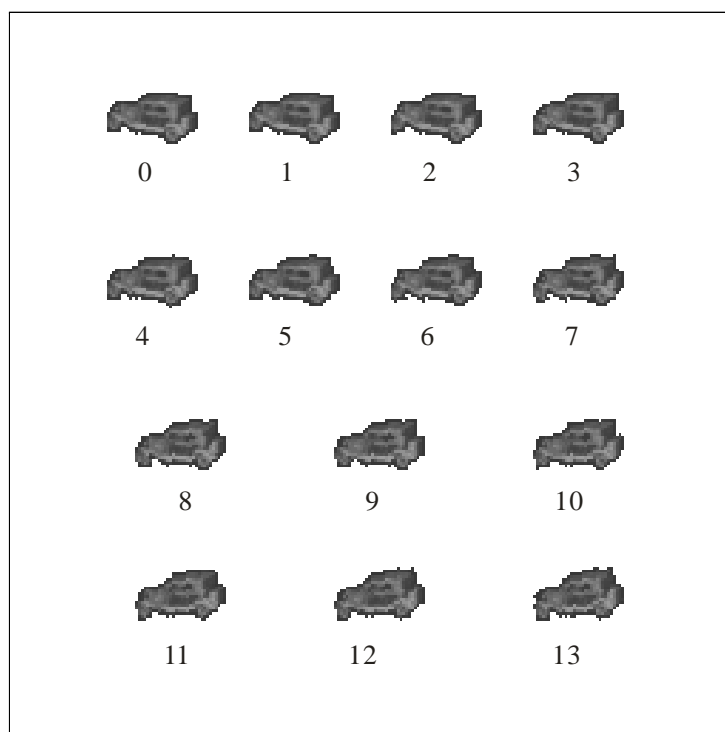


Figura 20: Diferentes versiones del objeto a reconocer. El número debajo de cada objeto indica el grado de rotación que se le aplicó al objeto original.

Invariancia a la rotación

Primero se investigó la distorsión geométrica representada por la rotación. Se generaron diferentes versiones rotadas del objeto de referencia, tal como se muestran en la Fig. 20. El rango de las rotaciones (en grados) del objeto a identificar fue de 0 a 13 con un incremento de un grado. Nuevamente, los resultados se compararon contra los filtros *POF*, *OF* y el *SDF* convencional. El filtro *SDF* convencional se diseñó con las versiones del objeto rotado por 0, 2, 4, 6, 8, 10, y 12 grados (ver Fig. 21). El filtro *ASDF* se entrenó con dos versiones del objeto: el primero fue rotado por 0 grados y el segundo por 8 grados. Después de 22 iteraciones el filtro *ASDF* que se obtuvo proporcionó una $DC = 0.97$. En la Fig. 21 se presenta una ampliación de los negativos de las

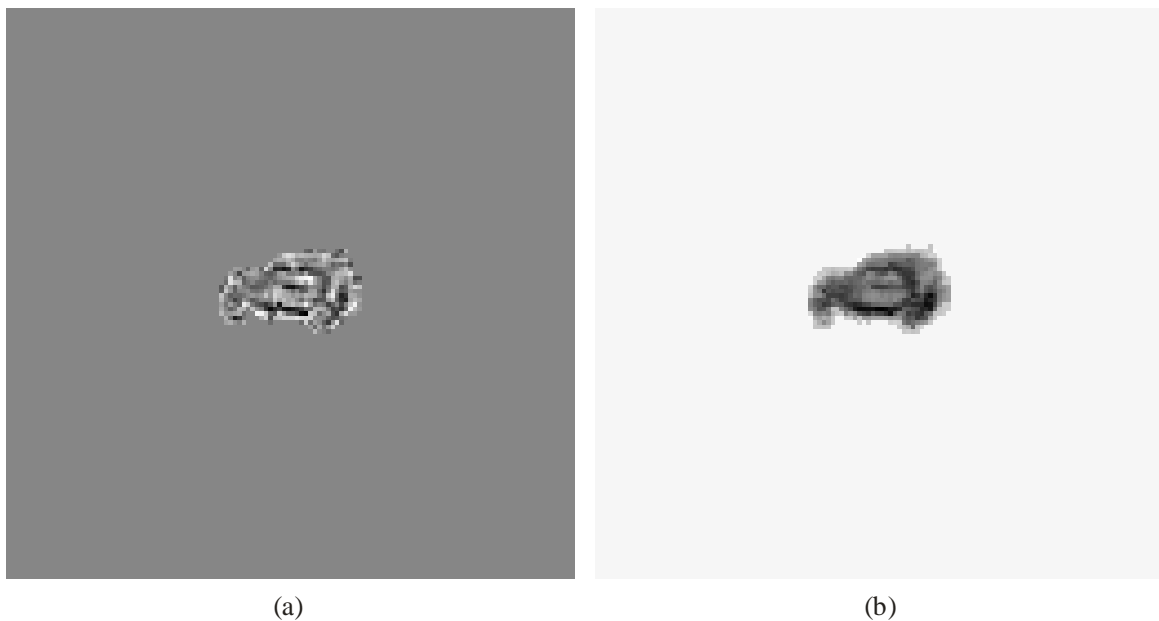


Figura 21: Respuestas impulsivas correspondientes a los filtros: (a) *ASDF* (generado con los objetos 0 y 8 de la Fig. 20) y (b) *SDF* convencional (generado con los objetos 0, 2, 4, 6, 8, 10, y 12 de la Fig. 20).

Tabla III: Desempeño de filtros de correlación en términos de su *DC* para objetos rotados en la escena de entrada de la Fig. 15(a)

Grado de rotación	OF	ASDF
0	0.83	0.95
1	0.68	0.94
2	0.32	0.93
3	0.26	0.91
4	0.00	0.90
5	0.00	0.91
6	0.00	0.93
7	0.00	0.94
8	0.00	0.95
9	0.00	0.94
10	0.00	0.91
11	0.00	0.88
12	0.00	0.83
13	0.00	0.81

imágenes que representan las respuestas impulsivas de los filtros *ASDF* y *SDF* convencional generados tal como se describió en este párrafo. En el caso del filtro *ASDF*, la región de soporte para generar las imágenes de rechazo se definió como unos al interior de la unión de los objetos 0 y 8, y ceros fuera de esta. Los resultados experimentales mostraron que el filtro *SDF* convencional siempre falla para detectar a los objetos de referencia rotados al incrustarlos en el fondo. El desempeño del *OF* y del *ASDF* en términos de la *DC* se proporciona en la Tabla III. Se puede apreciar que el desempeño del *OF* se degrada rápidamente cuando se incrementa la distorsión del objeto. El filtro propuesto se adapta bien cuando se entrena con versiones del objeto de referencia que presentan pequeñas rotaciones.

Invariancia al escalamiento

La tolerancia al escalamiento del objeto también se investigó. Para incorporar la invariancia a la escala se tuvo que buscar un mecanismo para intentar homogeneizar la energía contenida en todos los objetos. Esto se consiguió al incrustar los objetos dentro de un campo homogéneo generado con varios pixeles aleatorios que se tomaron del fondo. Este campo homogéneo tiene una región de soporte igual a la unión de los objetos involucrados en el diseño (ver Figs. 22 y 23). El rango del factor de escala del objetivo se varió de 0.8 a 1.2 con un incremento de 0.04. El filtro *SDF* convencional se diseñó con todas las versiones escaladas del objeto de la Fig. 22. El filtro *ASDF* se entrenó solo con cinco versiones del objeto escalado por los factores 0.8, 0.9, 1.0, 1.1, y 1.2. Después de 120 iteraciones el filtro *ASDF* resultante proporcionó una $DC = 0.96$. El filtro *SDF* convencional falló en el reconocimiento de los objetos escalados. El desempeño de los filtros *OF* y *ASDF* se proporciona en la Tabla IV. Se puede observar que el filtro *OF* es muy sensible a las distorsiones de escala. El filtro *ASDF* siempre detecta los objetos escalados. Nótese que en este caso el diseño del filtro fue computacionalmente

más costoso en comparación con aquel que se diseñó para las distorsiones de rotación. En una manera similar, el método propuesto se puede utilizar para diseñar un filtro adaptativo que posea una buena tolerancia a las distorsiones geométricas arbitrarias. La complejidad del diseño del filtro compuesto depende del número de imágenes de entrenamiento.

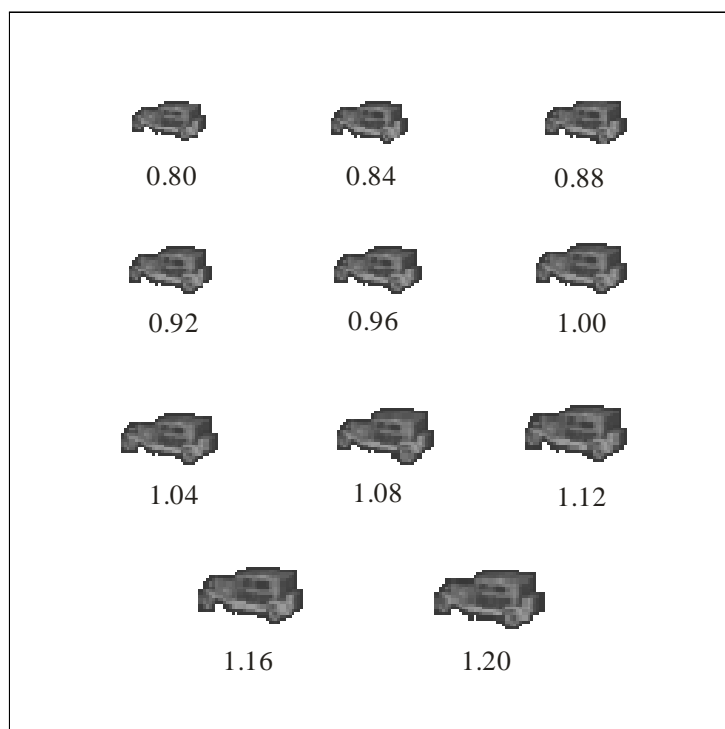


Figura 22: Diferentes versiones del objeto a reconocer. El número debajo de cada objeto indica el factor de escalamiento que se le aplicó al objeto original.

Tolerancia al ruido aditivo

Finalmente se probó la robustez de los filtros de correlación frente al ruido aditivo (que generalmente lo provocan los sensores durante la adquisición de las imágenes). La

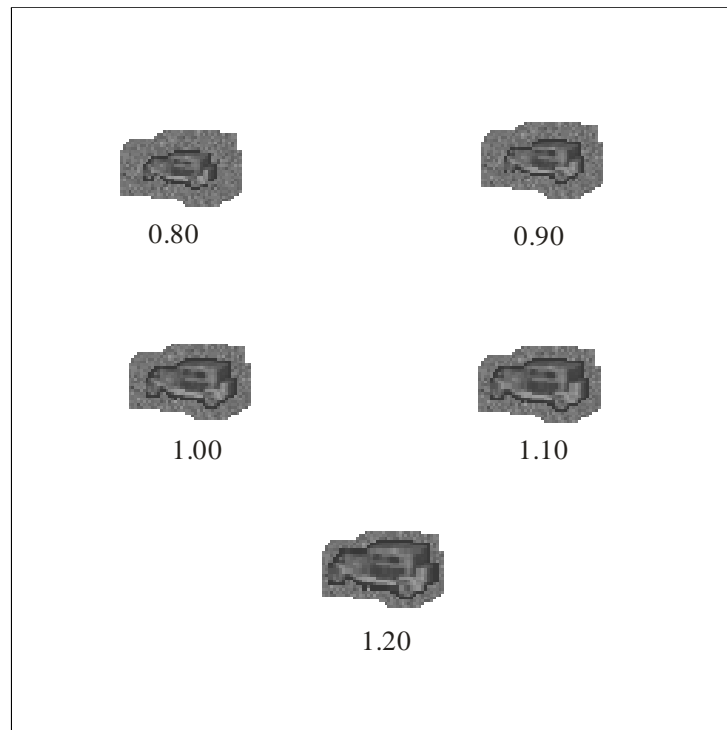


Figura 23: Objetos modificados, los cuales se utilizaron para diseñar el filtro *ASDF* con invariancia a la escala.

Tabla IV: Desempeño de varios filtros de correlación en términos de su *DC* para objetos escalados en la escena de entrada de la Fig. 15(a).

Factor de escala	OF	ASDF
0.80	0.00	0.92
0.84	0.00	0.87
0.88	0.00	0.89
0.92	0.00	0.89
0.96	0.08	0.88
1.00	0.83	0.92
1.04	0.10	0.88
1.08	0.00	0.89
1.12	0.00	0.90
1.16	0.00	0.89
1.20	0.00	0.93

Fig. 24 muestra a las escenas de las Figs. 15(b) y 16(b) alteradas por ruido aditivo con media cero, de naturaleza Gaussiana y con una desviación estándar de 40. De forma similar, los experimentos de reconocimiento con los filtros *OF* y *ASDF* se realizaron con varios valores de desviación estándar del ruido. La Fig. 25 presenta la tolerancia de los filtros frente al ruido aditivo para las escenas mostradas en la Fig. 24. Puesto que la síntesis de un *ASDF* considera al ruido aditivo durante el entrenamiento con una realización del ruido, entonces el filtro proporciona una buena robustez frente a este. En contraste, el desempeño del filtro *OF* se ve deteriorado rápidamente cuando el ruido se incrementa.

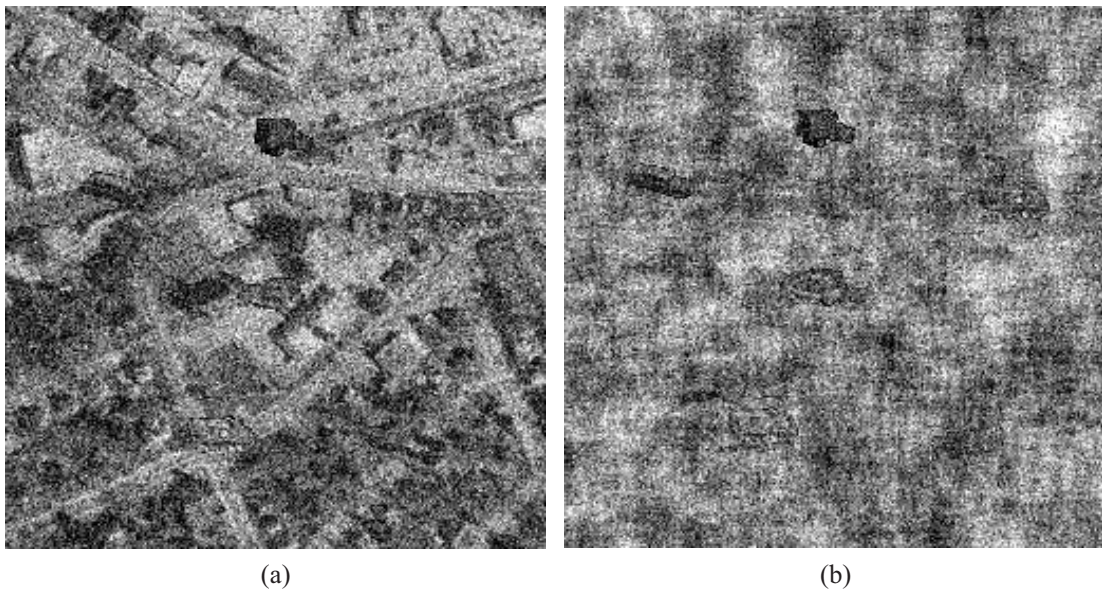


Figura 24: Escenas de prueba alteradas por ruido aditivo de media cero y $\sigma = 40$: (a) escena mostrada en la Fig. 15(b) y (b) escena mostrada en la Fig. 16(b) .

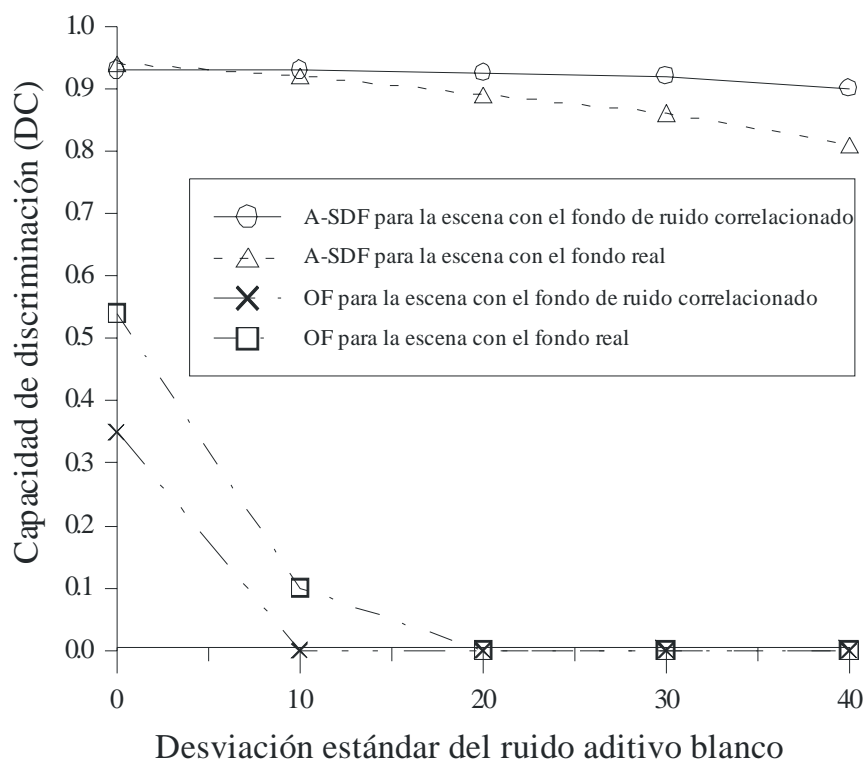


Figura 25: Tolerancia al ruido aditivo blanco de los filtros *OF* y *ASDF*. El desempeño de los filtros se probó con las escenas mostradas en la Fig. 24. El desempeño se midió en términos de la *DC* contra la desviación estándar del ruido.

V.3. Mejora al algoritmo básico del filtro ASDF

Nuestro problema de diseño de filtros *ASDF* se puede formular correctamente como un problema de optimización, ya que se desea maximizar la *DC* con respecto a ciertos patrones conocidos (fondo y objetos falsos). En cada iteración se desea encontrar la imagen de rechazo que mejore la *DC* de la iteración anterior, de tal manera que el comportamiento del filtro sea monotónicamente ascendente. La heurística utilizada para seleccionar las imágenes de rechazo en cada iteración, es conocida como “*algoritmos voraces*”. Ya que se aceptan, como primer ocurrencia, las coordenadas del valor de correlación cruzada más alto como el centro de la imagen de rechazo que proporcionará una *DC* mayor con respecto a la iteración anterior. Sin embargo, para diferentes fondos el algoritmo básico para el diseño del filtro no siempre converge al valor de *DC* deseado (que se espera sea el mismo valor en las pruebas experimentales).

En caso de que en una iteración en particular, la imagen de rechazo no proporcione una *DC* mayor, se implementó una sencilla heurística denominada “*ascenso de colina*” (mejor conocidos por algoritmos *hill-climbing* (Michalewicz y Fogel, 2000)). El método que se implementó, consiste en crear un vecindario (lista de coordenadas) de donde saldrá la imagen de rechazo que proporcionará la *DC* que satisfaga la condición buscada. El vecindario se crea al ordenar los valores de correlación cruzada en orden descendente y se delimita por los primeros k elementos. Dentro de cada iteración, para cada elemento de la lista de coordenadas se crea una imagen de rechazo y se prueba su capacidad de discriminación. Si la *DC* es mejor que la *DC* de la iteración anterior, entonces se acepta esa imagen de rechazo y se continua con la siguiente iteración del algoritmo.

Para poder diseñar el filtro *ASDF* con invariancia al escalamiento se utilizó esta nueva versión del algoritmo.

V.4. Resumen

En este capítulo se detalló un algoritmo propuesto para diseñar filtros adaptativos de correlación para el reconocimiento confiable de objetos incrustados en un fondo complicado. Estos filtros, a los que denominamos *ASDF*, están basados en el filtro *SDF* convencional. Para diseñar el filtro *ASDF*, la información que se conoce tanto del objeto de referencia como de los objetos no deseados y del fondo a rechazar, se utiliza en un procedimiento de entrenamiento y de manera iterativa para que nos permita alcanzar un valor de discriminación determinado. También se presentaron y se discutieron algunas simulaciones por computadora en las que se demuestra que el filtro *ASDF* propuesto tiene una mejor capacidad de discriminación y una mejor tolerancia al ruido aditivo y a las pequeñas distorsiones geométricas.

Se demostró que el algoritmo iterativo para el diseño del filtro requiere unas cuantas iteraciones para tomar el control sobre todo el plano de correlación. Las simulaciones mostraron la superioridad en el desempeño del filtro propuesto para el reconocimiento de patrones comparado contra los filtros *CMF*, *POF* y *OF*. Los filtros propuestos poseen una gran adaptabilidad a la escena y una buena robustez frente al ruido aditivo y a las pequeñas distorsiones geométricas en el objeto de referencia. Finalmente, nótese que el algoritmo iterativo propuesto para el diseño del filtro se puede emplear para mejorar el desempeño de cualquier filtro de correlación al adaptar el filtro a la escena de entrada. Además se puede usar cualquier otra métrica de calidad para optimizar el filtro.

Capítulo VI

Filtros adaptativos para reconocer objetos parcialmente ocultos

El reconocimiento de patrones por métodos de correlación es un campo en la computación y en la óptica que ha estado progresando continuamente. Se sabe que los filtros de correlación utilizan principalmente la información del contorno de los objetos para realizar el reconocimiento. Sin embargo, en muchas aplicaciones del mundo real, los objetos y sus contornos pueden desaparecer parcialmente, esto se puede deber a posibles oclusiones con otros cuerpos u objetos. De aquí que los filtros de correlación convencionales pueden tener un desempeño pobre para reconocer fragmentos de objetos y objetos parcialmente ocultos. Durante las dos últimas décadas, se han propuesto varios métodos para el reconocimiento de objetos parcialmente ocultos (Mardia *et al.*, 1997; Wiskott y Malsburg, 1993; Campos *et al.*, 1994; Khoury *et al.*, 1998b; Gonzalez-Fraga *et al.*, 2005b; Villalobos-Flores *et al.*, 2002; Javidi *et al.*, 2006).

En el capítulo anterior se propusieron nuevos filtros de correlación, los cuales están basados en el filtro convencional *SDF*. Estos filtros son adaptables a las escenas de entrada que se crean con los objetos de referencia y de rechazo, estos últimos pueden ser objetos falsos y un fondo. Se resalta que estos filtros son capaces de suprimir los picos de correlación falsos correspondientes tanto al fondo como a los objetos falsos. En otras palabras, el filtro toma un control total sobre el plano de correlación. En este capítulo se explota la misma idea para diseñar filtros adaptativos para el reconocimiento de objetos fragmentados y parcialmente ocultos, los cuales se encuentran incrustados en escenas

de la realidad y en presencia de ruido aditivo. Los filtros de correlación propuestos son adaptables a una escena de entrada, la cual esta compuesta por fragmentos de un objeto de referencia, objetos falsos y un fondo que se desea rechazar. Los filtros propuestos en este capítulo tienen capacidad para suprimir los picos de correlación correspondientes a un fondo determinado y a los objetos falsos. Además, tienen la característica de proporcionar picos grandes de correlación correspondientes a algunos fragmentos del objeto de referencia. El desempeño de los filtros propuestos se compara contra varios filtros de correlación en términos de su capacidad de discriminación y de su robustez frente al ruido aditivo en la entrada.

Dentro del reconocimiento de patrones dos problemas importantes son la detección y la clasificación, estas tareas se pueden resolver empleando filtros de correlación. La localización de un objeto en la escena se encuentra al buscar los picos de correlación en el plano de salida. La clasificación se consigue si, para cierta clase, el valor del pico de autocorrelación excede determinado umbral. Aunque la manera tradicional de diseñar filtros de correlación es optimizando diferentes criterios, también es posible mejorar algunos criterios empleando el concepto de adaptabilidad (Yaroslavsky, 1993; Gonzalez-Fraga *et al.*, 2006). De acuerdo a este concepto, en este capítulo se propone un algoritmo para diseñar un filtro de reconocimiento, con un buen desempeño para una escena en particular, es decir, con un conjunto de patrones y un fondo fijo que se desea rechazar, considerando que el objeto que se desea reconocer se encuentra parcialmente oculto o bien, fragmentado.

VI.1. Filtros adaptativos para reconocer objetos fragmentados y parcialmente ocultos

Para diseñar un filtro que garantice un gran pico de correlación, correspondiente a la información parcial que representa al objeto que se desea reconocer, es necesario reducir los niveles de la función de correlación en todos los picos correspondientes a los objetos falsos o no deseados excepto en el origen del plano de correlación. De aquí que, para un objeto de referencia, objetos falsos y un fondo a rechazar, se propone un algoritmo iterativo, el cual en cada iteración suprime el pico de correlación más alto, correspondiente al fondo, y por consiguiente se incrementará la capacidad de discriminación de manera monótona hasta llegar a un valor determinado.

Se tiene interés en un filtro de correlación que identifique una porción disponible de un objeto de referencia, dentro de escenas complicadas y ruidosas, y con una discriminación alta. Los filtros convencionales proporcionan un desempeño pobre en estas situaciones (Yaroslavsky, 1993).

El algoritmo para diseñar el filtro requiere del conocimiento de una imagen del fondo y un objeto de referencia $t(x, y)$. Este último se divide en N fragmentos independientes $t_i(x, y)$ que serán utilizados como imágenes de entrenamiento, es decir

$$t(x, y) = \sum_{i=1}^N t_i(x, y) \quad (105)$$

Con esta división del objeto, se espera que al menos uno de los fragmentos reaccione a la información disponible del objeto. Al hablar de información disponible de un objeto nos referimos a que en la escena de entrada aparece una porción del objeto con una forma desconocida. De aquí que, buscamos un fragmento del objeto de referencia, del cual también se desconoce su ubicación dentro del fondo. Es posible que el fondo contenga objetos falsos con coordenadas arbitrarias. El primer paso en el algoritmo, consiste en

realizar la correlación entre el fondo y un filtro *ECP-SDF* básico (tal como se describe en la sección IV.4.1). El cual inicialmente se construye con los fragmentos $\{t_i(x, y); i = 1, 2, \dots, N\}$ del objeto de referencia y fijando los valores del vector $\mathbf{u} = 1$. Luego, el máximo en la salida del filtro se fija como el centro para construir un nuevo objeto de rechazo a partir del fondo. Este nuevo objeto tiene exactamente la misma forma que el objeto de referencia $t(x, y)$ y se agrega a la clase falsa. El valor del vector \mathbf{u} para cada imagen de la clase falsa se fija a 0. Este procedimiento se repite hasta alcanzar un valor determinado de *DC*.

En la práctica fué necesario incorporar la mejora del algoritmo básico que se describe al final del capítulo V. El diagrama de bloques del algoritmo propuesto para diseñar filtros *ASDF* que permitan reconocer objetos fragmentados y parcialmente ocultos se muestra en la Fig. 26.

Al igual que en el algoritmo básico del capítulo anterior, en cada iteración se selecciona de entre todos los picos falsos aquel que se suprimirá en el siguiente paso para asegurar un incremento monótono en la *DC*. La salida del algoritmo presentado, es el filtro *ASDF* creado en la última iteración. Con el algoritmo básico, el desempeño del filtro durante la etapa del reconocimiento se espera que sea similar a la que se obtuvo en la etapa de diseño. Obviamente, dado que el problema es reconocer fragmentos de objetos con una forma desconocida, la *DC* en las escenas de prueba será más baja en comparación con la obtenida en el diseño, pero suficiente para efectuar un reconocimiento exitoso.

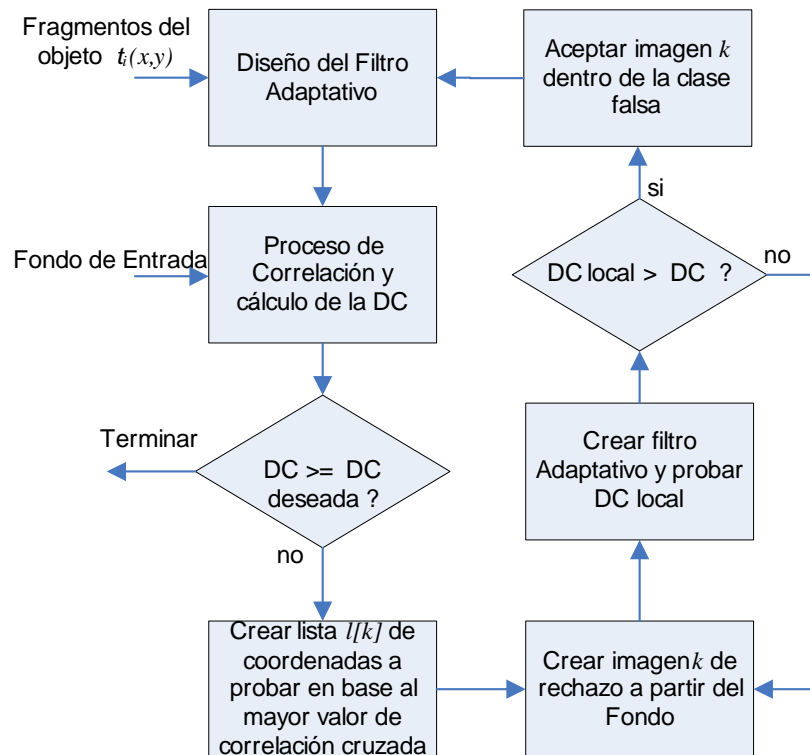


Figura 26: Diagrama de bloques del algoritmo iterativo para diseñar el filtro *ASDF* para reconocer objetos parcialmente ocultos.

Resultados experimentales

En esta sección se presentan los resultados que se obtuvieron al aplicar diferentes filtros de correlación, primero para reconocer objetos fragmentados y luego para reconocer objetos parcialmente ocultos. En el primer caso, se consideran los subproblemas de detección y clasificación por separado. Los resultados obtenidos con los filtros adaptativos propuestos, se comparan con los resultados de los filtros *SDF* convencional, el *POF* y el *MACE* para cada problema en particular.

VI.1.1. Reconocimiento de objetos fragmentados

El objeto de referencia que se utilizó para los experimentos en que se desea reconocer objetos fragmentados es la mariposa en la Fig. 27(a), la cual se particionó en $N = 6$ imágenes independientes como se muestra en la Fig. 28. Cada fragmento independiente debe permanecer en su posición original con el fin de que su contribución en la correlación sea en el mismo punto. El valor medio y la desviación estándar sobre el área del objeto de referencia son 107 y 51 respectivamente. El tamaño del objeto es de aproximadamente 52×33 pixeles. Para los experimentos se utilizó un fondo no homogéneo de la realidad (ver Fig. 30(a)). El valor medio y la desviación estándar del fondo son 116 y 40 respectivamente.

DetECCIÓN

Se estudiaron dos problemas de suma importancia en el ámbito del reconocimiento de patrones, estos son la detección y la clasificación. El objetivo que se persigue en el problema de detección es determinar si se encuentra un fragmento del objeto de referencia en el fondo. Para tal efecto, se diseñó un filtro *ASDF* entrenado con los seis fragmentos de la Fig. 28 y con imágenes de rechazo generadas a partir del fondo. La Fig. 29 muestra el desempeño del filtro *ASDF* en términos de la *DC* contra el índice

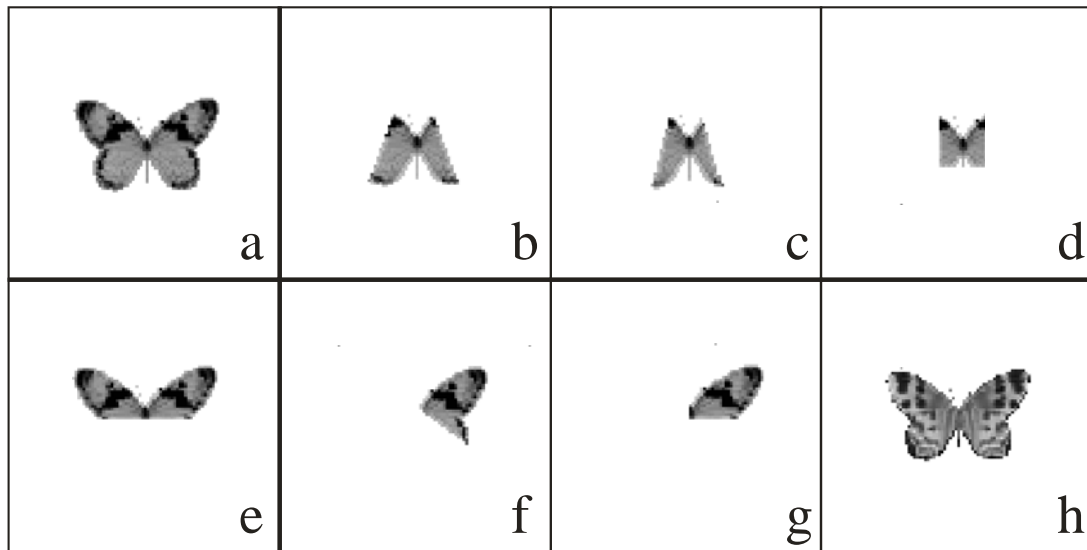


Figura 27: Imágenes de prueba: (a) objeto de referencia, (b-g) fragmentos del objeto de referencia a ser detectados y (h) mariposa falsa.

de iteraciones. En la primera iteración se obtuvo una $DC = -0.52$, en la gráfica este valor fue sustituido por cero. Después de 20 iteraciones, el *ASDF* proporcionó una $DC = 0.982$. Lo que constituye un gran control sobre el plano de correlación para la escena de entrada que se construirá con el fondo y fragmentos del objeto de referencia con una forma desconocida. Se construyeron diferentes escenas de prueba con los objetos que se muestran en la Fig. 27. Los fragmentos se incrustaron en el fondo en posiciones arbitrarias. Las Tablas V y VI muestran el desempeño de los filtros *ASDF*, *POF* y *MACE* en términos de falsas alarmas y DC , respectivamente. El filtro *MACE* se construyó con los mismos fragmentos de la Fig. 28. Se puede apreciar que el filtro propuesto proporciona el menor número de clasificaciones erróneas. El número de pruebas estadísticas para cada fragmento que se colocó en posiciones arbitrarias fue de 50.

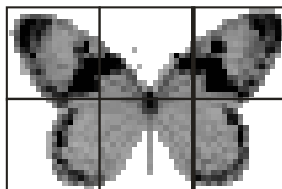


Figura 28: Ampliación del objeto de referencia particionado en seis imágenes independientes.

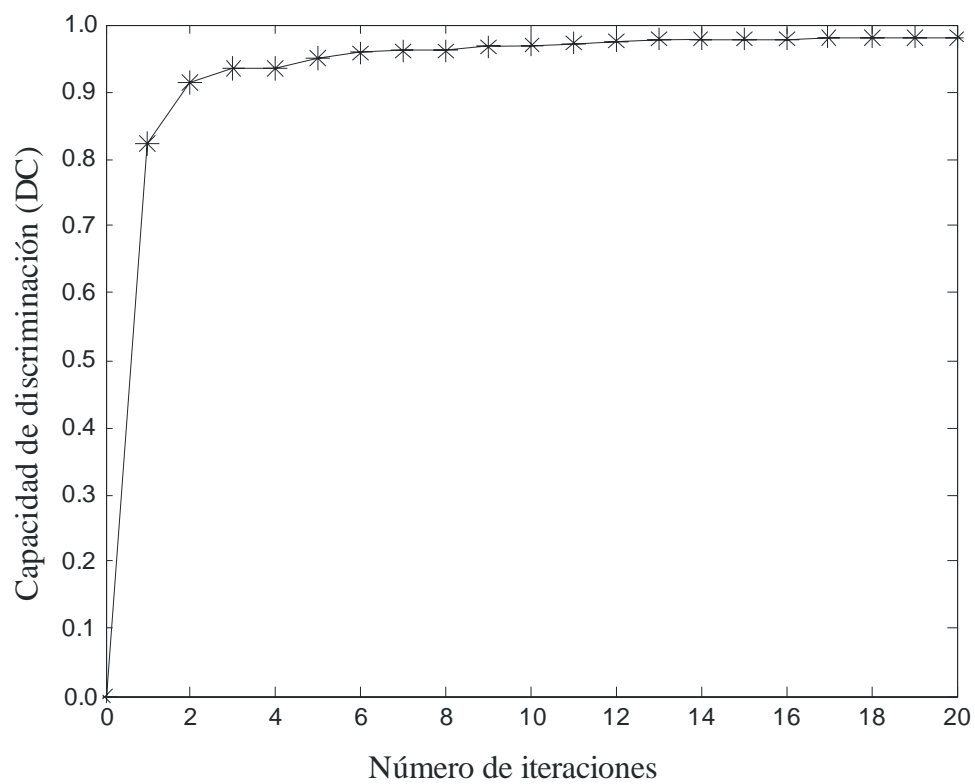


Figura 29: Desempeño del filtro *ASDF* para reconocer objetos fragmentados durante la etapa de diseño.

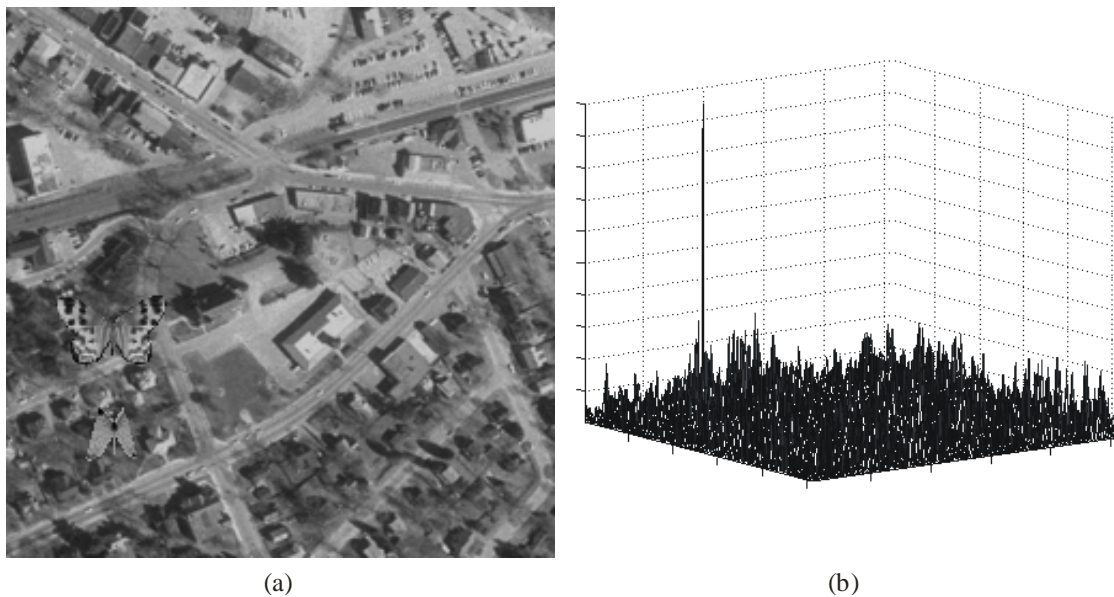


Figura 30: Imagen de prueba para el problema de Clasificación: (a) la escena contiene al objeto fragmentado (b) que se muestra en la Fig. 27 y a la mariposa falsa. (b) Imagen tridimensional de la intensidad correlación resultante para la escena de prueba.

Tabla V: Desempeño de los filtros de correlación en términos de las falsas alarmas (clasificación errónea) para los problemas de Detección y Clasificación. Se usaron 50 pruebas para distintas posiciones de los objetos de la Fig. 27.

Objetos en la escena de prueba	Problema de Detección			Problema de Clasificación		
	POF	MACE	A-SDF	POF	MACE	A-SDF
a	0	0	0	0	0	0
b	2	0	0	12	5	0
c	14	9	0	21	37	0
d	17	8	0	29	42	0
e	0	0	0	0	0	0
f	0	0	0	5	0	0
g	0	0	0	7	0	0

Tabla VI: Desempeño de los filtros adaptativos en términos de su DC para los problemas de Detección y Clasificación. Se usaron 50 pruebas estadísticas para distintas posiciones de los objetos en la Fig. 27.

Objetos en la escena de prueba	Porción disponible del objeto de referencia	DC al 95 % de confianza	
		Detección	Clasificación
a	100 %	0.976 ± 0.001	0.969 ± 0.002
b	45 %	0.869 ± 0.009	0.771 ± 0.013
c	32 %	0.705 ± 0.024	0.446 ± 0.043
d	25 %	0.719 ± 0.021	0.489 ± 0.034
e	62 %	0.934 ± 0.006	0.921 ± 0.005
f	35 %	0.872 ± 0.009	0.794 ± 0.014
g	30 %	0.843 ± 0.011	0.744 ± 0.018

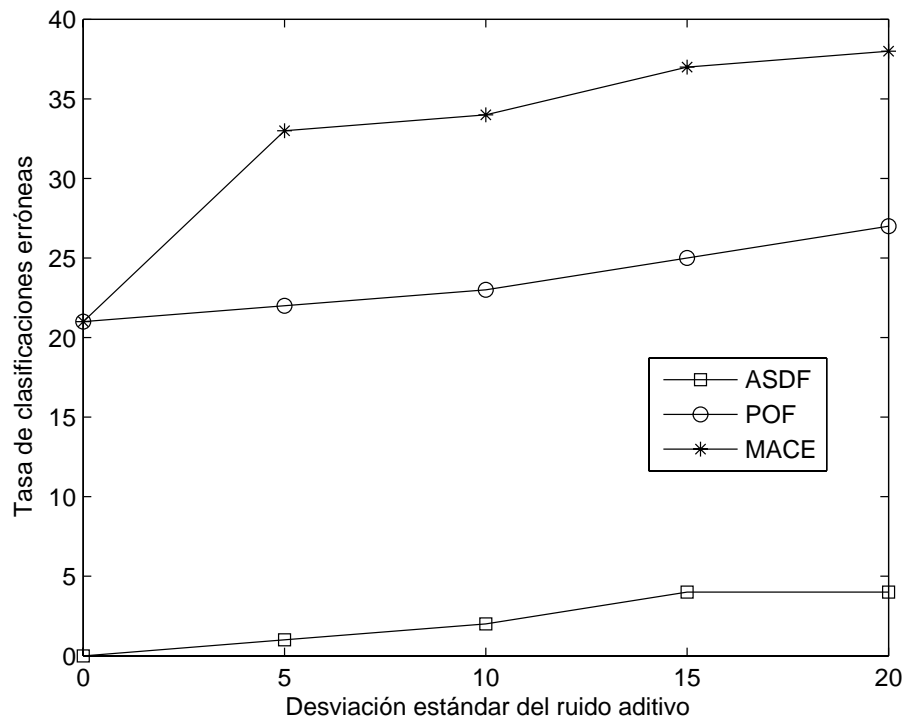


Figura 31: Tolerancia de diferentes filtros hacia el ruido aditivo en términos de falsas alarmas. Se usó la escena de la Fig. 30(a).

Clasificación

Para el problema de clasificación, en el fondo se colocaron tanto los fragmentos a reconocer como una mariposa falsa (ver Fig. 27(h)), la cual tiene la misma forma que el objeto de referencia pero diferente información al interior. El valor medio y la desviación estándar de la mariposa falsa son 109 y 49 respectivamente. Se construyó un nuevo filtro *ASDF* para este problema, es decir para reconocer un objeto fragmentado y para rechazar tanto al fondo como a la mariposa falsa. Después de 38 iteraciones en la etapa de diseño del filtro *ASDF*, se alcanzó una $DC = 0.985$. La Tabla V muestra los resultados de las simulaciones en términos de falsas alarmas, donde se usaron 50 pruebas estadísticas para diferentes posiciones de los objetos. El filtro *ASDF* no produce errores de clasificación erróneas, mientras que los filtros clásicos fallan.

La Fig. 30(a) muestra un ejemplo de las escenas de prueba utilizadas en el problema de clasificación. El objeto fragmentado se muestra en la Fig. 27(c). Para este fragmento, los filtros *POF* y *MACE* tienen la tasa más alta de clasificaciones erróneas. El filtro *ASDF* propuesto es capaz de reconocer acertadamente a estos objetos. Para garantizar resultados correctos, se realizaron 50 pruebas experimentales para diferentes posiciones de los objetos (objeto fragmentado y mariposa falsa), y con un 95 % de confianza la DC resultó ser igual a 0.446 ± 0.043 para el objeto (c). La Fig. 30(b) muestra la distribución de intensidad de la salida de correlación obtenida con el filtro *ASDF*.

La Tabla VI proporciona el desempeño de los filtros adaptativos (*ASDF*) en términos de la DC para los problemas de detección y clasificación. También se utilizaron 50 pruebas estadísticas para diferentes posiciones de los objetos. En las simulaciones presentadas en esta sección el área de la posición del objeto de referencia se definió con un círculo de 10 píxeles de radio, que aproximadamente representa el 2 % del área total del objeto.

Finalmente, se probó la robustez de los filtros de correlación contra el ruido aditivo.

Se usó la escena que se muestra en la Fig. 30. Esta escena se alteró con ruido aditivo blanco con distribución Gaussiana de media cero. El valor de la desviación estándar del ruido se fue variando de 0 hasta 20. La Fig. 31 presenta estos resultados en términos de la tasa de clasificaciones erróneas. Nuevamente, el filtro *ASDF* propuesto presenta una buena tolerancia al ruido aditivo, mientras que el desempeño de los filtros *POF* y *MACE* se ve rápidamente deteriorado cuando la fluctuación del ruido se incrementa.

VI.1.2. Reconocimiento de objetos parcialmente ocultos

Siguiendo el mismo esquema de diseño de filtros *ASDF*, se diseñaron filtros para reconocer objetos parcialmente ocultos. Para tener mayor diversidad en las demostraciones, se optó por cambiar los objetos de prueba. Se construyeron varias escenas similares a la mostrada en la Fig. 36, donde el objeto que se desea localizar se encuentra parcialmente oculto por otro objeto similar, con la misma forma pero diferente contenido. El objeto de referencia que se utilizó en estas nuevas simulaciones es la mariposa en la Fig.32(a), la cual se dividió en $N = 4$ fragmentos y los cuales se muestran en la Figura 33. El valor medio y la desviación estándar sobre el área del objeto a reconocer son 108 y 51 respectivamente. El tamaño de las mariposas es de aproximadamente 90×60 pixeles. Para nuestros experimentos se utilizó el fondo no homogéneo que se muestra en la Fig. 34, cuyos valores estadísticos son, media igual a 89 y desviación estándar igual a 47. El valor medio y la desviación estándar de la mariposa falsa son 103 y 53, respectivamente.

Se utilizó el algoritmo para construir un filtro *ASDF* con el propósito de que reconozca fragmentos del objeto de referencia y rechace la información tanto del fondo como de la mariposa falsa. El desempeño del filtro en la etapa de entrenamiento se puede apreciar en la Fig. 35, donde inicialmente el filtro presentó una $DC = 0.215$ y se fué incrementando hasta la iteración número 30 donde el valor de la DC fué de 0.97. Como ya se mencionó, una DC cercana a la unidad se puede interpretar como que el

filtro tiene un buen control sobre el plano de correlación para la escena de entrada que se construye con el fondo y los fragmentos del objeto de referencia. Una vez que se obtuvo el filtro, se realizaron diferentes pruebas de reconocimiento para múltiples escenas que contienen a los objetos de la Fig.32. Estos objetos se incrustaron en el fondo en una posición arbitraria. El desempeño de los filtros *ASDF*, *POF* y *MACE* con respecto a las falsas alarmas se reporta en la Tabla VII. El filtro *MACE* se construyó con los cuatro fragmentos del objeto de referencia. Se puede apreciar que el mejor desempeño en términos de las falsas alarmas lo obtiene el filtro *ASDF* propuesto. El número de pruebas estadísticas fué de 30. La Fig. 36(a) muestra un ejemplo de las escenas de prueba empleadas para el problema de reconocimiento. Para esta escena, los filtros *POF* y *MACE* proporcionan la tasa más alta en términos de clasificaciones erróneas. El filtro *ASDF* es capaz de reconocer correctamente a los objetos. Para garantizar resultados correctos, se realizaron 30 pruebas para diferentes posiciones de los objetos dentro de la escena de prueba. Se usó un intervalo de confianza al 95%. En particular para los objetos de la Fig. 32(b) se obtuvo una *DC* igual a 0.62 ± 0.01 . La Fig. 36(b) muestra la distribución de intensidad que se obtiene con el filtro *ASDF*.

La Tabla VIII proporciona el desempeño del filtro *ASDF* en términos de la *DC*. También se realizaron 30 pruebas para diferentes posiciones de los objetos.

Por último, se prueba la robustez de los filtros de correlación frente a pequeñas distorsiones geométricas. Se usaron los objetos de la escena de prueba que se muestran en la Fig. 36(a). La Fig. 37 presenta la tolerancia del filtro frente a la rotación y la Fig. 38 frente al escalamiento. Cabe señalar que el filtro propuesto presenta la mejor tolerancia frente a las pequeñas distorsiones geométricas, ya que el desempeño de los filtros *POF* y *MACE* se ve fuertemente deteriorado cuando se presentan estas distorsiones en el objeto a reconocer.

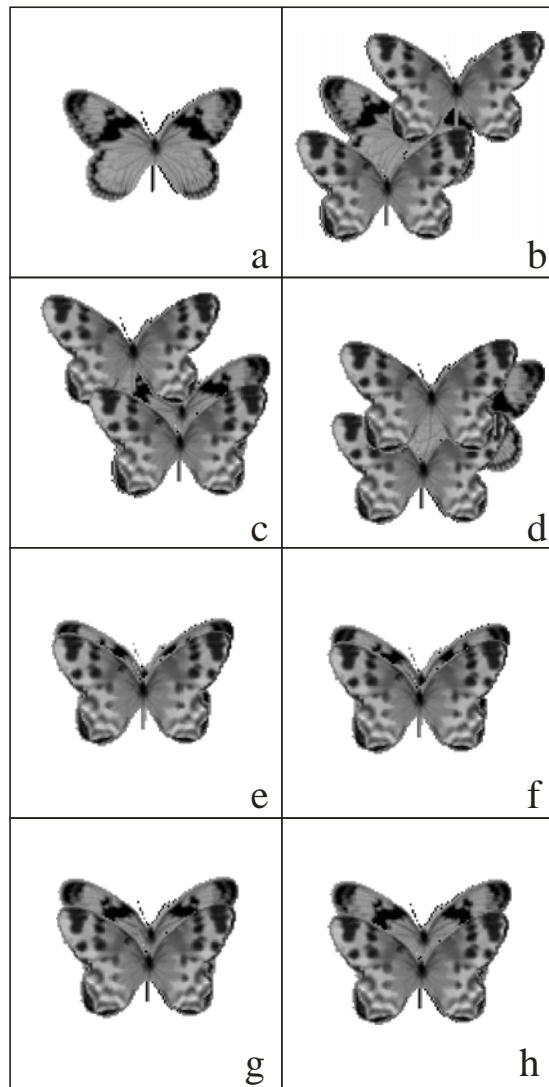


Figura 32: Objetos de prueba: (a) Objeto de referencia, (b-g) objeto de referencia parcialmente oculto por otros objetos similares.

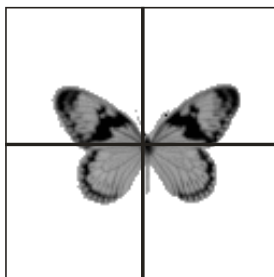


Figura 33: Objeto de referencia, dividido en cuatro fragmentos (imágenes independientes).



Figura 34: Fondo utilizado en los experimentos

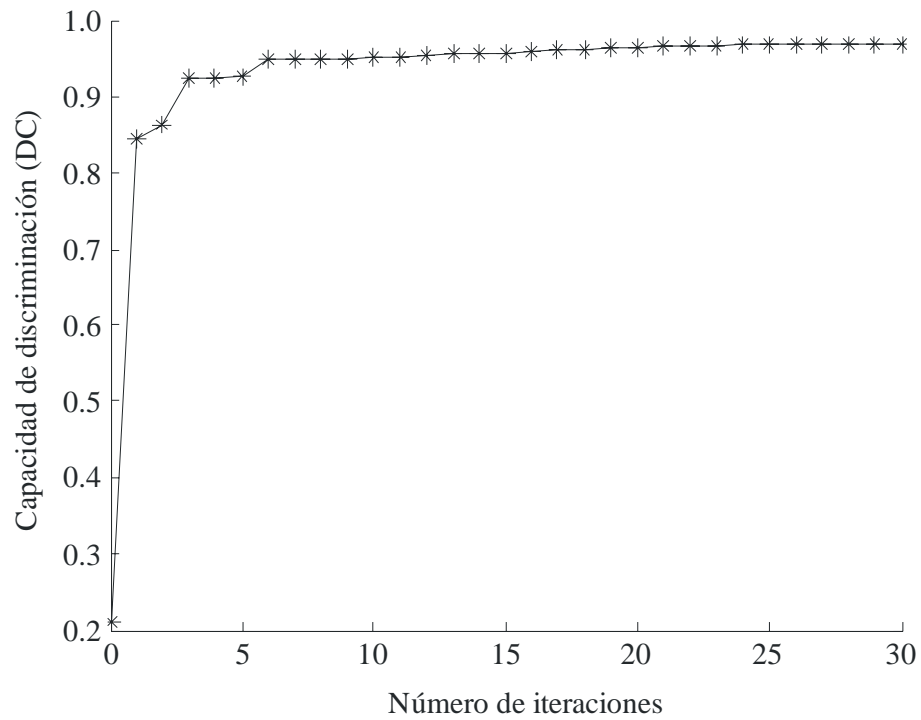


Figura 35: Desempeño del filtro Adaptativo SDF en el proceso de diseño.

Tabla VII: Desempeño de los filtros de correlación en términos de falsas alarmas (errores de detección) para las escenas de la Fig. 32

Objetos en la escena de prueba	POF	MACE	ASDF
a	7	7	0
b	30	30	0
c	18	22	0
d	6	8	0
e	1	0	0
f	0	0	0
g	0	0	0

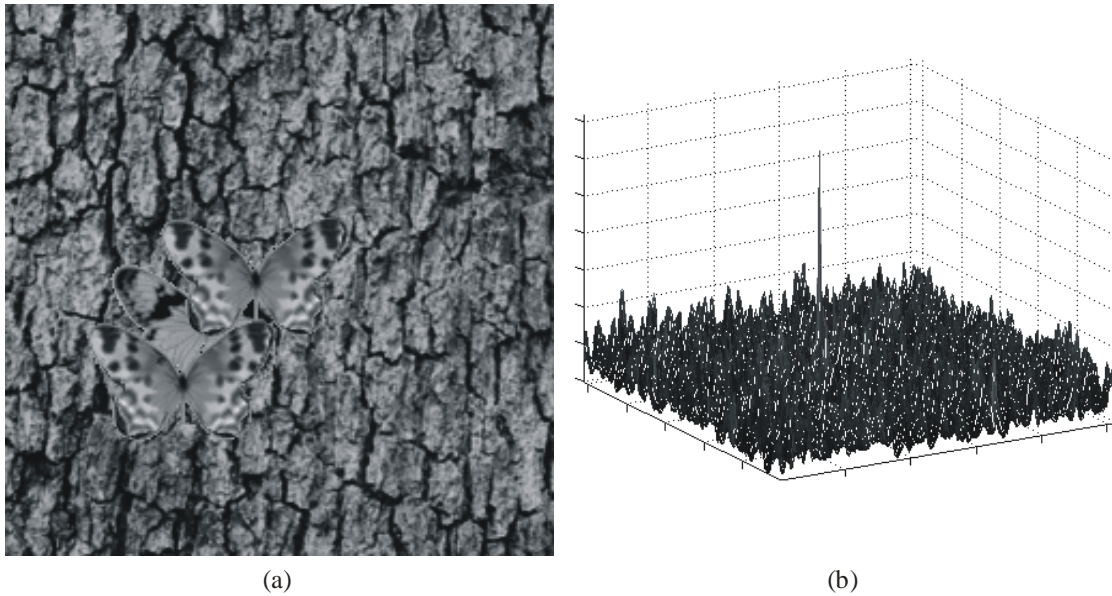


Figura 36: Escena de prueba para detectar objetos parcialmente ocultos. (a) objetos de la Fig. 32.(b) incrustados en un fondo de la realidad, (b) plano de correlación en 3D, el cual se obtuvo con el filtro *AF*.

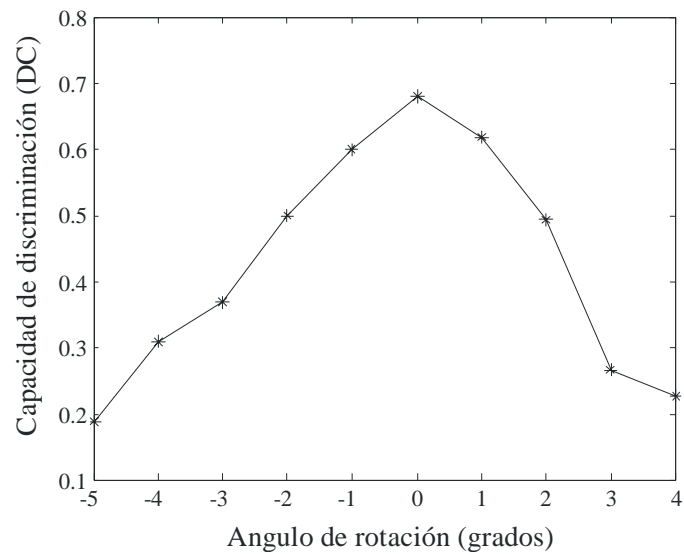


Figura 37: Tolerancia del filtro *ASDF* frente a pequeñas distorsiones geométricas: rotación.

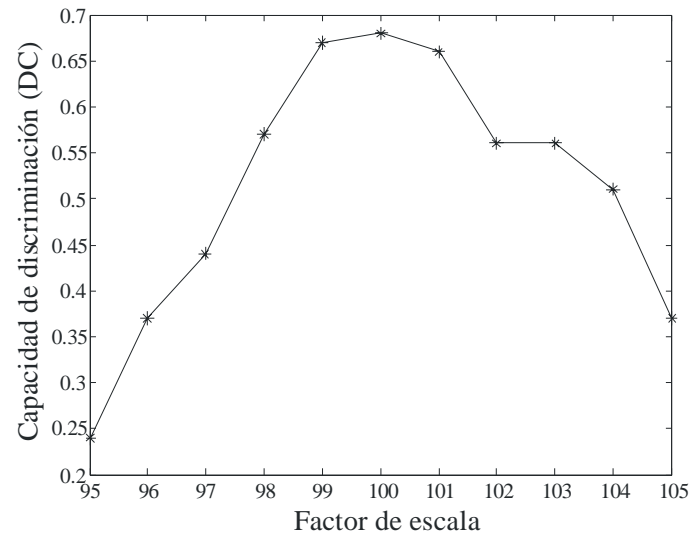


Figura 38: Tolerancia del filtro *ASDF* frente a pequeñas distorsiones geométricas: escalamiento.

Tabla VIII: Desempeño de varios filtros de correlación en términos de su *DC* para las escenas de prueba de la Fig. 32.

Objetos en la escena de prueba	Porción disponible del objeto de referencia	DC con un intervalo de confianza al 95 %
a	43 %	0.62 ± 0.014
b	29 %	0.30 ± 0.017
c	29 %	0.28 ± 0.016
d	19 %	0.26 ± 0.022
e	25 %	0.45 ± 0.017
f	38 %	0.68 ± 0.011
g	46 %	0.70 ± 0.010

VI.2. Resumen

Uno de los principales problemas en el reconocimiento de patrones y en visión de máquina, se presenta cuando los objetos aparecen parcialmente ocultos, ya que estos pierden parte de la información que los describe.

En este capítulo se presentó un algoritmo para diseñar un filtro adaptativo de correlación, el cual permite reconocer objetos parcialmente ocultos y objetos fragmentados que se encuentran incrustados en un fondo conocido. El filtro adaptativo está basado en el filtro compuesto denominado *SDF*. La información conocida, tanto del objeto de referencia como del fondo, se utiliza dentro de un procedimiento iterativo para entrenar al filtro y alcanzar una capacidad de discriminación deseada. Los resultados obtenidos con los filtros adaptativos se comparan con los resultados proporcionados por los filtros *SDF* convencional, *POF* y *MACE*. El filtro propuesto proporcionó los mejores resultados respecto a la capacidad de discriminación y presentó el menor número de errores de detección.

Cabe señalar que el problema de reconocimiento para los objetos fragmentados que se presentaron en este capítulo, se puede englobar como el reconocimiento de objetos parcialmente ocultos. Puesto que el área del fondo que circunda al fragmento oculta ciertas porciones de su composición.

Capítulo VII

Aplicaciones

En los capítulos precedentes se han descrito las técnicas para diseñar filtros basados en la correlación así como sus bases matemáticas. El objetivo de este capítulo es proporcionar algunos ejemplos del reconocimiento de patrones aplicados a problemas del mundo real. Como se ha venido mencionando, el filtrado basado en correlación tiene un uso en muchas aplicaciones diferentes, pero en los ejemplos que se presentarán se le presta principal atención al reconocimiento de partículas biogénicas.

Una de las razones que ha permitido el crecimiento en el interés de diseño de filtros de correlación, surge de la necesidad de lidiar con imágenes más complejas para sistemas automatizados de reconocimiento de imágenes y también de la necesidad de procesar imágenes demasiado grandes en tiempo real. Aún y cuando los patrones u objetos pueden presentar distorsiones geométricas (al ser capturados por medio de una cámara desde diferentes puntos de vista se pueden presentar desplazamientos, rotaciones, cambios de escala, etc.), diferencias en el contenido de la señal (textura, intensidad o contraste), y degradaciones por el ruido ya sea dependiente de la señal o independiente, se deben reconocer o clasificar como a un mismo objeto. Algunos ejemplos de tales aplicaciones son la visión por computadora y el procesado automático de imágenes biológicas. Para algunas aplicaciones de vital importancia, las técnicas ópticas u opto-digitales híbridas permiten un procesamiento de imágenes más rápido. Esta es la razón principal de que nuestro enfoque esté basado en filtros por correlación para esta área, además de que los filtros poseen buenos fundamentos matemáticos (Horner y Gianino, 1984) y pueden implementarse eficazmente de manera digital u opto-digital (Moreno

et al., 1998).

En este trabajo se consideran algunas aplicaciones de los filtros adaptativos de correlación que hemos propuesto (Gonzalez-Fraga *et al.*, 2005a, 2006) para el procesamiento de datos biológicos. El desempeño de los filtros adaptativos en las escenas se compara con otros filtros por correlación en términos de la capacidad de la discriminación y de la robustez al ruido.

VII.1. El problema de detección

Existe una gran cantidad de partículas biogénicas las cuales son muy útiles para determinar el estado en que se encuentra un ecosistema. Estas pueden mostrar cualquier desequilibrio provocado por la intrusión de algún agente extraño al sistema. Como algunos ejemplos de esta definición se pueden mencionar la bacteria responsable de la enfermedad del cólera, la bacteria de la tuberculosis, cuerpos de inclusión de virus en tejido de camarón y parásitos en peces, sólo por mencionar algunos, sin embargo cualquier entidad biológica que cumpla con la definición anterior puede ser considerada una partícula biogénica (Pech-Pacheco y Álvarez-Borrego, 1998). En particular el plancton (fitoplancton y zooplancton) es una de las partículas más importantes y es indicador del estado ecológico del planeta. Pueden ser utilizados para identificar regiones naturales de los océanos, las cuales a su vez pueden ser caracterizadas por una especie típica o por un conjunto de especies. En esta sección se estudian dos problemas de reconocimiento para este tipo de microorganismos, los cuales de manera general son llamados copépodos (pertenecientes al zooplancton). Se presentan algunos resultados obtenidos con los filtros adaptativos *SDF* por medio de simulaciones por computadora. Los resultados son comparados con aquellos que arrojan los filtros *POF*, *OF*, *MACE* y el *SDF* convencional.

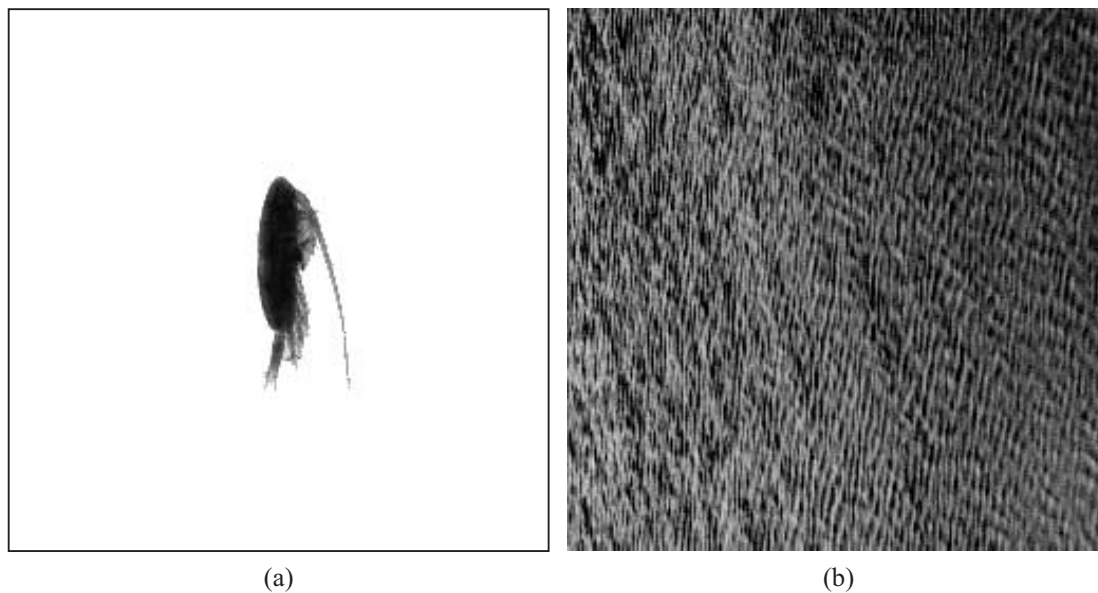


Figura 39: Imágenes de prueba: (a) Copépodo macho como el objeto a reconocer. (b) Fondo a rechazar.

Para ilustrar el diseño y el desempeño del filtro *ASDF* en un problema de detección, el cual se define como “encontrar al objeto de referencia en una posición arbitraria dentro de la escena”, se usó el copépodo macho y el fondo que se muestran en la Fig. 39. El tamaño de todas las imágenes usadas en los experimentos que se muestran en esta sección es de 256×256 píxeles. El rango de la señal se encuentra en $[0 - 255]$ (imágenes en escala de gris).

El tamaño del pico en la salida de los filtros *SDF* y *CMF* dependen de la cantidad de energía que contengan los objetos. Para tratar de homogeneizar la cantidad de energía en todos los objetos involucrados se utilizó una región de soporte con la forma de la unión de los objetos involucrados pero dilatada por unos cuantos píxeles, de esta manera se espera que los picos de correlación de los objetos a reconocer presenten la

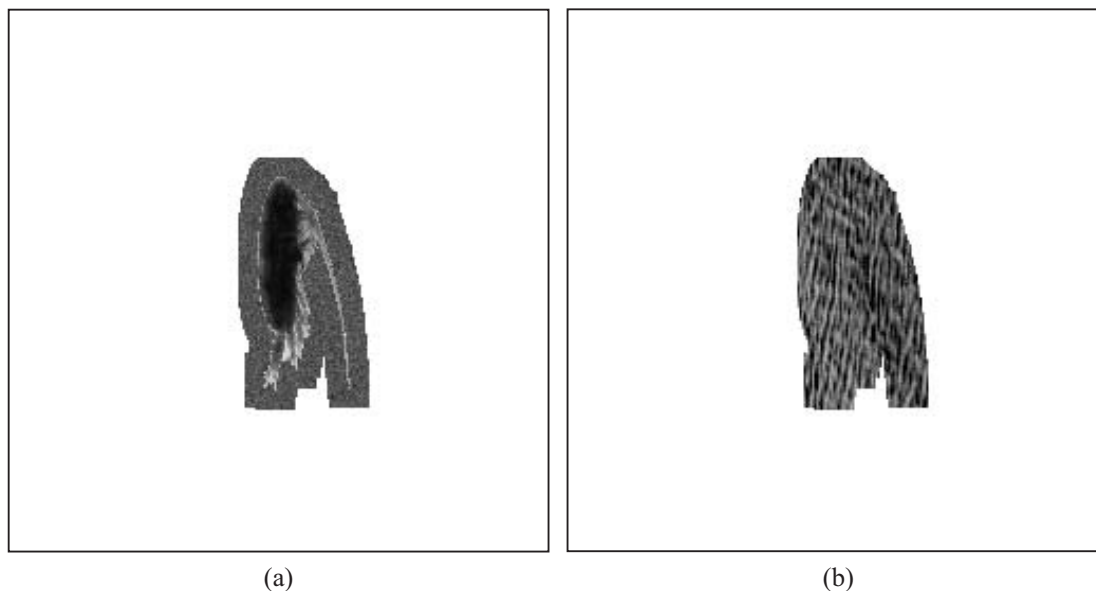


Figura 40: Objetos modificados: (a) El copépodo macho de la Fig. 39 como nuevo objeto a reconocer. (b) Objeto de rechazo del fondo, creado durante el proceso iterativo del diseño del filtro *ASDF*.

misma altura. De aquí que, para mejorar el reconocimiento dentro del entrenamiento se construye un objeto modificado al usar un campo homogéneo, creado con el promedio de la imagen de fondo. Este nuevo objeto a reconocer se vería como en la Fig. 40(a) dentro de la región de soporte, la cual es la forma del objeto dilatada por 7 píxeles.

Al iniciar la ejecución del algoritmo, la iteración 0 indicó una DC igual a -0.2354 (valores negativos indican que falló el reconocimiento), se fijó como meta alcanzar una DC igual a 0.996. En cada iteración el algoritmo selecciona de entre todos los lóbulos un pico que se suprimirá en la siguiente etapa para asegurar un comportamiento monótonamente creciente de la función DC contra el número de iteraciones, La Fig. 40(b) muestra un objeto de rechazo generado a partir del fondo. El comportamiento

creciente de la función DC se puede apreciar en la Fig. 41 en donde los valores negativos se reemplazaron por ceros. Al probar el filtro con la escena de la Fig. 42(a) se obtuvo una DC igual a 0.944 mientras que en la etapa de optimización se alcanzó una DC igual a 0.991. Se realizaron 30 pruebas colocando al objeto en diferentes coordenadas en la escena y se calculó que el intervalo de confianza del 95 % para la DC es 0.93 ± 0.01 . La Fig. 42(b) muestra la distribución de la correlación obtenida con la escena de prueba, recuérdese que la posición del pico indica la posición del objeto.

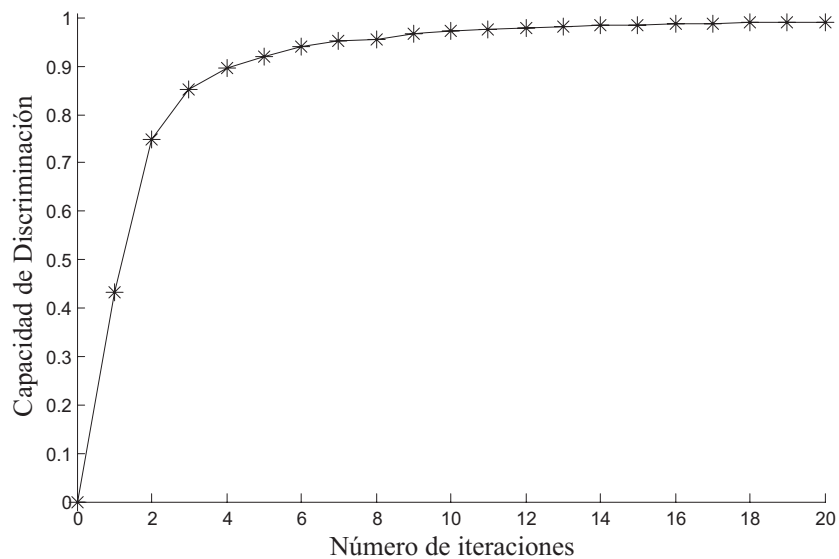
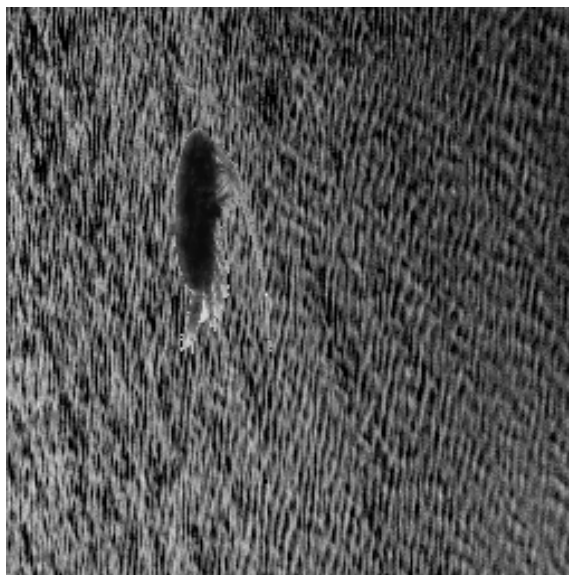
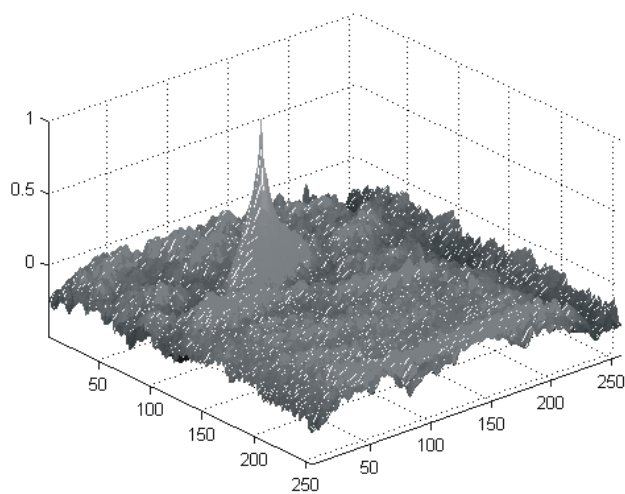


Figura 41: Desempeño del filtro $ASDF$ en la etapa de diseño.

Como un segundo ejemplo se mostrará que los filtros $ASDF$ tienen una mejor tolerancia a las distorsiones geométricas y al ruido que generalmente se presentan en los sensores de los dispositivos empleados para capturar las imágenes. Se diseñó un filtro con dos objetos, el copépodo de la Fig. 39(a) y una versión rotada 10° en sentido de las



(a)



(b)

Figura 42: Imágenes de prueba: (a) Escena con un copépodo macho como el objeto a reconocer. (b) Distribución de la salida de correlación correspondiente a la escena de prueba.

Tabla IX: Desempeño de diferentes filtros para la escena de la Fig. 43(a).

Objeto	Angulo de rotación	Factor de escala	DC				
			POF	OF	SDF	MACE	A-SDF
T_1	0.0	1.00	0.940	0.992	-0.149	0.933	0.948
T_2	8.0	1.02	-4.476	-4.194	-0.904	-0.314	0.799
T_3	-1.0	1.10	-0.052	-0.368	-0.236	-2.061	0.906
T_4	12.0	0.95	-1.901	-4.733	-0.099	-0.731	0.880
T_5	10.0	1.00	-2.945	-1.561	-0.497	0.957	0.945

manecillas del reloj (objeto T_5 de la escena de la Fig. 43). Tanto el filtro *POF* como el *OF* se diseñaron con el objeto T_1 de la Fig. 43(a) y como se podrán apreciar los resultados de la Tabla IX, estos filtros solo pueden reconocer al objetos T_1 que es el objeto para el cual están entrenados. El filtro *SDF* convencional y el *MACE* se diseñaron con los objetos T_1 y T_5 . El filtro *ASDF* también emplea los objetos T_1 y T_5 pero enmascarados con la región de soporte y el campo homogéneo. Al probar la escena de la Fig. 43, los resultados demostraron que los filtros adaptativos *ASDF* presentaron la mejor tolerancia a pequeñas distorsiones geométricas, ya que fué capaz de reconocer a todos los objetos en la escena. Mientras que el filtro *MACE* solo reconoció a los objetos con que fué entrenado y el filtro *SDF* convencional falló en todos los casos.

La Fig. 43(b) muestra la escena de la Fig. 43(a), pero distorsionada con ruido aditivo con media cero y desviación estándar 15. Se compararon varios filtros y los resultados se resumen en la Tabla X. Nuevamente, el filtro *ASDF* demostró tener mejor tolerancia ya que los picos en la salidas de los otros filtros decaen más rápidamente dando como resultado una discriminación más pobre.

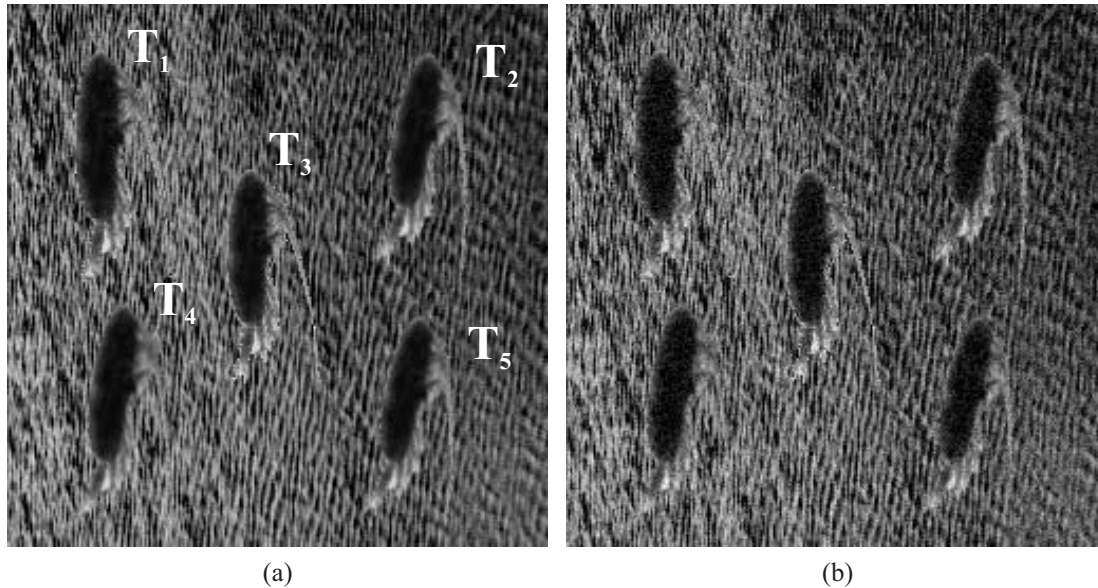


Figura 43: Imágenes de prueba: (a) T_2 , T_3 , T_4 y T_5 representan diferentes versiones del objeto T_1 que es el mismo copépodo de la Fig. 39, ver la Tabla IX para más información. (b) La escena distorsionada con ruido aditivo de naturaleza Gaussiana con media cero y desviación estándar 15.

Tabla X: Desempeño de diferentes filtros para la escena de la Fig. 43(a), alterada por ruido aditivo de media cero y desviación estándar 15.

Objeto	Angulo de rotación	Factor de escala	DC				
			POF	OF	SDF	MACE	A-SDF
T_1	0.0	1.00	0.937	0.973	-0.139	0.860	0.948
T_2	8.0	1.02	-3.548	-1.544	-0.891	-1.323	0.796
T_3	-1.0	1.10	-0.076	-0.678	-0.226	-1.408	0.904
T_4	12.0	0.95	-1.812	-3.741	-0.099	-1.396	0.880
T_5	10.0	1.00	-3.161	-2.216	-0.493	0.928	0.944



Figura 44: Imágenes de prueba: (a) Objetos a reconocer (copéodos macho). (b) Objetos a rechazar (copéodos hembra).

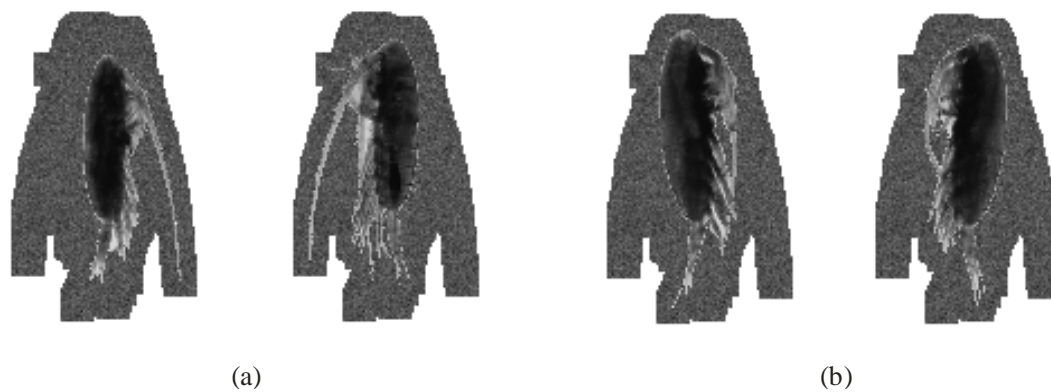


Figura 45: Imágenes de prueba: (a) Objetos modificados a reconocer (copéodos macho). (b) Objetos modificados a rechazar (copéodos hembra).

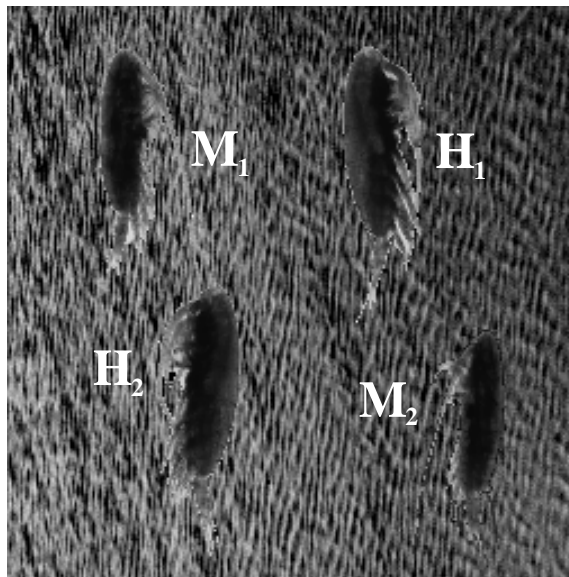


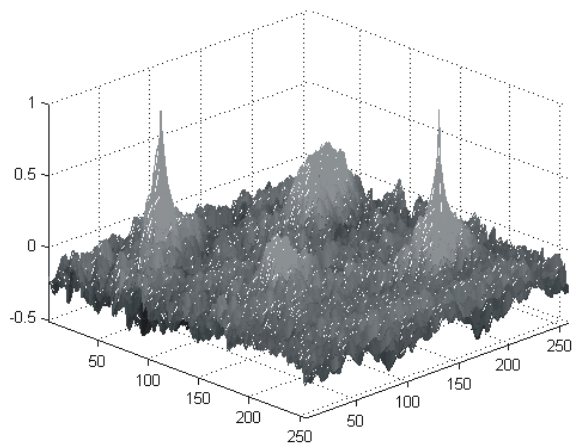
Figura 46: Escena de prueba, los copéodos macho están marcados con la letra M.

VII.2. El problema de clasificación

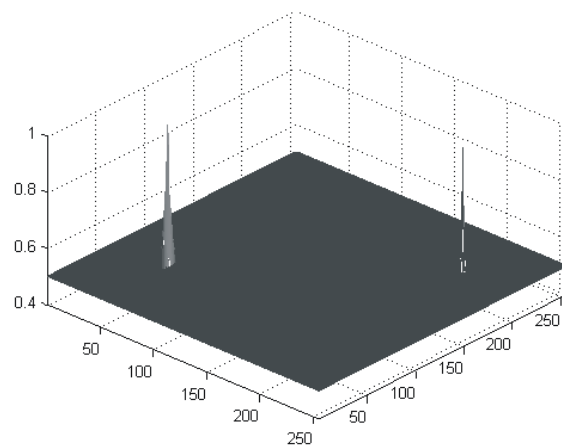
En el siguiente ejemplo se consideran como objetos a reconocer los copéodos machos de la Fig. 44(a) y las hembras de la Fig. 44(b) son objetos de rechazo así como también lo es el fondo mostrado en la Fig. 39(b). Para distinguir los machos de las hembras se diseña un filtro para cada género considerando el opuesto como objetos de rechazo.

Tal como se indicó en la sección anterior, se usa una región de soporte correspondiente a la unión de las formas de los objetos involucrados para enmascarar los nuevos objetos a buscar y los objetos de rechazo que se crean a partir del fondo. La Fig. 45 muestra como se verían estos nuevos objetos. Obsérvese que los objetos originales se encuentran incrustados en un campo homogéneo, el cual se creó al promediar diez puntos del fondo y la forma que contiene al objeto es la forma de la región de soporte indicada.

El algoritmo usó 30 imágenes de rechazo del fondo para el filtro de los machos y 25 para el filtro de las hembras. Para ambos filtros se alcanzó una DC de 0.996.



(a)



(b)

Figura 47: Distribuciones de intensidad de correlación: (a) salida de la escena de la Fig. 46. (b) salida después de aplicar un umbral del 50% del valor del pico más alto.

Para eliminar los lóbulos laterales producidos por los objetos de rechazo (machos o hembras, según sea el caso), se incorporaron 4 imágenes correspondientes a los objetos de rechazo pero desplazados a la izquierda, derecha, arriba y abajo. Estas imágenes fueron centradas en las fronteras de tales objetos y enmascaradas con la región de soporte y el campo homogéneo. Lo anterior corresponde a una optimización de los objetos que se desean reconocer contra los objetos que se desean rechazar.

La escena de prueba se muestra en la Fig. 46, la cual contiene a todos los objetos de la Fig. 44. Al probar el filtro de los machos se calcularon valores de discriminación de 0.915 y 0.925 para los machos M_1 y M_2 respectivamente. El filtro de las hembras reportó que las discriminaciones fueron 0.899 y 0.879 para las hembras H_1 y H_2 respectivamente. La Fig. 47 muestra los planos de correlación resultado de aplicar el filtro de los machos.

VII.3. Reconocimiento de fragmentos de diatomeas fósiles

Las diatomeas son de los organismos más comunes en depósitos pantanosos (los hay en el fondo del mar o de los lagos). Puesto que poseen un esqueleto siliceo que se acumula en el fondo cuando mueren, sus restos se utilizan para hacer interpretaciones ambientales. Es importante identificar los géneros y especies, así como su abundancia, presentes en los estratos formados en antiguos lagos y capas de sedimento en el mar para interpretar su ambiente natural en el pasado en términos de historia geológica y climática. Debido a esto, el estudio de estos fósiles es muy útil en la industria petrolera para determinar la presencia de este preciado combustible. Gracias a este estudio las compañías determinan si es costoso excavar en ciertos lugares.

En muchas ocasiones se encuentran estos fósiles fragmentados o pulverizados. Lo

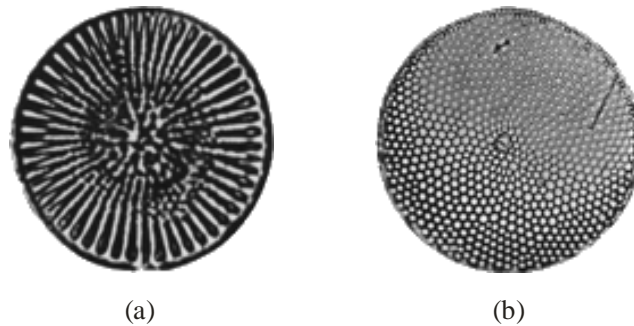


Figura 48: Diatomeas fósiles: (a) *Stephanodiscus sp* y (b) *Aspeitia nodulifer*.

cual nos puede conducir a la pérdida de información para poder ser clasificados correctamente. El nombre científico de las diatomeas mostradas en la Fig. 48 fueron tomados del trabajo de Villalobos-Flores *et al.* (2002). En el ejemplo que se proporciona en la Fig. 49 se muestra una imagen típica (a nivel microscopio) con fragmentos de la diatomea denominada “*Stephanodiscus sp*”. La diatomea que aparece completa (objeto no deseado) en la imagen se catalogó como “*Aspeitia nodulifer*” aparece obstruyendo otro fragmento del objeto que se desea reconocer y clasificar (Villalobos-Flores *et al.*, 2002). Se diseñó un filtro ASDF con fragmentos de “*Stephanodiscus sp*”. También se utilizó el fondo y el objeto no deseado en el diseño del filtro para llevar a cabo un reconocimiento exitoso. La Fig. 49(c) muestra la salida de correlación resultante, la cual también se graficó en 3D. Para automatizar se requiere solamente de definir un umbral en la salida de correlación, aquellos puntos que excedan tal umbral corresponderán a los objetos que se desean identificar.

VII.4. Resumen

El reconocimiento adaptativo de patrones continúa en un estado de rápida evolución. Los filtros de correlación son populares gracias a que pueden realizar las tareas

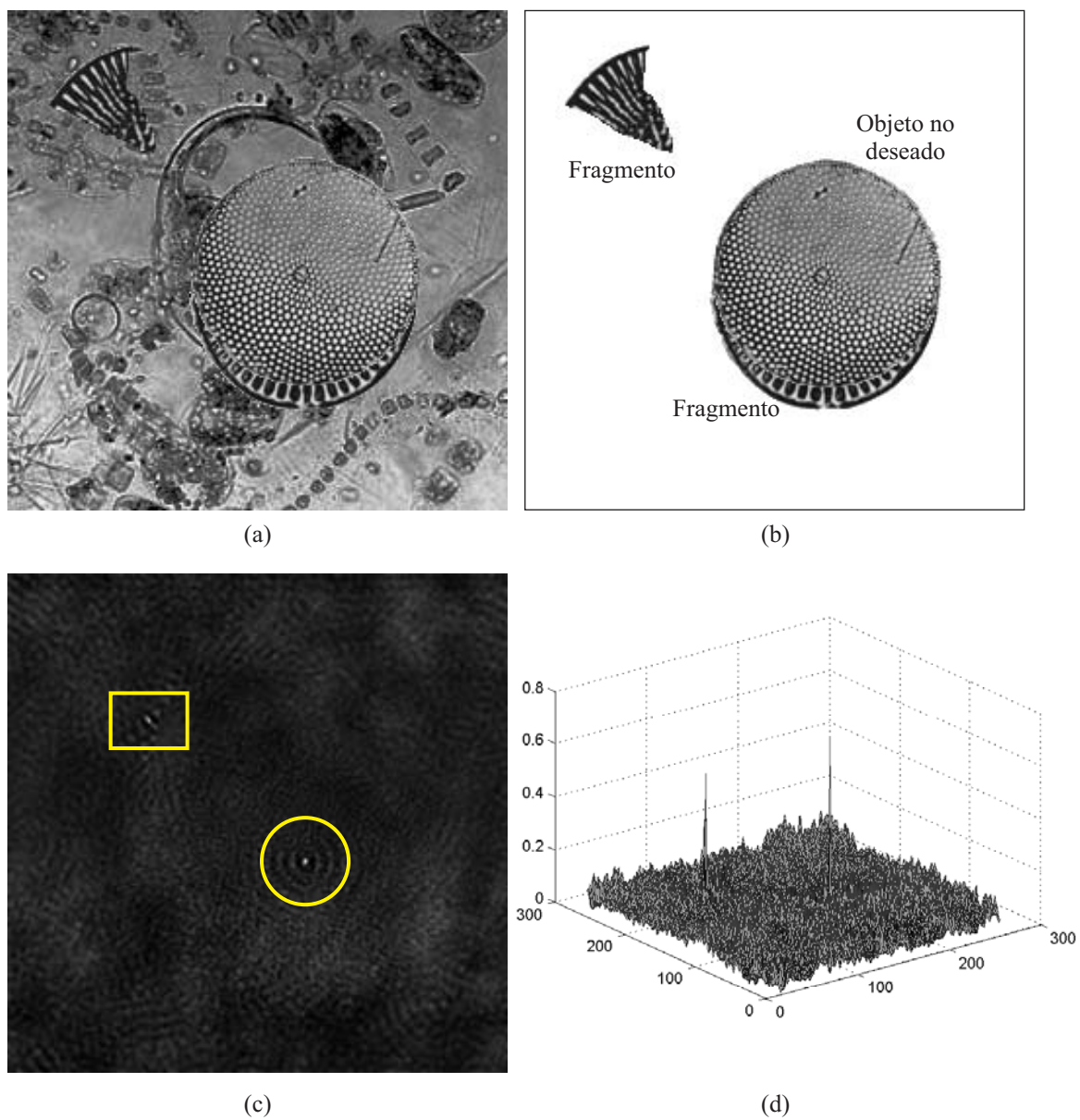


Figura 49: Imagen de prueba: (a) Escena de prueba, (b) posición de los objetos a reconocer, (c) salida de la correlación con el filtro *ASDF* y (d) gráfica 3D de la salida de correlación.

de localización e identificación sin necesidad de segmentar o de preprocesar la escena a analizar. En este capítulo se han propuesto sistemas digitales diseñados en base a filtros adaptativos para mejorar el reconocimiento de objetos biológicos que pueden encontrarse en un fondo real. Se demostró que los algoritmos de diseño de los filtros pueden ayudar a tomar el control sobre todo el plano de correlación con pocas iteraciones para el entrenamiento. Los sistemas digitales están basados en un entrenamiento iterativo de filtros *SDF* y como consecuencia, poseen los beneficios de tener una buena tolerancia a las distorsiones geométricas y al ruido aditivo. Los sistemas digitales pueden implementarse fácilmente en una computadora para realizar el reconocimiento de patrones. Las simulaciones por computadora demostraron que los filtros propuestos para el reconocimiento de datos biológicos poseen un desempeño muy bueno comparado con aquellos filtros por correlación clásicos. Los filtros recomendados poseen una buena adaptividad a la escena y una buena robustez al ruido de entrada.

Capítulo VIII

Conclusiones y áreas de oportunidad

El uso de filtros convencionales de correlación para situaciones de la realidad, muchas veces resulta ser impráctico. Lo anterior puede deberse a diversos factores, como que el objeto a identificar presente alguna deformación, que los factores de iluminación no sean adecuados, que se presente algún tipo de ruido en la captura de la escena, etc. Este trabajo de investigación estuvo motivado por encontrar y desarrollar nuevos métodos por correlación para efectuar el reconocimiento confiable de objetos, en particular, cuando los objetos que se desean reconocer se encuentran parcialmente ocultos en una escena real. Este problema es de mucha importancia en varias áreas como la robótica, visión de máquina e inteligencia artificial.

En este manuscrito se mostró que el reconocimiento de objetos en imágenes por medio de la correlación corresponde a un estimador estadístico óptimo, por lo que es bastante utilizado y confiable. También se mostró que los filtros clásicos de correlación no funcionan correctamente bajo el modelo de imagen disjunto (cuando el objeto se incrusta en un fondo).

Existen diversos filtros y técnicas de transformación que optimizan las métricas de desempeño, sin embargo, los filtros optimizados dentro de un enfoque adaptativo presentan mejoras en escenas particulares. Hay diferentes técnicas para introducir invariancias a los filtros, pero la que mejor se adaptó a las necesidades del problema de reconocer objetos parcialmente ocultos fueron las funciones discriminantes sintéticas

o filtros *SDF*.

La aportación de este trabajo se centra en el desarrollo de nuevos filtros adaptativos de correlación. Estos filtros, a los que denominamos *ASDF* debido a que están basados en el filtro *SDF* convencional, resultaron tener características de desempeño muy superior a los filtros convencionales. Para diseñar un filtro *ASDF*, la información que se conoce tanto del objeto de referencia como de los objetos no deseados y del fondo a rechazar, se utilizó en un proceso de entrenamiento y de manera iterativa para que nos permita alcanzar un valor de discriminación determinado.

Se presentaron y discutieron algunas simulaciones por computadora en las que se demostró que el filtro *ASDF* propuesto tiene el mejor desempeño en términos de capacidad de discriminación y de errores de detección frente a los filtros convencionales. También presenta la mejor tolerancia al ruido aditivo y se mostró como se le pueden incorporar invariancias al hacer uso de imágenes de entrenamiento. Principalmente se consideraron las invariancias a la rotación y al escalamiento, pero es posible de considerar cualquier tipo de distorsión (como por ejemplo tener objetos fragmentados).

Se usó una versión del algoritmo utilizado para la generación del filtro *ASDF* básico para diseñar filtros adaptativos que nos permiten detectar y clasificar objetos parcialmente ocultos y fragmentados. Aún y cuando la escena analizada presente ruido aditivo y sea una escena complicada. Es importante recalcar que la mayoría de los filtros convencionales solo pueden controlar un solo punto en la salida de la correlación. Por medio de simulaciones, se demostró que los algoritmos de diseño de los filtros pueden tomar todo el control sobre el plano de correlación con unas cuantas iteraciones.

Los filtros propuestos tienen capacidad para reconocer y discriminar entre objetos similares aún y cuando la información disponible del objeto sea de alrededor del 20 % (en el caso de los objetos fragmentados y parcialmente ocultos).

Finalmente, nótese que el algoritmo iterativo propuesto para el diseño del filtro se

puede emplear para mejorar el desempeño de cualquier filtro de correlación al adaptar el filtro a la escena de entrada.

El diseño de los filtros *ASDF* se puede formular como un problema de optimización, por lo que es deseable explorar otras heurísticas, como los algoritmos genéticos, para la selección de las mejores imágenes de rechazo. Otra alternativa que se puede contemplar cuando se tienen imágenes muy grandes, como suele suceder en las aplicaciones reales, es utilizar algoritmos paralelos en el diseño del filtro

Bibliografía

- Ahouzi, E., J. Campos, y M. Yzuel 1994. “Phase-only filter with improved discrimination”. *Optics Letters*, 18(17):1340–1342 p.
- Arsenault, H. H., T. Szoplik, y B. Macukow 1989. “Optical processing and computing”. Birkhauser. Primera Edición. Boston. 493 p.
- Billet, O. y L. Singher 2002. “Adaptive multiple filtering”. *Optical Engineering*, 41(1):55–68 p.
- Bovik, A. 2000. “Handbook of image and video processing”. Academic Press. Primera Edición. San Diego. 891 p.
- Campos, J., K. Styczynski, M. J. Yzuel, y K. Chalasinska-Macukow 1994. “Recognition of partially occluded objects by correlation methods”. *Optics Communications*, 106(1–3):45–51 p.
- Casasent, D. 1984. “Unified synthetic discriminant function computational formulation”. *Applied Optics*, 23(10):1620–1627 p.
- Casasent, D. y D. Psaltis 1976a. “Position, rotation, and scale invariant optical correlation”. *Applied Optics*, 15(7):1795–1799 p.
- Casasent, D. y D. Psaltis 1976b. “Scale invariant optical correlation using mellin transforms”. *Optics Communications*, 17(1):59–63 p.
- Casasent, D. y D. Psaltis 1976c. “Scale invariant optical transform”. *Optical Engineering*, 15(3):258–261 p.
- Castro-Longoria, E., J. Álvarez-Borrego, y J. L. Pech-Pacheco 2001. “Identification of species of calanoid copepods using a new invariant correlation algorithm”. *Crustaceana*, 74(10):1029–1039 p.

- Duda, R. O., P. E. Hart, y D. G. Stork 2001. "Pattern classification". John Wiley and Sons, Inc., Segunda Edición, New York. 654 p.
- Garcia-Martinez, P., J. Garcia, y C. Ferreira 1995. "A new criterion for determining the expansion center for circular-harmonic filters". *Optics Communications*, 117(5-6):399–405 p.
- Gonzalez, R. C. y R. E. Woods 2002. "Digital image processing". Prentice Hall, Segunda Edición. 793 p.
- Gonzalez-Fraga, J., V. Diaz-Ramirez, V. Kober, y J. Alvarez-Borrego 2005a. "Improving the discrimination capability with an adaptive synthetic discriminant function filter". *Lecture Notes in Computer Science*, 3523:83–90 p.
- Gonzalez-Fraga, J., V. Kober, y J. Alvarez-Borrego 2005b. "Recognition of partially occluded objects using correlation filters with training". En: Tescher, A. G., (ed.). "Proceedings of SPIE. Applications of Digital Image Processing XXVIII", San Diego, California, USA. SPIE-The International Society for Optical Engineering. 59091X(1–7) p.
- Gonzalez-Fraga, J., V. Kober, y J. Alvarez-Borrego 2006. "Adaptive synthetic discriminant function filters for pattern recognition". *Optical Engineering*, 45(5):057005(1–10) p.
- González-Fraga, J. A. 2004. "Reconocimiento de objetos parcialmente ocultos". Tesis de Maestría, CICESE, México. 117 p.
- Horner, J. L. y P. D. Gianino 1984. "Phase-only matched filtering". *Applied Optics*, 23(6):812–816 p.
- Hsu, Y. y H. Arsenault 1982. "Optical pattern recognition using circular harmonic expansion". *Applied Optics*, 21(222):4016–4019 p.
- Jain, A. 1986. "Fundamentals of digital image processing". Prentice Hall. Primera Edición. 569 p.

- Javidi, B., R. Ponce-Diaz, y S. Hong 2006. "Three-dimensional recognition of occluded objects by using computational integral imaging". *Optics Letters*, 31:1106–1108 p.
- Javidi, B. y J. Wang 1994. "Design of filters to detect a noisy target in nonoverlapping background noise". *Journal OSA (A)*, 11:2604–2612 p.
- Khoury, J., P. Gianino, J. Kane, y C. Woods 1998a. "Near-optimal binary phase-only filter correlator for obscured inputs". *Optical Engineering*, 37(1):123–131 p.
- Khoury, J., P. Gianino, y C. Woods 1998b. "Adaptive optimal filters for correlators operating on obscured inputs". *Optical Engineering*, 37(1):112–122 p.
- Khoury, J., P. Gianino, y C. Woods 2002a. "Algorithms for correlating severely obscured images in nonoverlapping zero-mean noise". *Optical Engineering*, 41(1):17–25 p.
- Khoury, J., P. Gianino, y C. Woods 2002b. "Effect of disjoint illumination in optimal correlation filters with obscured inputs". *Optical Engineering*, 41(1):118–125 p.
- Kober, V. y J. Campos 1996. "Accuracy of location measurement of noisy target in a nonoverlapping background". *J. Opt. Soc. Am. A.*, 13(8):1653–1666 p.
- Kober, V., L. Yaroslavsky, J. Campos, y M. Yzuel 1994. "Optimal filter approximation by means of a phase only filter with quantization". *Optics Letters*, 19(13):978–980 p.
- Mahalanobis, A., B. Vijaya-Kumar, y D. Casasent 1987. "Minimum average correlation energy filters". *Applied Optics*, 26(17):3633–3640 p.
- Mardia, K. V., W. Quian, D. Shah, y K. M. A. de Souza 1997. "Deformable template recognition of multiple occluded objects". *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19:1035–1042 p.
- Michalewicz, Z. y D. B. Fogel 2000. "How to solve it: Modern Heuristics". Springer. Berlin. 467 p.
- Moreno, I., J. Campos, M. Yzuel, y V. Kober 1998. "Implementation of bipolar real-valued input scenes in a real-time optical correlator: application to color pattern recognition". *Optical Engineering*, 37(1):144–150 p.

- Papoulis, A. 1962. "The fourier integral and its applications". McGraw-Hill Book Company, Inc., Primera Edición, New York. 318 p.
- Pech-Pacheco, J. L. y J. Álvarez-Borrego 1998. "Optical-digital system applied to the identification of five phytoplankton species". *Marine Biology*, 132(3):357–365 p.
- Press, W. H., B. P. Flannery, S. A. Teukolsky, y W. T. Vetterling 1988. "Numerical Recipes in C - The Art of Scientific Computing". Cambridge University Press. Primera Edición. New York. 735 p.
- Sekuler, A. B., J. M. Gold, R. F. Murray, y P. J. Bennett 2000. "Visual completion of partly occluded objects: Insights from behavioral studies". *Neuro-Ophthalmology*, 23(3,4):165–168 p.
- Suetens, P., P. Fua, y A. J. Hanson 1992. "Computational strategies for object recognition". *ACM Computing Surveys*, 24(1):5–61 p.
- Van Trees, H. L. 1968. "Detection, estimation and modulation theory". John Wiley and Sons, Inc. New York. 697 p.
- VanderLugt, A. 1964. "Signal detection by complex filters". *IEEE Trans. Inf. Theory*, 10(6):139–145 p.
- Vanderlugt, A. 1992. "Optical signal processing". John Wiley and Sons, Inc. New York. 603 p.
- Vijaya-Kumar, B. y L. Hassebrook 1990. "Performance measures for correlation filters". *Applied Optics*, 29(20):2997–3006 p.
- Vijaya-Kumar, B. V. K. 1986. "Minimum-variance synthetic discriminant functions". *J. Opt. Soc. Am. A.*, 3(10):1579–1584 p.
- Vijaya-Kumar, B. V. K. 1992a. "Tutorial survey of composite filter designs for optical correlators". *Applied Optics*, 31(23):4773–4801 p.

- Vijaya-Kumar, B. V. K. 1992b. "Tutorial survey of composite filter designs for optical correlators". *Applied Optics*, 31(23):4773–4801 p.
- Vijaya-Kumar, B. V. K. 1994. "Optical pattern recognition". En: Javidi, B. y J. L. Horner, (eds.). "Real-Time Optical information processing". Academic Press Limited, San Diego, 39–88 p.
- Villalobos-Flores, C. E., J. Alvarez-Borrego, V. Kober, G. Cristóbal, y E. Castro-Longoria 2002. "Study of 21 fragmented fossil diatoms using a digital invariant correlation". En: "The International Symposium on Optical Science and Technology", Seattle Washington. SPIE-The International Society for Optical Engineering. 528–533 p.
- Villalobos-Flores, C. E., J. Alvarez-Borrego, J. L. Pech-Pacheco, G. Cristóbal, y E. Castro-Longoria 2001. "Study of fragmented fossil diatoms using an invariant correlation method". En: Brudny, V. L., S. A. Ledesma, y M. C. Marconi, (eds.). "Proceedings 4th Iberoamerican Meeting on Optics and 7th Latin American Meeting on Optics, Lasers and their Applications", Tandil, Argentina (del 3 al 7 de septiembre). 18–21 p.
- Wiskott, L. y C. V. D. Malsburg 1993. "A neural system for the recognition of partially occluded objects in cluttered scenes (A Pilot Study)". *Int. J. of Pattern Recognition and Artificial Intelligence*, 7(4):935–948 p.
- Wozencraft, J. M. y I. M. Jacobs 1965. "Principles of Communication Engineering". John Wiley and Sons. Primera Edición. New York. 720 p.
- Yaroslavsky, L. y M. Eden 1996. "Fundamentals of digital optics: Digital signal processing in optics and holography". Birkhäuser, Inc. Primera Edición. Boston. 362 p.
- Yaroslavsky, L. 1993. "The theory of optimal methods for localization of objects in pictures". En: "in progress in Optics XXXII, E. Wolf, Ed., Elsevier". 145–201 p.

Apéndice A

Ruido aditivo

Uno de los modelos más utilizados en la observación de señales es el modelo aditivo de ruido Gaussiano, representado por:

$$s(x) = t(x) + n(x), \quad (106)$$

donde la señal original $t(x)$ es independiente del ruido $n(x)$, y la señal afectada por el ruido es $s(x)$. El ruido $n(x)$ se considera con media cero y varianza σ^2 , con una densidad de distribución Gaussiana.

Recuérdese que una variable aleatoria x de media μ y desviación estándar σ tiene una densidad de probabilidad Gaussiana o normal, si su función de probabilidad es de la forma:

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{(x - \mu)^2}{2\sigma^2} \right]. \quad (107)$$

Los parámetros de una distribución normal en la ecuación anterior, satisfacen las siguientes propiedades:

$$\begin{aligned} E[1] &= \int_{-\infty}^{\infty} p(x) dx, \\ E[x] &= \int_{-\infty}^{\infty} xp(x) dx = \mu, \\ E[(x - \mu)^2] &= \int_{-\infty}^{\infty} (x - \mu)^2 p(x) dx = \sigma^2. \end{aligned} \quad (108)$$

donde el operador $E[\cdot]$ es la esperanza matemática.

A.1. Generación de ruido con distribución Gaussiana

Para generar ruido de naturaleza Gaussiana con una media $\mu = 0$ y una desviación estándar $\sigma = 1$ se emplea la rutina *gasdev* que se describe en Press *et al.* (1988). Si se desea generar un valor con distribución Gaussiana y con una desviación estándar y una media dadas, solo se multiplica la variable por dicha desviación y se suma la media deseada. Para afectar una imagen con ruido, primero se genera una imagen que contenga sólo ruido de las dimensiones de la original, posteriormente se suman pixel a pixel.

A.1.1. Ruido correlacionado

Puede existir correlación entre los pixeles que conforman la matriz del ruido aditivo en una imagen. El modelo de función de correlación del ruido utilizado en los experimentos presentados en esta tesis se expone en (Jain, 1986) y se calcula a partir de la siguiente función de covarianza:

$$r(m_1, n_1, m_2, n_2) = \sigma^2 \rho_V^{|m_1 - m_2|} \rho_H^{|n_1 - n_2|}, \quad |\rho_V| < 1, |\rho_H| < 1, \quad (109)$$

donde σ^2 representa la varianza del campo aleatorio, ρ_V y ρ_H indican la correlación entre los pixeles adyacentes en la dirección vertical y horizontal respectivamente. Los valores (m_1, n_1) y (m_2, n_2) son las coordenadas de los pixeles que se encuentran correlacionados. Si la media μ de las variables aleatorias es igual a cero, entonces la covarianza se convierte en correlación, es decir:

$$Cov[X(m, n)X(k, l)] = E[(X(m, n) - \mu)(X(k, l) - \mu)] = E[X(m, n)X(k, l)], \quad (110)$$

donde $X(i, j)$ denota al pixel ubicado en la coordenada (i, j) .

Considérese la siguiente configuración de pixeles vecinos:

$X(i-1, j-1)$	$X(i-1, j)$.
$X(i, j-1)$	$X(i, j)$.
.	.	.

Utilizando el modelo de covarianza de la ecuación (109) y con $\sigma^2 = 1$, la correlación entre $X(i, j)$ con los tres vecinos restantes se expresa por el siguiente conjunto de ecuaciones:

$$\begin{aligned}
 E[X(i, j)X(i, j)] &= 1, \\
 E[X(i, j)X(i-1, j)] &= \rho_V, \\
 E[X(i, j)X(i, j-1)] &= \rho_H, \\
 E[X(i, j)X(i-1, j-1)] &= \rho_V\rho_H, \\
 E[X(i-1, j)\xi] &= 0, \\
 E[X(i, j-1)\xi] &= 0, \\
 E[X(i-1, j-1)\xi] &= 0, \\
 E[\xi\xi] &= 1,
 \end{aligned} \tag{111}$$

donde ξ es una variable aleatoria con densidad de probabilidad Gaussiana de media $\mu = 0$ y varianza $\sigma^2 = 1$ y es independiente de los tres vecinos de $X(i, j)$. Para generar ξ se emplea la rutina *gasdev*. La dependencia de $X(i, j)$ con sus tres vecinos y con ξ se puede expresar como

$$X(i, j) = AX(i-1, j) + BX(i, j-1) + CX(i-1, j-1) + D\xi. \tag{112}$$

La ecuación anterior se multiplica a su vez por $X(i, j)$, $X(i-1, j)$, $X(i, j-1)$ y $X(i-1, j-1)$, con la ayuda de las ecuaciones (111) y aplicando el operador valor

esperado, se obtiene el sistema de ecuaciones:

$$\begin{aligned}
 E[X(i, j)X(i, j)] &= A\rho_V + B\rho_H + C\rho_V\rho_H + D^2 = 1, \\
 E[X(i, j)X(i-1, j)] &= A + B\rho_V\rho_H + C\rho_H = \rho_V, \\
 E[X(i, j)X(i, j-1)] &= A\rho_V\rho_H + B + C\rho_V = \rho_H, \\
 E[X(i, j)X(i-1, j-1)] &= A\rho_H + B\rho_V + C = \rho_V\rho_H.
 \end{aligned} \tag{113}$$

Se resuelve el sistema para:

$$\begin{aligned}
 A &= \rho_V, \\
 B &= \rho_H, \\
 C &= -\rho_V\rho_H, \\
 D &= \sqrt{(1 - \rho_V^2)(1 - \rho_H^2)},
 \end{aligned} \tag{114}$$

y reemplazando estos coeficientes en la ecuación (112) finalmente se obtiene la expresión que se utiliza para generar ruido con estructura de covarianza:

$$X(i, j) = \rho_V X(i-1, j) + \rho_H X(i, j-1) - \rho_V\rho_H X(i-1, j-1) + \sqrt{(1 - \rho_V^2)(1 - \rho_H^2)}\xi. \tag{115}$$

En la Figura 50 se presenta un ejemplo de función de correlación con parámetros $\rho_V = 0.95$ y $\rho_H = 0.95$, también se presenta su respectiva función de densidad espectral, es decir su *FT*.

Algunos ejemplos de imágenes afectadas por ruido aditivo Gaussiano se muestran en la Figura 51. Para generar estas imágenes primero se generó una matriz bidimensional de ruido correlacionado, empleando la ecuación (115), cuyos valores se adicionaron a la imagen original pixel a pixel.

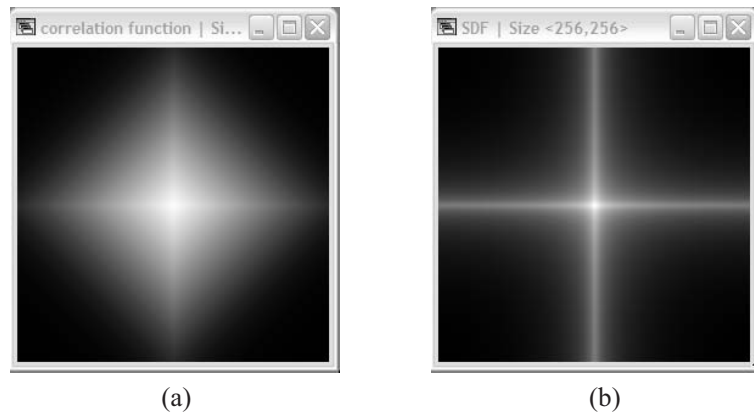


Figura 50: (a) Función de correlación de ruido con parámetros $\rho_V = 0.95$, $\rho_H = 0.95$ y (b) Función de densidad espectral de (a).



Figura 51: Imágenes afectadas con ruido aditivo de diferentes valores de correlación. El ruido es de media cero y desviación estándar de 20.