



# INSTITUTO POLITECNICO NACIONAL

CENTRO DE INVESTIGACIÓN Y DESARROLLO  
DE TECNOLOGÍA DIGITAL



MAESTRÍA EN CIENCIAS EN SISTEMAS DIGITALES

"DISEÑO DE FILTROS DIGITALES PARA EL RECONOCIMIENTO Y LA  
MULTI-CLASIFICACIÓN DE OBJETOS EN ESCENAS REALES"

TESIS

QUE PARA OBTENER EL GRADO DE  
MAESTRO EN CIENCIAS EN SISTEMAS DIGITALES

PRESENTA:

ING. OLIVER GAMALIEL CAMPOS TRUJILLO

BAJO LA DIRECCIÓN DE:  
DR. VÍCTOR HUGO DÍAZ RAMÍREZ Y  
DR. JOSÉ CRUZ NÚÑEZ PÉREZ

DICIEMBRE DEL 2011

TIJUANA, B. C., MÉXICO



# INSTITUTO POLITÉCNICO NACIONAL SECRETARÍA DE INVESTIGACIÓN Y POSGRADO

## ACTA DE REVISIÓN DE TESIS

En la Ciudad de Tijuana, B.C. siendo las 15:00 horas del día 6 del mes de diciembre del 2011 se reunieron los miembros de la Comisión Revisora de Tesis, designada por el Colegio de Profesores de Estudios de Posgrado e Investigación de CITEDI para examinar la tesis titulada:

**“DISEÑO DE FILTROS DIGITALES PARA EL RECONOCIMIENTO Y LA MULTI-CLASIFICACIÓN DE OBJETOS EN ESCENAS REALES”.**

Presentada por el alumno:

**CAMPOS**

Apellido paterno

**TRUJILLO**

Apellido materno

**OLIVER GAMALIEL**

Nombre(s)

Con registro: 

B	0	9	1	7	0	4
---	---	---	---	---	---	---

aspirante de:

**MAESTRÍA EN CIENCIAS EN SISTEMAS DIGITALES**

Después de intercambiar opiniones, los miembros de la Comisión manifestaron **APROBAR LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

### LA COMISIÓN REVISORA

Directores de tesis

DR. VÍCTOR HUGO DÍAZ RAMÍREZ

DR. JOSÉ CRUZ NÚÑEZ PÉREZ

DR. ROBERTO SEPÚLVEDA CRUZ



M. C. ERNESTO EDUARDO QUIROZ MORONEZ

M. C. ANDRÉS CALVILLO TÉLLEZ

S. E. P.  
INSTITUTO POLITÉCNICO NACIONAL  
CENTRO DE INVESTIGACION Y DESARROLLO  
DE TECNOLOGÍA DIGITAL  
DIRECCIÓN

PRESIDENTE DEL COLEGIO DE PROFESORES

DR. LUIS ARTURO GONZÁLEZ HERNÁNDEZ



**INSTITUTO POLITÉCNICO NACIONAL**  
**SECRETARÍA DE INVESTIGACIÓN Y POSGRADO**

*CARTA CESIÓN DE DERECHOS*

En la Ciudad de Tijuana, B.C. el día 9 del mes diciembre del año 2011, el (la) que suscribe Oliver Gamaliel Campos Trujillo alumno (a) del Programa de Maestría en ciencias en sistemas digitales con número de registro BO91704, adscrito a Centro de Investigación y Desarrollo de Tecnología Digital, manifiesta que es autor (a) intelectual del presente trabajo de Tesis bajo la dirección de Dr. Víctor H. Díaz Ramírez y Dr. José Cruz Nájera Pérez y cede los derechos del trabajo intitulado Diseño de filtros digitales para el reconocimiento y la multi-clasificación de objetos, al Instituto Politécnico Nacional para su difusión, con fines académicos y de investigación <sup>en escenas reales</sup>.

Los usuarios de la información no deben reproducir el contenido textual, gráficas o datos del trabajo sin el permiso expreso del autor y/o director del trabajo. Este puede ser obtenido escribiendo a la siguiente dirección Av. del Parque No. 1310, Meza de Otay, Tijuana, B.C. Mexico. Si el permiso se otorga, el usuario deberá dar el agradecimiento correspondiente y citar la fuente del mismo.

  
Oliver Gamaliel Campos Trujillo

Nombre y firma

# Dedicatoria

*En memoria de mi padre,  
que me enseñó siempre con el ejemplo el camino de la honestidad y la bondad.  
Gracias papá, por aguantarme tantos años,  
espero que encuentres la paz,  
que no encuentraste en este mundo.*

# Agradecimientos

Primero que nada quisiera agradecer al **INSTITUTO POLITECNICO NACIONAL** y al **CITEDI** por darme la oportunidad de haber estudiado la maestría en sus aulas. Espero sinceramente no haberlos defraudado.

**A mi familia**, mi madre Elba Alicia y a mis hermanos Daniel y Abdel, por soportarme.

Al **CONACYT** por apoyarme financieramente.

Al **Dr. Víctor Hugo Díaz Ramírez**, por todo el apoyo, la confianza, la tolerancia, los consejos brindados. Gracias Profe.

Al **Profe. Andrés Calvillo** por la confianza, y la amistad brindada durante mi estancia en el centro, y por las sugerencias hechas en los primeros intentos de redacción.

A los demás miembros del mi comité **Dr. José Cruz Núñez Pérez** , **Dr. Roberto Sepúlveda** y **M.C. Ernesto E. Quiroz Morones** por sus valiosas sugerencias en la revisión de esta tesis.

A mis compañeros del área y amigos, **José Luis y Francisco Ramírez**, con los que compartí muchas desveladas, y cafés. Por enseñarme y compartir conmigo, sus perspectivas sobre el doctorado de la vida.

**A mis demás amigos:** Chely, Cristian, jorel, Diana, Sergio, Colors y a mis compañeros de control por esos buenos momentos de alegría, distracción y esos cafecillos ocasionales.

*Y todas las personas que directa o indirectamente contribuyeron en la realización de la presente tesis.*

# Resumen

Hoy en día, la necesidad de técnicas de reconocimiento de objetos, se hace cada vez más evidente en diferentes aplicaciones importantes como seguridad, inspección industrial automática, y diagnóstico médico, entre otras. Los sistemas de reconocimiento de objetos basados en características son quizás los más populares, sin embargo, estos sistemas son computacionalmente intensivos y más aún, su desempeño general depende de la toma de decisiones subjetivas por parte del diseñador. El reconocimiento de objetos por correlación, es una alternativa importante a los sistemas basados en características, ya que poseen sólidas bases matemáticas y pueden implementarse de forma eficiente en hardware opto-digital y en sistemas digitales con múltiples procesadores a gran velocidad. El problema general del reconocimiento de objetos consiste en la identificación, localización, y clasificación de diferentes objetos de interés a partir de una escena observada. Este problema puede resolverse al utilizar un banco de filtros de correlación donde se requiere de al menos un filtro por cada clase diferente de objeto. Sin embargo, bajo este enfoque la complejidad computacional del sistema puede incrementarse significativamente conforme aumenta el número de diferentes clases de objetos. En esta tesis, se propone el diseño de nuevos filtros adaptativos de correlación con restricciones complejas para el reconocimiento y la multclasificación de objetos en escenas reales. Nuestra hipótesis consiste en que, es posible reconocer y clasificar múltiples objetos con una sola operación de correlación, haciendo uso de los planos de intensidad y de fase compleja generados por el filtro de correlación, en una misma etapa. De esta forma, se podrá reducir considerablemente la complejidad computacional de un sistema de reconocimiento y clasificación de objetos basado en correlación. La construcción de los filtros de correlación adaptativos se realiza a través de un algoritmo de adaptación que optimiza las métricas de desempeño del filtro. El impacto de la propuesta es analizado respecto a la tolerancia al ruido, la complejidad computacional y la confiabilidad en el reconocimiento y clasificación, respecto a los filtros de correlación convencionales.

**Palabras clave:** Filtros adaptativos de correlación, Filtros invariantes a distorsiones geométricas, Multclasificación de objetos, Reconocimiento de objetos con correladores ópticos.

# Abstract

Nowadays, the need for object recognition techniques has become increasingly evident in various important applications such as security, automatic industrial inspection, and medical diagnostics, among others. Feature-based object recognition systems are perhaps the most popular, however, these systems may be computationally intensive and, moreover, its overall performance depends on the some ad-hoc decisions made by the designer. Object recognition by correlation filtering represents important alternative to feature-based systems because they have strong mathematical basis and can be efficiently implemented in hybrid opto-digital correlators or in digital processing systems at high speed. The main problem of object recognition is the identification, localization, and classification of several objects of interest from an input observed scene. This problem can be solved by using a bank of correlation filters which requires at least one filter per each different class of objects. However, under this approach the computational complexity of the resultant pattern recognition system can be increased considerably as the number of classes of objects increases. In this thesis, we propose the design of new adaptive composite filters with complex constraints for multi-class pattern recognition in real scenes. The hypothesis consists in proving that it is possible to recognize and classify multiple objects with a single correlation operation, using the intensity and phase distributions at the output complex correlation plane at the filter output, the same stage. In this manner, the computational complexity of the pattern recognition system can be reduced considerably. The proposed adaptive filters are constructed with the help of an iterative algorithm optimizing performance metrics. The impact of the proposal is analyzed in terms of tolerance to additive noise, computational complexity, and recognition and classification efficiency, compared to conventional correlation filters.

**Keywords:** Adaptive correlation filters, distortion-invariant correlation filters, multi-classification of objects, optical pattern recognition.

# ÍNDICE

<b>ÍNDICE</b>	<b>i</b>
<b>LISTA DE FIGURAS</b>	<b>iv</b>
<b>LISTA DE TABLAS</b>	<b>vii</b>
<b>1. INTRODUCCIÓN</b>	<b>1</b>
1.1. PROBLEMÁTICA . . . . .	2
1.2. JUSTIFICACIÓN . . . . .	2
1.3. OBJETIVOS . . . . .	3
1.3.1. Objetivo general . . . . .	3
1.3.2. Objetivos específicos. . . . .	3
1.3.3. Organización de la tesis. . . . .	4
<b>2. MARCO TEÓRICO</b>	<b>5</b>
2.1. FORMACIÓN DE IMAGENES . . . . .	5
2.1.1. Cuantización y muestreo de una imagen . . . . .	6
2.2. SISTEMAS LINEALES . . . . .	7
2.3. CORRELACIÓN Y CONVOLUCIÓN EN SISTEMAS LINEALES . . . . .	8
2.3.1. Propiedades de la convolución y correlación . . . . .	10
2.4. PROPIEDADES DE LAS SEÑALES DE DOS DIMENSIONES . . . . .	10
2.4.1. Función de dispersión . . . . .	11
2.5. TRANSFORMADA DE FOURIER . . . . .	13
2.6. TRANSFORMADA DE FOURIER PARA SEÑALES 2D . . . . .	15
2.7. FILTRADO EN EL DOMINIO DE LA FRECUENCIA . . . . .	15
2.8. BANCOS DE FILTROS . . . . .	17
2.8.1. Aplicaciones de bancos de filtros . . . . .	18

2.8.2.	Bancos de filtros de correlación . . . . .	19
2.9.	ESTRATEGIAS DE OPTIMIZACIÓN . . . . .	19
2.9.1.	Búsqueda local . . . . .	21
2.9.1.1.	Escalando la colina . . . . .	22
2.9.2.	Búsqueda global . . . . .	24
2.9.2.1.	Temple simulado . . . . .	24
2.9.2.2.	Algoritmos genéticos . . . . .	25
<b>3.</b>	<b>FILTROS DE CORRELACIÓN</b>	<b>27</b>
3.1.	ARQUITECTURAS DEL PROCESADORES ÓPTICOS . . . . .	27
3.1.1.	Procesador óptico 4F . . . . .	29
3.1.2.	Procesador óptico JTC . . . . .	30
3.2.	RECONOCIMIENTO DE PATRONES CON FILTROS DE CORRE- LACIÓN . . . . .	32
3.3.	MÉTRICAS DE DESEMPEÑO . . . . .	34
3.3.1.	Relación señal a ruido . . . . .	34
3.3.2.	Error Cuadrático Promedio . . . . .	34
3.3.3.	Medida de similaridad promedio . . . . .	36
3.3.4.	Capacidad de discriminabilidad . . . . .	37
3.3.5.	Relación pico - lóbulo . . . . .	37
3.4.	TAXONOMÍA DE LOS FILTROS DE CORRELACIÓN . . . . .	38
3.4.1.	Filtro de comparación clásico . . . . .	39
3.4.2.	Filtro de funciones de discriminación sintética . . . . .	40
3.4.3.	Filtro de promedio mínimo de energía de correlación . . . . .	42
3.4.4.	Filtro de altura máxima promedio de correlación . . . . .	43
3.4.5.	Filtro de correlación clasificador por distancia . . . . .	43
3.5.	PROBLEMAS EN FILTROS DE CORRELACIÓN . . . . .	45
3.5.1.	Modelo disjunto . . . . .	45
3.5.2.	Ruido aditivo en filtros de correlación . . . . .	46
3.5.3.	Problema de la clasificación en filtros de correlación . . . . .	47
<b>4.</b>	<b>FILTRO PROPUESTO DE MULTICLASIFICACIÓN</b>	<b>49</b>
4.1.	FILTROS DE CORRELACIÓN ADAPTATIVOS PARA LA MULTI- CLASIFICACIÓN EN EL RECONOCIMIENTO DE PATRONES . . . . .	49
4.1.1.	Búsqueda en el proceso de adaptación . . . . .	55

<b>5. RESULTADOS</b>	<b>56</b>
5.1. DESCRIPCIÓN DE PRUEBAS DE DESEMPEÑO DEL FILTRO PROPUESTO . . . . .	56
5.1.1. Algoritmo de las pruebas para el desempeño del ruido en escenas reales . . . . .	56
5.1.2. Algoritmo para medir el comportamiento de la fase ante imágenes geoméricamente distorsionadas por rotación . . . . .	57
5.1.3. Algoritmo para la construcción de matriz de confusión . . . . .	57
5.1.4. Prueba de error de localización . . . . .	59
5.2. RESULTADOS DE LA SÍNTESIS DE BANCOS DE FILTROS . . . . .	60
5.3. RESULTADOS DEL DESEMPEÑO DEL FILTRO ANTE LA PRESENCIA DE RUIDO ADITIVO Y RUIDO DISJUNTO . . . . .	61
5.4. RESULTADOS DE LA CAPACIDAD DE DISCRIMINALIDAD DEL FILTRO PROPUESTO . . . . .	63
<b>6. CONCLUSIONES</b>	<b>70</b>
6.1. APORTACIONES . . . . .	71
6.2. TRABAJOS FUTUROS . . . . .	71
<b>REFERENCIAS BIBLIOGRAFICAS</b>	<b>73</b>

# LISTA DE FIGURAS

2.1.	Sistema básico de captura de imágenes: a la izquierda un sistema biológico, a la derecha un sistema de visión digital. . . . .	6
2.2.	Efectos de la cuantización y de la resolución en una imagen. . . . .	6
2.3.	Diagrama a bloques de un sistema general de procesamiento de señales. . . . .	7
2.4.	Invariancia en el tiempo. . . . .	8
2.5.	(a) Respuesta de un sistema a un impulso unitario. (b) Secuencia discreta compuesta por impulsos unitarios. (c) Secuencia descompuesta en impulsos unitarios. (d) Respuesta de cada pulso unitario de la secuencia. (e) Suma de las respuesta a los impulsos de la secuencia $i[n]$ . . . . .	9
2.6.	Mapeo del dominio espacial al dominio de la frecuencia. . . . .	12
2.7.	Representación de una señal por series de Fourier. . . . .	13
2.8.	Convolución en el espacio y óptica. . . . .	16
2.9.	Configuración óptica para el procesado en Fourier. . . . .	17
2.10.	Representación gráfica de la respuesta de un banco de filtros en el dominio de la frecuencia para señales unidimensionales. . . . .	17
2.11.	Filtro de correlación en una dimensión. . . . .	18
2.12.	Esquema básico de un banco de filtros de correlación. . . . .	19
2.13.	Distribución de técnicas de optimización computacional. . . . .	20
2.14.	Ejemplo de una función objetivo con solución global y soluciones locales. . . . .	21
2.15.	Problemas de <i>hill climbing</i> . . . . .	22
2.16.	Trayectorias de búsqueda de los algoritmos hill-climbing y temple simulado. . . . .	25
3.1.	Difracción de Fraunhofer producida por una transparencia. . . . .	28
3.2.	Diagrama a bloques del procesador 4f. . . . .	29
3.3.	Implementación de un correlador óptico 4f. . . . .	29
3.4.	Imagen conjunta para un correlador JTC. . . . .	30

3.5.	Diagrama a bloques del correlador de transformada conjunta. . . . .	31
3.6.	Implementación del correlador JTC. . . . .	32
3.7.	Proceso general del reconocimiento de patrones basado en características. . . . .	33
3.8.	Métrica de promedio de error cuadrático. . . . .	35
3.9.	Representación gráfica de la métrica ASM. . . . .	36
3.10.	Región para calcular la PSR. . . . .	38
3.11.	Esquema de operación de un filtro DCCF. . . . .	44
3.12.	Distancias entre clases por medio de la transformación $\mathbf{H}$ del filtro DCCF. . . . .	45
3.13.	Representación del modelo disjunto. . . . .	45
3.14.	Efecto en la salida de correlación de una escena bajo el modelo disjunto. . . . .	46
3.15.	Representación gráfica del ruido aditivo. . . . .	47
3.16.	El problema de clasificación clásico usando el filtro de correlación MSF. . . . .	48
4.1.	Diagrama a bloques del algoritmo propuesto para el reconocimiento y clasificación de objetos. . . . .	51
4.2.	Algoritmo iterativo para el diseño de filtros adaptativos compuestos con restricciones complejas. . . . .	52
4.3.	Algoritmo propuesto para la construcción de un banco de filtros con restricciones complejas. . . . .	53
4.4.	Imágenes de referencia de las clases de los objetivos: (a) Clase 1; (b) Clase 2; (c) Clase 3; (d) Clase 4. . . . .	54
4.5.	Puntos seleccionados de la escena para la construcción de filtro. . . . .	55
5.1.	Simulación del desempeño del filtro propuesto bajo distintos niveles de ruido aditivo. . . . .	57
5.2.	Algoritmo del desempeño del filtro ante distorsiones de rotación. . . . .	57
5.3.	Ejemplo de la escena de entrada de prueba con ruido aditivo de varianza $\sigma_n^2 = 4/256$ . TC1: Objeto de la clase 1; TC2: Objeto de la clase 2; TC3: Objeto de la clase 3; TC4: Objeto de la clase 4. . . . .	59
5.4.	Desempeño del banco de filtros propuesto en el proceso de diseño: (a) ACSDF1, (b) ACSDF2, (c) ACSDF4 . . . . .	60
5.5.	Plano de intensidad de la correlación obtenida con el filtro ACSDF1 para la escena de entrada de la Fig. 5.3. . . . .	62
5.6.	Desempeño del reconocimiento y clasificación del filtro ACSDF1 con versiones rotadas de objetos de todas las clases. . . . .	63

5.8.	Valores de fase $\phi_k$ estimados del pico de correlación con un intervalo de confianza del 95 % mientras la varianza del ruido $\sigma_n$ es cambiada. (a) ACSDF1, (b) ACSDF2, (c) ACSDF4. . . . .	64
5.7.	Desempeño de la DC con 95 % de confianza para los diferentes bancos de filtros mientras la varianza del ruido $\sigma_n$ es cambiada. (a) ACSDF1, (b) ACSDF2, (c) ACSDF4, (d) MACE, (e) MACH, (f) DCC. . . . .	65
5.9.	Desempeño del ERC con intervalo de confianza del 95 % para los diferentes bancos de filtros mientras la varianza del ruido $\sigma_n$ se incrementa.	66
5.10.	Desempeño de LE con intervalos de confianza del 95 % para los diferentes bancos de filtros mientras la varianza del ruido $\sigma_n$ crece. (a) ACSDF1, (b) ACSDF2, (c) ACSDF4, (d) MACE, (e) MACH, (f) DCC.	67

# LISTA DE TABLAS

2.1. Propiedades de la convolución. . . . .	10
2.2. Funciones separables típicas aplicadas en imágenes. . . . .	11
2.3. Transformada de Fourier para señales de una dimensión. . . . .	13
2.5. Propiedades principales de la TF. . . . .	14
3.1. Taxonomía de los filtros de correlación. . . . .	38
5.1. Ejemplo de matriz de confusión de tres clases. . . . .	58
5.2. Desempeño del ACSDF1 en términos de la DC y de los valores de fase $\phi^k$ estimados en el pico de correlación para la escena de entrada mostrada en la Fig.5.3. . . . .	64
5.3. Desempeño de clasificación del ACSDF1. . . . .	68
5.4. Desempeño de clasificación del ACSDF2. . . . .	68
5.5. Desempeño de clasificación del ACSDF4. . . . .	68
5.6. Desempeño de clasificación del MACE. . . . .	68
5.7. Desempeño de clasificación del MACH. . . . .	69
5.8. Desempeño de clasificación del DCCF. . . . .	69

# Índice de algoritmos

2.1.	Pseudo algoritmo <i>hill climbing</i> . . . . .	23
2.2.	Pseudocódigo de <i>hill climbing</i> estocástico . . . . .	23
2.3.	Seudocódigo para el algoritmo genético . . . . .	26
5.1.	Pseudocódigo para la construcción de la matriz de confusión. . . . .	58

# Nomenclatura

⊗	Correlación
★	Convolución
2D	Dos dimensiones
ACE	Promedio de la energía de correlación
ACSDF	Filtro SDF adaptativo con restricciones complejas
ASE	Error cuadrático promedio
ASM	Medida de promedio de similaridad
AWGN	Ruido blanco aditivo gaussiano
CMF	Classic Matched Filter
DC	Capacidad de discriminabilidad
DCCF	Filtro de correlación de clasificación de distancia
ERC	Tasa de error de clasificación
FFT	Transformada rápida de Fourier
JTC	Correlador de transformada conjunta
LE	Error de localización
LTI	Sistema lineal e invariante en el tiempo
MACE	Filtro de Promedio mínimo de energía de correlación
MACH	Filtro de altura máxima promedio de correlación
MEL	Modulador espacial de luz
MSE	Error cuadrático medio

PSR	Relación pico lóbulo
RP	Reconocimiento de patrones
SDF	Filtro de funciones de discriminación sintética
SNR	Relación señal a ruido
TF	Transformada de Fourier
TIF	Transformada inversa de Fourier

# 1

## INTRODUCCIÓN

EL reconocimiento de patrones (RP) ha sido uno de los tópicos de procesamiento de imágenes más investigados en las últimas décadas. Esto se debe a que el RP ha tenido gran impacto en actividades importantes del mundo real, tales como vigilancia, inspección industrial automática, diagnóstico clínico y tecnología militar. Básicamente, el RP consiste en la detección, identificación y clasificación de un conjunto de objetos dados en una escena observada. Típicamente, este problema es resuelto por medio de sistemas basados en el uso de extracción de características [1]. En esta aproximación, el RP es desarrollado en dos etapas. Primero, la segmentación de los objetos es llevada a cabo, seguida por la identificación y clasificación de los objetos segmentados usando sus características extraídas. Los sistemas basados en extracción de características pueden ser computacionalmente intensivos. Además, su desempeño general depende de la toma de decisiones subjetivas por parte del diseñador que podrían ser optimizadas a través de un procedimiento formal. Tenga en cuenta que en esta situación, los filtros de correlación representan una importante alternativa respecto a los métodos basados en extracción de características, (ya que no es necesaria la segmentación) debido a que los filtros están basados en métodos estadísticos y de optimización. Adicionalmente, es posible implementar los filtros de correlación de forma eficiente en sistemas de procesamiento digital o en correladores híbridos opto-digitales a una alta velocidad de operación [2, 3].

## 1.1. PROBLEMÁTICA

Los filtros de correlación son usualmente diseñados a través de la optimización de algún criterio estadístico [4]. Entonces estos filtros pueden ser robustos a las distorsiones típicas que afectan al objeto y a la escena de entrada [5, 6, 7]. Básicamente, un filtro de correlación es un sistema lineal e invariante en el espacio. En donde el valor máximo a la salida del sistema es la estimación de máxima verosimilitud de las coordenadas del objetivo en la escena observada [8]. Nótese que de esta forma, solo se considera la estimación de la posición de los objetos en la escena, por lo que este simple procedimiento no permite la clasificación. Por lo tanto, para superar este inconveniente, un banco de filtros de correlación es comúnmente usado. Para clasificar los objetos con un banco de filtros de correlación, se requiere de por lo menos un filtro de correlación por cada clase de objeto considerado [9, 10, 11]. Entonces, la clasificación es llevada a cabo a través de un post-procesamiento que da como resultado la obtención de un plano de correlación por cada filtro del banco. Un filtro compuesto [12] es sintetizado por la combinación de varias imágenes (representativas) de entrenamiento. Es de esperar que el filtro sea adecuado para el reconocimiento de versiones distorsionadas del objeto. Los filtros basados en funciones de discriminación sintética (SDF) [13], de energía promedio de correlación mínima (Minimum Average Correlation Energy, MACE) [14], de máxima altura de correlación promedio (MACH) [15], y clasificador de distancia de correlación (DCC) [16], son filtros compuestos exitosos que pueden ser usados para la clasificación de los objetos. Estos filtros producen buenos resultados en términos del desempeño del reconocimiento y clasificación. Sin embargo, la complejidad computacional del sistema de RP se incrementa considerablemente cuando el número de clases de objetos aumenta [11, 10]. Por esta razón, el diseño de un nuevo sistema de RP con un buen desempeño en el reconocimiento, clasificación y de baja complejidad computacional es deseable.

## 1.2. JUSTIFICACIÓN

Es importante notar que la salida del filtro de correlación es, en general, una función compleja. Sin embargo, la aproximación de RP convencional usa solo información real para estimar las coordenadas del objeto y un post-procesamiento adicional para llevar a cabo la clasificación. En esta tesis, proponemos un nuevo enfoque adaptativo utilizando filtros compuestos con restricciones complejas para el reconocimiento y clasificación confiable de múltiples objetivos en escenas ruidosas. Mediante el uso de las distribuciones de intensidad y fase a la salida del plano complejo de correlación, diversos objetos pueden ser reconocidos y clasificados en un procedimiento de dos pasos.

Primero, una estimación de la localización de los objetos es llevada a cabo mediante el uso de un estimador de máxima verosimilitud convencional en la salida compleja del plano de correlación. A continuación, los valores de fase en la ubicación estimada de los objetivos indican la clase para cada objeto que ha sido reconocido. Como resultado, esta propuesta es capaz de reconocer y clasificar las versiones geoméricamente distorsionadas de los objetos que pertenecen a diferentes clases con un solo filtro de correlación. Tenga en cuenta que con el enfoque propuesto mediante correlaciones complejas, la complejidad computacional del RP puede ser reducida considerablemente en comparación con el método basado solo en el uso de correlaciones reales. Además, el sistema propuesto es capaz de solventar el problema del ruido causado por la aparición de lóbulos de correlación, que son generados por estructuras similares a los patrones objetivo presentes en el área del fondo. Este problema puede resolverse al utilizar un enfoque adaptativo para el diseño del filtro [17, 18], donde el mecanismo de adaptación se basa en maximizar una o varias métricas por medio de la información presentada por el fondo, con el fin de eliminar la mayor parte de esta información, de manera que los picos sean mejor discriminados por el filtro, ante escenas reales con fondos complicados.

## **1.3. OBJETIVOS**

### **1.3.1. Objetivo general**

Diseñar filtros digitales para reconocer y clasificar objetos en escenas reales con invariancia a distorsiones geométricas de los objetos, ruido de fondo y ruido del sensor.

### **1.3.2. Objetivos específicos.**

- I Diseñar nuevos filtros adaptativos de correlación para la multclasificación de objetos que tengan como base, el uso de toda la respuesta compleja de correlación a la salida del filtro.
- II Diseñar filtros para la multclasificación de objetos optimizados respecto a métricas de desempeño.
- III Proponer e implementar pruebas que midan el desempeño del uso de la fase como herramienta para clasificar múltiples objetos, ante las situaciones más comunes con son el ruido aditivo y el ruido disjunto, así como también tolerancia a pequeñas distorsiones geométricas de los objetos..
- IV Explorar el enfoque de la adaptación local con la aproximación propuesta.

V Implementar pruebas que muestren el desempeño, respecto a los filtros tradicionales para la clasificación de objetos y tolerancia al ruido.

### **1.3.3. Organización de la tesis.**

La presente tesis está organizada de la siguiente manera: En el Capítulo 2 presentamos el marco teórico necesario para comprender el enfoque de reconocimiento de objetos por filtros de correlación. En el Capítulo 3, presentamos el estado del arte de los filtros de correlación más exitosos, así como también de una manera más extensa las problemáticas existentes de los filtros actuales. En el Capítulo 4, se desarrolla una versión extensa del diseño del filtro propuesto para la multclasificación de objetos y el procedimiento de adaptación del mismo. En el Capítulo 5, se presenta la descripción de las pruebas de desempeño ejecutadas, así como también los resultados arrojados por tales pruebas. Finalmente, en el Capítulo 6 se presentan las conclusiones generales del análisis e implementación del filtro propuesto en la presente tesis.

# 2

## MARCO TEÓRICO

### 2.1. FORMACIÓN DE IMAGENES

Una imagen se puede definir de varias maneras, desde una perspectiva biológica una imagen es el reflejo de la radiación electromagnética captada por la retina del ojo humano (véase Fig. 2.1) y posteriormente convertida en impulsos nerviosos [19]. Desde un punto de vista técnico una imagen puede considerarse como una representación de señales independientes que componen una función bidimensional [20]. Esta función puede representar distintos fenómenos físicos como el escaneo de rayos X, la captura de la sección transversal de un radar, y la energía captada por una cámara infrarroja, entre otras. Es fácil darse cuenta, que en los ejemplos mencionados, las imágenes resultantes son funciones discretas que se obtienen a partir de imágenes continuas del fenómeno capturado; la función discreta saliente  $f_d(x, y)$  comúnmente se expresa de la siguiente manera:

$$f_d(x, y) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} f(j\Delta x, k\Delta y) \delta(x - j\Delta x, y - k\Delta y), \quad (2.1.1)$$

donde  $j\Delta x$ ,  $k\Delta y$  representan incrementos discretos de las variables continuas " $x$ ", " $y$ ", y  $\delta(x, y)$  es la función delta de Dirac. Un sistema de captura de imágenes digitales (vease Fig. 2.1) está compuesto por una escena física, una sistema de captura donde cada elemento de la imagen resultante representa una muestra que es cuantizada a

un valor discreto, donde está muestra es comúnmente conocida como pixel (picture element). La imagen resultante, puede ser almacenada, procesada, y/o desplegada, según convenga para una aplicación en particular.

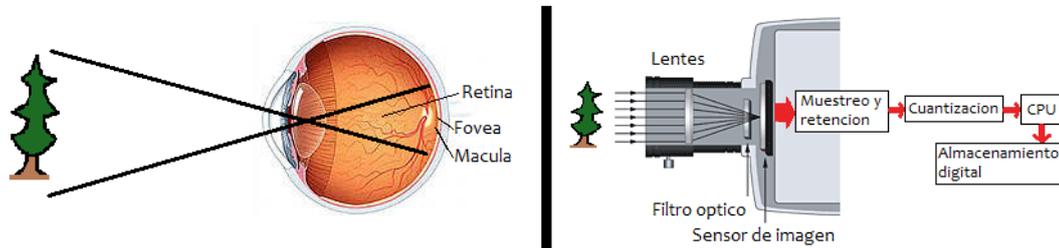


Figura 2.1: Sistema básico de captura de imágenes: a la izquierda un sistema biológico, a la derecha un sistema de visión digital.

### 2.1.1. Cuantización y muestreo de una imagen

La resolución de una imagen está dada por el número de muestras que el sistema de captura es capaz de registrar. Además, la capacidad de captar detalles finos de las imágenes depende también del número de niveles de cuantización del sistema de captura. Esto quiere decir, que la calidad de la imagen depende de la resolución y de la profundidad de cuantización, y sus efectos son mostrados en la Fig. 2.2. Nótese que la forma del objeto en la figura no cambia, sin embargo la textura se va degradando gradualmente conforme se reducen los niveles de cuantización. Adicionalmente, observemos como el contorno del objeto se va deformando gradualmente conforme la resolución del sistema disminuye.

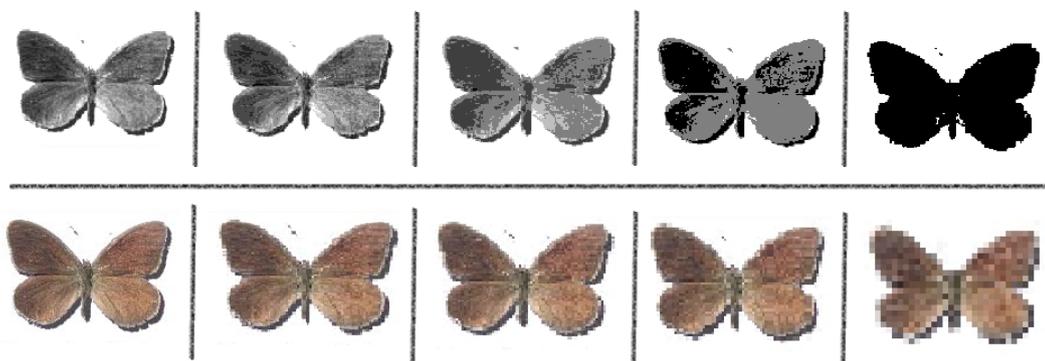


Figura 2.2: Efectos de la cuantización y de la resolución en una imagen.

## 2.2. SISTEMAS LINEALES

Uno de los conceptos básicos utilizados en esta tesis es el uso de sistemas lineales. En términos generales un sistema se puede definir como un conjunto de objetos que interactúan entre sí, para conformar un objeto mucho más complejo que está destinado a un fin específico [21]. Cualquier sistema está compuesto por una estructura básica, tal y como se puede ver en la Fig. 2.3, donde el sistema puede estar constituido por una o varias entradas, donde dichas entradas usualmente son funciones de una o más variables independientes que contienen información acerca de la naturaleza o comportamiento de algún fenómeno físico. Los sistemas reciben estas señales como entrada y responden a ellas produciendo señales a la salida.



Figura 2.3: Diagrama a bloques de un sistema general de procesamiento de señales.

En el ambiente del procesamiento de señales, existen sistemas con diferentes características (discretos, continuos, determinísticos, aleatorios) los cuales son ampliamente mencionadas en literatura [22]. Los sistemas lineales, han sido ampliamente estudiados debido a que poseen diferentes propiedades atractivas para aplicaciones reales. Un sistema lineal debe de cumplir con los principios de superposición y homogeneidad, que son descritos a continuación:

$$\textit{Propiedad de superposicion} \quad \rightarrow \quad f(x + y) = f(x) + f(y), \quad (2.2.1)$$

$$\textit{Propiedad de homogeneidad} \quad \rightarrow \quad f(\lambda x) = \lambda f(x). \quad (2.2.2)$$

Las propiedades de superposición y homogeneidad son importantes en aplicaciones reales, ya que simplifican la manipulación de dichos sistemas, permiten caracterizar la salida del sistema a partir de la entrada y de una función característica del propio sistema conocida como respuesta al impulso. Al saber que el sistema es lineal, este puede ser caracterizado separando las entradas y resolviendo cada una de ellas de forma independiente, entonces por medio de las propiedades de los sistemas lineales (superposición y homogeneidad) es posible conocer el resultado global del sistema. Para comprobar la linealidad de un sistema de una función arbitraria  $f(x)$ , basta con mostrar que  $f(ax + by) = af(x) + bf(y)$ , entonces la propiedad para dicha función

arbitraria  $f(x)$  queda demostrada.



Figura 2.4: Invariancia en el tiempo.

La propiedad de la invariancia en el tiempo (vease Fig. 2.4), se describe como la forma en que el sistema responde a entradas desplazadas en el tiempo. El significado físico de esta propiedad establece que la salida del sistema registrará un desplazamiento en el tiempo igual al desplazamiento en el tiempo de la señal de entrada. Esta propiedad es importante, porque en conjunto con la propiedad de linealidad nos da la facilidad de poder analizar el sistema en términos de la respuesta frecuencial, lo cual es visto en secciones posteriores.

### 2.3. CORRELACIÓN Y CONVOLUCIÓN EN SISTEMAS LINEALES

Los sistemas lineales e invariantes en el tiempo (LTI) son importantes porque pueden caracterizarse a través de su respuesta al impulso [23]. Nótese que la respuesta al impulso de un sistema LTI, está dada por la función de salida del sistema cuando la entrada es la función delta de Dirac. Además, es fácil darse cuenta que una señal discreta puede representarse como una combinación lineal de funciones impulso con diferentes pesos. De esta manera, haciendo uso de la propiedad de linealidad, la salida del sistema está dada por la suma de las respuestas individuales de cada una de las funciones impulso ponderadas que componen a la señal de entrada. Esto último puede verse en la Fig. 2.5 donde en la Fig. 2.5(a) se muestra la respuesta al impulso del sistema  $h[n]$ . Además en la Fig. 2.5(b),(c), se muestra una señal discreta de entrada y su descomposición en impulsos unitarios, mientras que la Fig. 2.5(d),(e) muestra las salidas individuales del sistema en respuesta a las diferentes funciones impulso de entrada, y muestra la superposición de las salidas individuales para formar la salida completa del sistema. El proceso descrito gráficamente en la Fig. 2.5 se le conoce como sumatoria de convolución y puede ser realizado por sistemas LTI. Su representación

matemática para el caso discreto está dada por

$$o[n] = \sum_{k=-\infty}^{\infty} i[k] h[n-k] = i[n] \star h[n], \quad (2.3.1)$$

donde  $i[k]$  es una señal arbitraria,  $h[n-k]$  es la respuesta al impulso de nuestro sistema que se va desplazando  $k$  muestras por la señal  $i[k]$ , y  $o[n]$  es la salida del sistema. Una manera simplificada de ver la operación de convolución es como una suma de la respuestas de las muestras de la señal con el sistema. Un proceso estrechamente ligado con la convolución es la operación de correlación, definida por:

$$c[n] = \sum_{k=-\infty}^{\infty} i[k] h[n+k] = i[n] \otimes h[n]. \quad (2.3.2)$$

Nótese que la diferencia entre convolución y correlación radica en que la operación de la convolución invierte una de la señales para desplazarla, mientras que en la correlación la señal se desplaza de manera directa. Sin embargo, el resultado de ambas operaciones es totalmente diferente.

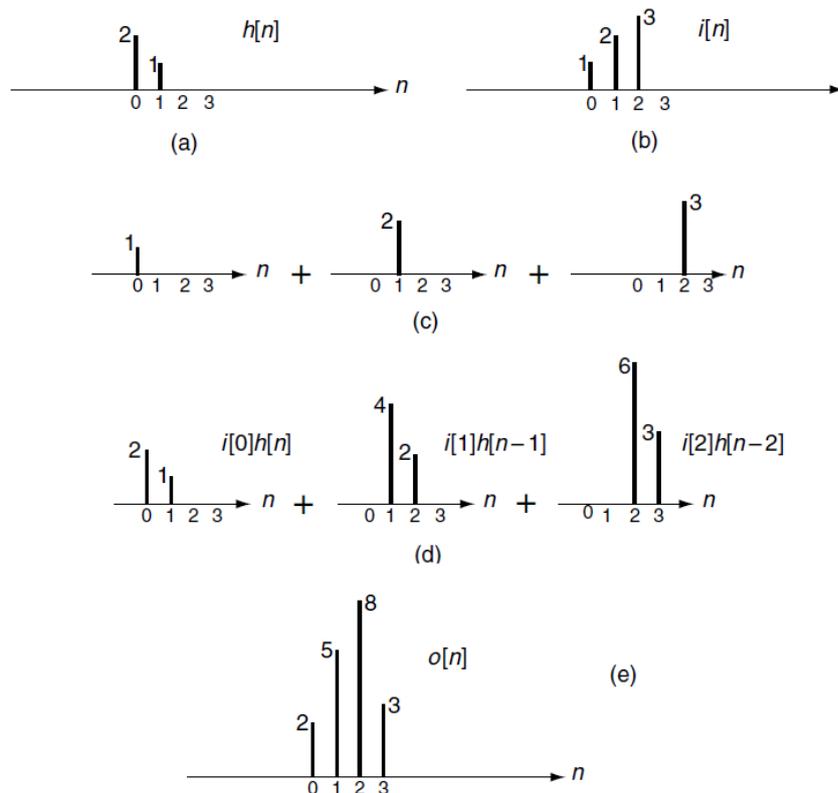


Figura 2.5: (a) Respuesta de un sistema a un impulso unitario. (b) Secuencia discreta compuesta por impulsos unitarios. (c) Secuencia descompuesta en impulsos unitarios. (d) Respuesta de cada pulso unitario de la secuencia. (e) Suma de las respuesta a los impulsos de la secuencia  $i[n]$ .

La convolución entrega la respuesta del sistema a una señal arbitraria y la correlación cuantifica la similitud entre las señales  $i[k]$  y  $h[k]$ . La operación de convolución se describe implícitamente con el símbolo  $\star$ , mientras que la correlación es descrita con el símbolo  $\otimes$ , tal y como se puede ver en la Ecs. (2.3.1) y (2.3.2).

### 2.3.1. Propiedades de la convolución y correlación

Tanto la convolución como la correlación tienen propiedades similares, como la linealidad e invariancia en el tiempo; ambas asociadas a los sistemas LTI. En la Tabla 2.1 se pueden ver las propiedades de la convolución.

Tabla 2.1: Propiedades de la convolución.

Propiedad	Descripción
Linealidad	$h(x) \star [ai_1(x) + bi_2(x)] = \{ah(x) \star i_1(x) + bh(x) \star i_2(x)\}$ , donde $a$ y $b$ son escalares
Invariancia al desplazamiento	$o(x) = i(x) \star h(x) \implies i(x - x_0) \star h(x) = o(x - x_0)$
Conmutatividad	$o(x) = i(x) \star h(x) = \int i(\tau) \cdot h(x - \tau) d\tau = \int i(x - u) \cdot h(u) du = h(x) \star i(x)$
Convolución con un desplazamiento en la función delta	$\delta(x - x_0) \star h(x) = h(x - x_0)$

## 2.4. PROPIEDADES DE LAS SEÑALES DE DOS DIMENSIONES

Usualmente las señales en dos dimensiones (2D) pueden representarse como el producto de dos señales unidimensionales [24]. Las señales 2D resultantes a través de esta práctica, son conocidas como señales separables. En algunos casos, las imágenes pueden modelarse también como señales separables, tal como se muestra enseguida,

$$i(x, y) = i_x(x)i_y(y); \quad i[n, m] = i_n[n]i_m[m]. \quad (2.4.1)$$

Las imágenes separables, tienen la ventaja de que pueden ser procesadas con mayor facilidad que las imágenes no separables. Mientras que las imágenes separables pueden procesarse manipulando dos vectores unidimensionales, las imágenes no separables tienen que procesarse como matrices. Sin embargo, las imágenes capturadas en

la vida real son raramente separables e incluso, estas imágenes, raramente permiten representaciones analíticas. Esto se debe a que las señales que componen las imágenes reales son mutuamente independientes entre si [25]. En la Ec. (2.4.1), la señal  $i(x, y)$  representa a una imagen continua separable y la señal  $i[n, m]$  a una imagen digital separable. Contrariamente a lo visto anteriormente, las imágenes representadas por funciones analíticas, usualmente resultan separables, como a continuación:

$$f(x, y) = f_1(x) f_2(y). \quad (2.4.2)$$

Algunos ejemplos típicos de funciones separables, se presentan en la Tabla 2.2, cabe mencionar que cada una de estas funciones cumplen con la propiedad de linealidad al igual que en el caso de una dimensión.

Tabla 2.2: Funciones separables típicas aplicadas en imágenes.

Función	Definición	Función	Definición
Delta de dirac	$\delta(x) = 0, x \neq 0$ $\lim_{x \rightarrow 0} \int_{-a}^a \delta(x) dx = 1$	Rectangulo	$\text{rect}(x) =$ $\begin{cases} 1, &  x  \leq 1/2 \\ 0, &  x  > 1/2 \end{cases}$
Propiedad de filtrado	$\int_{-\infty}^{\infty} f(x') \delta(x - x') dx' = f(x)$	Signo	$\text{sgn}(x) =$ $\begin{cases} 1, & x > 0 \\ 0, & x = 0 \\ -1, & x < 0 \end{cases}$
Propiedad de es- calamiento	$\delta(ax) = \frac{\delta(x)}{ a }$	Sinc	$\text{sinc}(x) = \frac{\sin \pi x}{\pi x}$
Delta de Kronecker	$\delta(n) = \begin{cases} 0, & n \neq 0 \\ 1, & n = 0 \end{cases}$	Peine	$\text{comb}(x) =$ $\sum_{n=-\infty}^{\infty} \delta(x - n)$
Propiedad de filtrado	$\sum_{m=-\infty}^{\infty} f(m) \delta(n - m) = f(n)$	Triangular	$\text{tri}(x) =$ $\begin{cases} 1 -  x , &  x  \leq 1 \\ 0, &  x  > 1 \end{cases}$

### 2.4.1. Función de dispersión

Es llamada función de dispersión a la respuesta impulso de un sistema, donde las entradas y las salidas representan una cantidad positiva tal como la intensidad de la luz en un sistema de imágenes [26]. El termino respuesta al impulso es mas general y este permite valores negativos, así como también valores complejos. La región de soporte de una respuesta al impulso es una pequeña región de tamaño  $n, m$ , cerrada en

el plano. Sea  $o[n, m]$  la salida de un sistema de imágenes, dada por

$$o[n, m] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} i[k, l] \cdot h[n - k, m - l] = i[n, m] \star \star h[n, m], \quad (2.4.3)$$

donde  $i[n, m]$  es la imagen de entrada al sistema, y  $h[n, m]$  es la función de dispersión. Nótese que en un sistema de imágenes, la operación de convolución entre la imagen de entrada y la función de dispersión se representa implícitamente como  $i[n, m] \star \star h[n, m]$ . Si la imagen de entrada y la función de dispersión son separables, entonces la salida del sistema también es separable, tal y como se muestra a continuación:

$$\begin{aligned} o[n, m] &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} i[k, l] \cdot h[n - k, m - l], \\ &= \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} \{i_n[k] \cdot i_m[l]\} \cdot \{h_n[n - k] \cdot h_m[m - l]\}, \quad (2.4.4) \\ &= \sum_{k=-\infty}^{\infty} \{i_n[k] \cdot h_n[n - k]\} \cdot \sum_{l=-\infty}^{\infty} \{i_m[l] \cdot h_m[m - l]\} = o_n[n] \cdot o_m[m]. \end{aligned}$$

La Ec. (2.4.4) establece que en los casos donde las señales son separables, la convolución en 2D es posible manejarla como dos convoluciones en una sola dimensión, llevadas a cabo de forma secuencial.

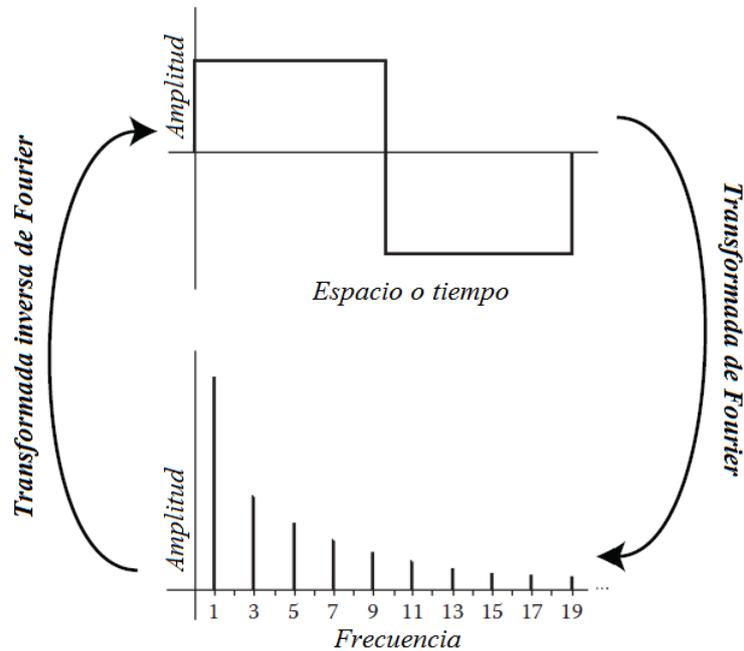


Figura 2.6: Mapeo del dominio espacial al dominio de la frecuencia.

## 2.5. TRANSFORMADA DE FOURIER

La transformada de Fourier (TF) es una operación lineal que nos permite mapear una función del dominio temporal o espacial al dominio de la frecuencia (véase Fig. 2.6). La TF es ampliamente usada en diversas áreas como en compresión de imágenes y video, selección de características, mejora en imágenes, entre otras. Esto es debido a que ciertos procesos son más fáciles de manejar y analizar en el dominio de la frecuencia que en el dominio espacial.

Tabla 2.3: Transformada de Fourier para señales de una dimensión.

Transformada continua de Fourier (TCF)	Inversa de TCF
$F(u) = \int_{-\infty}^{+\infty} f(x)e^{-2\pi iux} dx$	$f(x) = \int_{-\infty}^{+\infty} F(u)e^{2\pi iux} dx$
Transformada discreta de Fourier (TDF)	Inversa de TDF
$F(u) = \frac{1}{N} \sum_{x=0}^{N-1} f(x)e^{-\frac{i2\pi ux}{N}}$	$f(x) = \sum_{u=0}^{N-1} F(u)e^{\frac{i2\pi ux}{N}}$

El concepto de la TF se basa en que es posible representar cualquier señal periódica por medio de una combinación lineal de señales elementales [27]. Una definición más formal sobre esta idea es que una señal puede ser reconstruida por una sumatoria ponderada de funciones base (el caso más sencillo es con funciones ortogonales como las funciones seno y coseno).

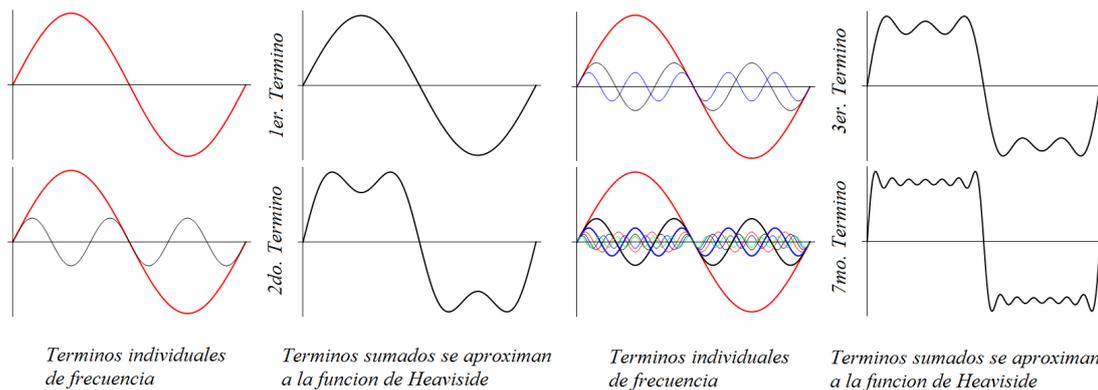


Figura 2.7: Representación de una señal por series de Fourier.

Bajo este enfoque, cada termino conocido nos acerca mas y mas a una representación exacta de la señal (vease Fig. 2.7), sin embargo, la representación por medio de series de Fourier no resulta conveniente cuando se aplica a señales aperiódicas o con periodo infinito dado a que requiere de un número infinito de coeficientes, lo que provoca que el espectro se convierta en una función continua. Una manera más simple de representar

este espectro es por medio de la transformada de Fourier, que es una integral como se ve en la Tabla 2.3, o por una sumatoria para el caso de señales discretas.

En la Tabla 2.3, se puede observar que la transformada de Fourier, además de contar con la ventaja de ser lineal, posee la operación inversa. Esto quiere decir, que la transformada inversa de Fourier (TIF) nos puede regresar del dominio frecuencial al dominio espacial. Este hecho es tremendamente útil, si lo que deseamos es realizar operaciones en el dominio de la frecuencia, y visualizar resultados en el dominio espacial.

Sea  $f(x)$  una función real, representando por ejemplo, un voltaje variable, el brillo en una imagen (variación espacial), la variación de señales sísmicas, entre otros casos. La TF de  $f(x)$ , estará dada por la función  $F(w)$  que describe la concentración de cada componente en frecuencia sobre la que esta distribuida la función  $f(x)$ . Generalmente,  $F(w)$  es una función compleja de la forma  $F(u) = R(u) + jI(u)$  donde  $I(u)$  y  $R(u)$  representan la parte imaginaria y la parte real, respectivamente. Usualmente resulta mas conveniente expresar el resultado en forma polar, como en la siguiente ecuación:

$$F(u) = \|F(u)\| e^{i\varphi(u)}. \quad (2.5.1)$$

Donde  $\|F(u)\|$  es la magnitud y  $\varphi(u)$  es la fase. La visualización de estos dos parámetros son importantes en diversas operaciones de filtrado. La expresion  $\|F(u)\|^2$  es comúnmente llamada potencia espectral o densidad espectral, la cual es útil para visualizar características relacionadas con la energía de las señales. Otra propiedad útil de la transformada de Fourier es la separabilidad; anteriormente se había mencionado que las señales reales son rara vez separables, sin embargo, con la representación polar es posible representar señales no separables de forma separable. Otras propiedades no menos importantes de la TF, se pueden ver en la Tabla 2.5.

Tabla 2.5: Propiedades principales de la TF.

Propiedad	Descripción
Dualidad	$i(x) \leftrightarrow I(f) \Rightarrow I(-x) \leftrightarrow i(f)$
Linealidad	$FT \{ai_1(x) + bi_2(x)\} = aI_1(f) + bI_2(f)$
Desplazamiento en el tiempo	$i(x - x_0) \longleftrightarrow I(f) \cdot e^{-j2\pi f x_0}$
Derivación del tiempo	$di(x)/dx \longleftrightarrow (j2\pi f I(f))$ y $d^n i(x)/dx^n \longleftrightarrow (j2\pi f)^n I(f)$
Simetría conjugada	$i(x) \text{ real} \Rightarrow I(f) = I^*(-f) \Rightarrow  I(f)  =  I(-f) $ y $\theta(f) = -\theta(-f)$
Operación inversa	$FT \{i(-x)\} = I(-f) \therefore$ para $i(x) \text{ real}$ $FT \{i(-x)\} = I(-f) = I^*(f)$
Modulación	$t(x) = i(x) \cdot h(x) \longleftrightarrow I(f) \star H(f) = T(f)$
Convolución	$i(x) \star h(x) \longleftrightarrow I(f) \cdot H(f)$
Correlación	$i(x) \otimes h(x) \longleftrightarrow I(f) \cdot H^*(f)$

## 2.6. TRANSFORMADA DE FOURIER PARA SEÑALES 2D

Al analizar la transformada de Fourier en señales unidimensionales, es claro ver que es posible extender su uso a dos o más dimensiones. En esta tesis solo se tomara en cuenta la transformada de Fourier para señales bidimensionales, particularmente, para imágenes digitales. La transformada de Fourier de una imagen  $f(x, y)$ , está dada por

$$F(u, v) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(\frac{ux}{M} + \frac{vy}{N})}, \quad (2.6.1)$$

donde M y N son constantes que indican el tamaño de la señal bidimensional, las variables  $(x, y)$  son coordenadas espaciales en la imagen, y las variables  $(u, v)$  representan las coordenadas en el dominio de la frecuencia. La transformada inversa de Fourier para señales 2D, está dada por:

$$f(x, y) = \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) \exp^{j2\pi(\frac{ux}{M} + \frac{vy}{N})}. \quad (2.6.2)$$

Parte del éxito de la transformada de Fourier en el procesamiento de señales, se debe a la variedad de algoritmos para acelerar su cálculo. Para esto, se utiliza una aproximación más rápida para el cálculo de la transformada basada en el método de divide y conquista [28] con el cual se reduce la complejidad computacional de  $O(N^2)$  operaciones a  $O(N \log_2 N)$  operaciones [29]. Este algoritmo es conocido como Transformada rápida de Fourier (FFT por sus siglas en inglés). El algoritmo de la FFT es muy útil y su aplicación se extiende por todas las ciencias y en la ingeniería. En el procesamiento de señales digitales es donde ha tenido su mayor impacto, de hecho se pueden encontrar diversas implementaciones en la literatura con esta herramienta.

## 2.7. FILTRADO EN EL DOMINIO DE LA FRECUENCIA

Una de las propiedades más útiles de la transformada de Fourier es su relación con las operaciones de convolución y correlación. Esto es importante porque para aplicar un filtro lineal a una imagen se tiene que convolucionar el filtro deseado con la imagen de la escena, donde podemos observar que cuando este proceso es realizado en el dominio espacial resulta computacionalmente demandante. En vez de esto la operación de convolución se realiza en el dominio de la frecuencia donde solo se requiere hacer la multiplicación punto a punto entre ambas señales. En este proceso intervienen la

propiedad de desplazamiento de la TF de manera implícita. Es sabido como ya se menciono en los capítulos anteriores, que la operación de convolución es lineal, lo que permite que puedan sumarse las respuestas del filtro al desplazarse por la señal de entrada, obteniendo así la salida del sistema. En el dominio de la frecuencia no hay necesidad de hacer esto, ya que solo se requiere un desplazamiento en la fase permitiendo así que solo se tenga que multiplicar la magnitud de los coeficientes de las señales. Esta equivalencia, está dada por

$$\begin{aligned}
 F[f(t) \star f(t)] &= F\left[\int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau\right] \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau)g(t - \tau)d\tau e^{-j\omega t}dt = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(\tau)g(t - \tau)e^{-j\omega t}dtd\tau \\
 \int_{-\infty}^{\infty} f(\tau) \int_{-\infty}^{\infty} g(t - \tau)e^{-j\omega t}dtd\tau &= \int_{-\infty}^{\infty} f(\tau) \int_{-\infty}^{\infty} g(t - \tau)e^{-j\omega t}dtd\tau \quad (2.7.1) \\
 &= \int_{-\infty}^{\infty} f(\tau)G(\omega)e^{-j\omega\tau}d\tau = G(\omega) \int_{-\infty}^{\infty} f(\tau)e^{-j\omega\tau}d\tau = G(\omega) \cdot F(\omega).
 \end{aligned}$$

De manera física el proceso descrito en la Ec. (2.7.1), lo podemos ver en un sistema óptico para señales de 2D. Cada punto de la imagen pasa por el origen del filtro de manera paralela sin necesidad de un desplazamiento por la imagen. En un sistema óptico, la trasformada de Fourier se lleva a cabo a través de una configuración adecuada de polarizadores y lentes, tal y como se ve en la Fig. 2.8 y en la implementación de la Fig. 2.9.

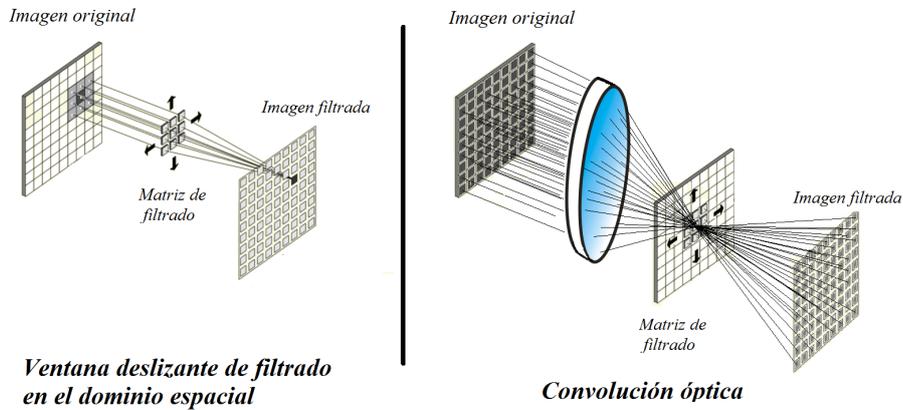


Figura 2.8: Convolución en el espacio y óptica.

La correlación en el dominio espacial se lleva a cabo de forma similar, sin embargo, la señal no se invierte para desplazarse. Visto de otro modo, la señal se desplaza de derecha a izquierda, lo que produce que en el dominio de la frecuencia, el espectro de la imagen de entrada tenga que ser multiplicada por el complejo conjugado de la respuesta en frecuencia del filtro o sistema. Nótese que este hecho solo afecta a la fase de la respuesta en frecuencia del sistema original.

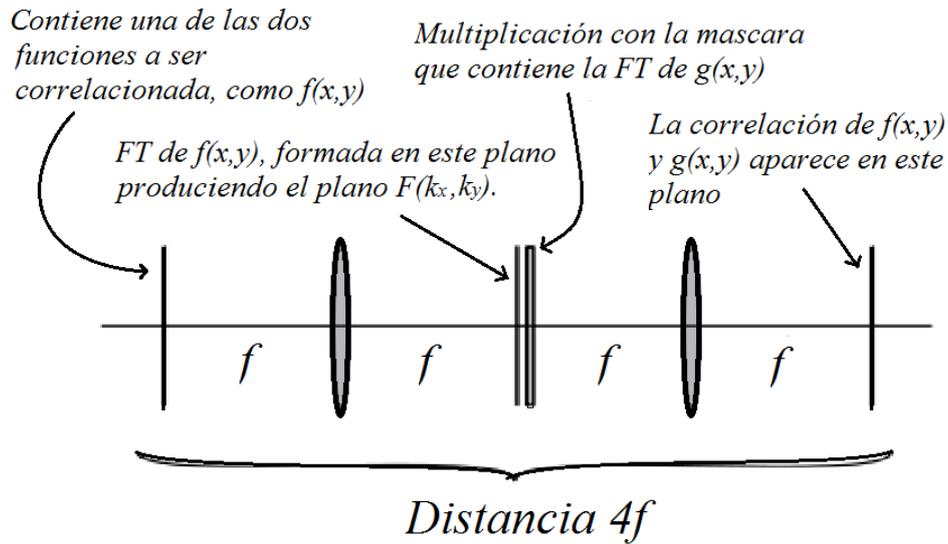


Figura 2.9: Configuración óptica para el procesamiento en Fourier.

## 2.8. BANCOS DE FILTROS

En la teoría de filtrado existen dos problemas básicos, el primero de ellos hace referencia a que los filtros ideales no existen en la práctica, tal y como puede verse en la teoría matemática [30]. Por lo tanto, diferentes técnicas de aproximación deben ser utilizadas dependiendo el tipo de ruido o características que deseamos preservar en la señal.

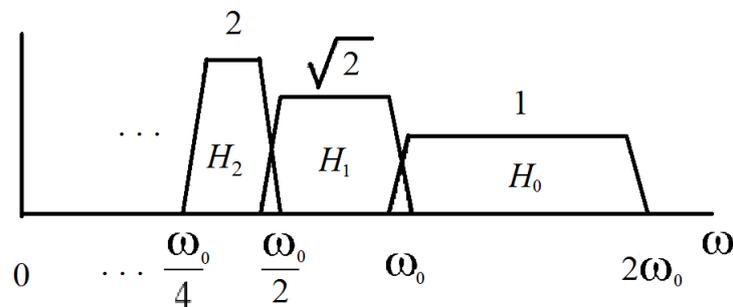


Figura 2.10: Representación gráfica de la respuesta de un banco de filtros en el dominio de la frecuencia para señales unidimensionales.

El otro problema, plantea que habrá situaciones donde se requiere anular o dejar pasar ciertas bandas de frecuencias, tal y como se ve en la Fig. 2.10. Observe que el efecto deseado, mostrado por la Fig. 2.10 no puede obtenerse a través de un solo filtro, en lugar de esto, se requieren de varios filtros que se estructuran de forma conveniente para dar una sola respuesta del sistema a la señal de entrada. A esta técnica se le denomina bancos de filtros, la cual es un filtrado selectivo compuesto por varios filtros que discriminan partes específicas de la señal. La aplicación más común de esta técnica se encuentra en la ecualización del sonido [31]. En esta tesis, estamos interesados en el uso de bancos de filtros, para el reconocimiento y clasificación confiable de objetos en imágenes.

### 2.8.1. Aplicaciones de bancos de filtros

Uno de los propósitos más comunes de un filtro es eliminar el ruido de una señal, en donde el ruido es cualquier proceso no deseado que distorsiona la señal de interés. El ruido puede ser cualquier proceso electromagnético, incluyendo alguna banda del espectro de la luz visible.

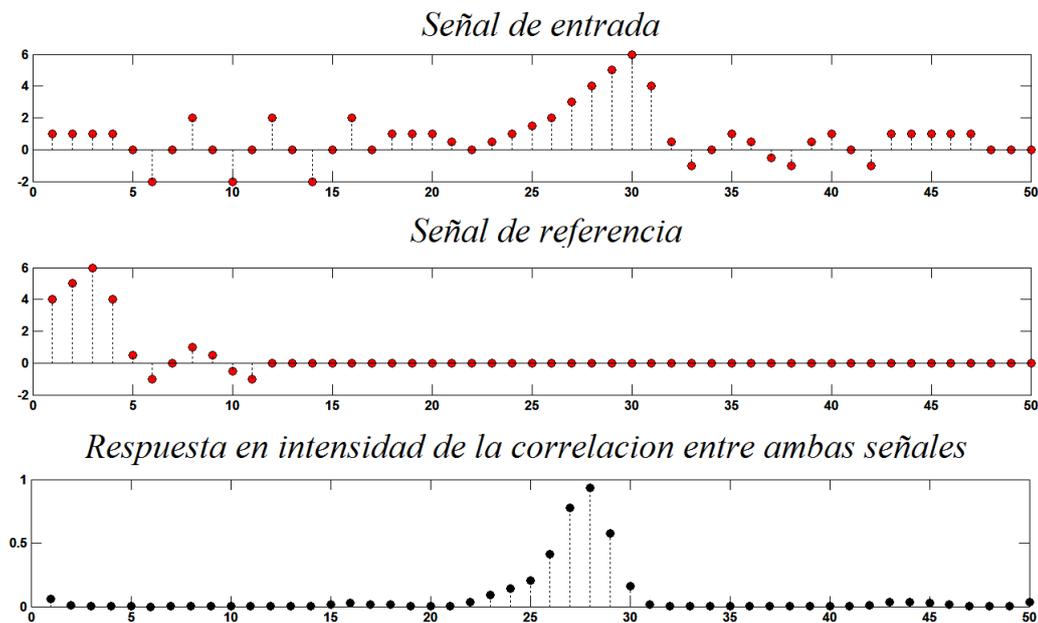


Figura 2.11: Filtro de correlación en una dimensión.

En imágenes por ejemplo, el ruido puede significar corrupción de zonas de alta energía (bordes) y de baja energía (texturas), así como la presencia de alguna distribución de ruido aleatorio en la señal. El ejemplo más común para la reducción de ruido utilizando bancos de filtros es el uso de la transformada Wavelet [32]. En esta transformada, la imagen de entrada se descompone en diferentes sub-bandas de frecuencias tanto en

el eje vertical, como en el horizontal, para su posterior análisis y procesamiento en diferentes bandas de frecuencia y en diferentes resoluciones, de forma independiente.

### 2.8.2. Bancos de filtros de correlación

En el presente trabajo los filtros diseñados tienen como objetivo eliminar toda la información de la señal, excepto en las coordenadas donde se encuentra ubicado el patrón deseado (objeto a ser reconocido). Esto puede lograrse con la ayuda de la operación de correlación, la que sirve como un estimador estadístico de máxima-verosimilitud de la posición del objeto deseado en la escena [33] (véase Fig. 2.11).

En situaciones en la que varios objetos de diferentes clases están incluidos en la misma escena y surge la necesidad de reconocerlos y clasificarlos, es necesario utilizar bancos de filtros (ver Fig. 2.12). En este caso, se requiere crear un filtro por cada clase de objetos que se quiera reconocer, entonces esto implica que se necesita hacer al menos  $N$  correlaciones para reconocer y clasificar  $N$  diferentes clases de objetos.

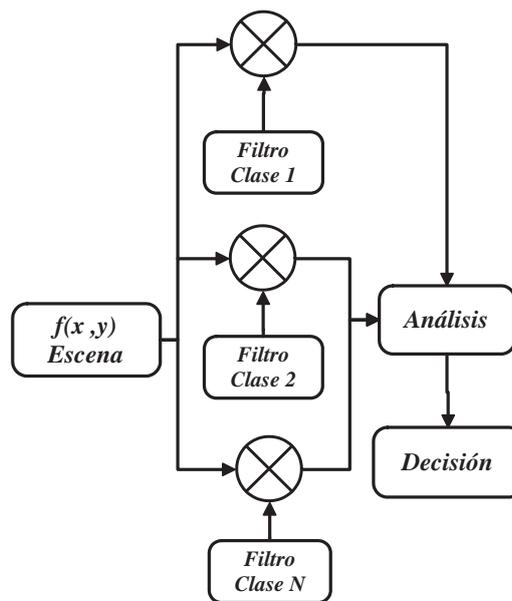


Figura 2.12: Esquema básico de un banco de filtros de correlación.

## 2.9. ESTRATEGIAS DE OPTIMIZACIÓN

La optimización matemática tiene como objetivo encontrar el mejor elemento entre un conjunto de alternativas disponibles [34]. Esto significa el poder resolver problemas en los cuales se busca maximizar o minimizar una función real que representa la calidad de cada solución factible de un problema, a través de la selección de valores de forma sistemática de variables de un conjunto conocido. En la formulación de un problema de optimización intervienen dos elementos básicos:

- Una *función objetivo* que representa al espacio de búsqueda, que generalmente contiene variables reales.
- *Restricciones* que pueden o no, ser funciones que restringen el conjunto de valores que puede tomar la función objetivo.

A modo de ejemplo, considere el siguiente problema de optimización.

Minimizar  $f(x)$   
 Sujeto a  $x \in \Omega$  donde "Ω" es el conjunto de restricciones o conjunto factible.

La función  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  que se desea minimizar es una función de valores reales, y es llamada función objetivo, o función de costo. El vector  $x$  es un vector de  $n$  variables independientes, tal que  $x = [x_1, x_2, \dots, x_n]^T \in \mathbb{R}^n$ . Las variables  $x_1, x_2, \dots, x_n$  son usualmente conocidas como variables de decisión. El conjunto factible  $\Omega$  es un subconjunto de  $\mathbb{R}^n$ . El problema de optimización anterior puede ser visto como un problema de decisión que involucra encontrar el mejor vector  $x$  de las variables de decisión sobre todos los posibles vectores en  $\Omega$ . Por el "mejor" vector, nos referimos al resultado de menor valor en la función objetivo [35]. Este vector es llamado minimizador de  $f$  sobre  $\Omega$ .

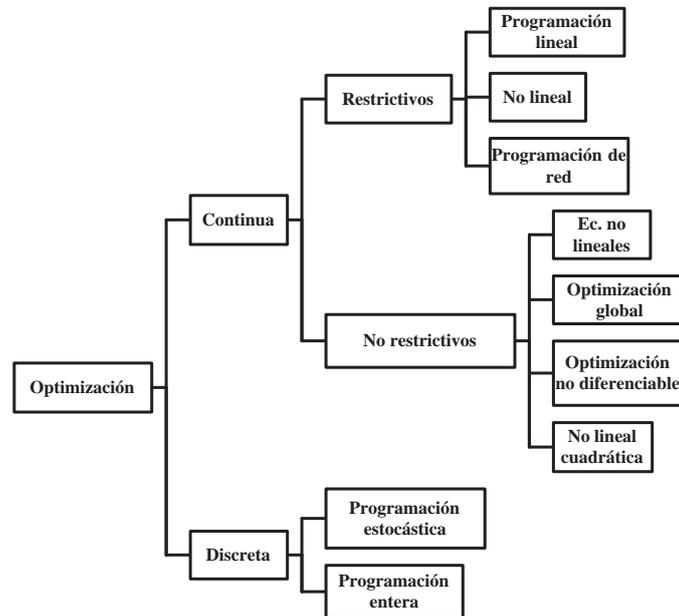


Figura 2.13: Distribución de técnicas de optimización computacional.

Es posible que tal vez existan varios minimizadores. En este caso, la búsqueda de cualquiera de estos minimizadores sera suficiente. Hay también problemas de optimización que requieren de la maximización de la función objetivo. Sin embargo, estos problemas, pueden ser representados en la forma anterior por que la maximización  $f$  es equivalente a minimizar  $-f$ . Por lo tanto, podemos limitar nuestra atención a los

problemas de minimización sin pérdida de generalidad. El problema anterior es una forma general de un problema de optimización con restricciones, por que las variables de decisión están restringidas a un conjunto restringido  $\Omega$ . Si  $\Omega = \mathbb{R}^n$ , entonces, el problema de optimización pasaría a ser sin restricciones. En el área de optimización se encuentra una gran variedad de problemas con diversas características. En la Fig. 2.13 se muestran las principales áreas de investigación dentro de la optimización. Un gran número de algoritmos propuestos para resolver problemas no convexos [36], no son capaces de hacer distinción entre un óptimo local y soluciones óptimas globales.

En el espacio de búsqueda de una función objetivo suele haber dos tipos de soluciones, soluciones locales y una solución global. Consideremos una función objetivo unidimensional dentro del espacio de búsqueda mostrado en la Fig. 2.14, que contiene tres soluciones para el caso de la minimización de la función  $f(x)$ . Es de notar que  $f(x)$  tiene tres mínimos, donde el primero representado por  $x_1$  es la solución global por que es la mejor solución para toda la función  $f(x)$ ;  $x_1$  y  $x_2$  son soluciones locales.

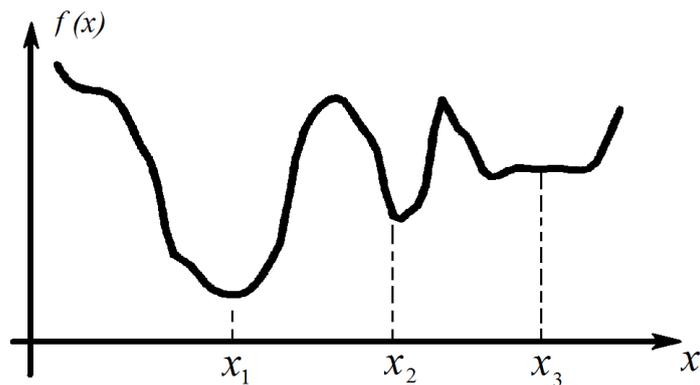


Figura 2.14: Ejemplo de una función objetivo con solución global y soluciones locales.

### 2.9.1. Búsqueda local

En ocasiones el espacio de búsqueda suele ser tan grande que buscar la solución óptima puede llevar mucho tiempo (días o tal vez meses). Hay situaciones en donde no es prioritario encontrar la mejor solución. A veces es preferible encontrar una solución suficientemente buena que pueda obtenerse en un tiempo razonable. Dicha solución puede ser evaluada para que cumpla con una función de costo. Si la solución alcanza el valor deseado en la función de costo, se puede decir que hemos encontrado una solución óptima “localmente”, lo que significa que es suficientemente buena pero no necesariamente la mejor de todas. La búsqueda local se enfoca en algoritmos que encuentran soluciones suficientemente buenas en tiempos razonables. Algunos ejemplos de este tipo de algoritmos son escalando la colina [37] (*hill climbing*), gradiente descendiente [38], búsqueda tabú [39], entre otros. Cada uno de estos algoritmos tiene características

que los hacen los mas adecuados para cierta clase de problemas de optimización. Aquí se describirá el algoritmo de *hill climbing*, y su versión probabilística.

### 2.9.1.1. Escalando la colina

El algoritmo escalando la colina (*hill climbing* por sus siglas en inglés), es un algoritmo iterativo de búsqueda local, el cual está basado en una idea simple: si se encuentra una solución que suponga una mejora respecto a la solución inicial, entonces dicha solución se toma como la solución actual, se desecha la anterior, y se busca nuevamente una solución que mejore la solución actual; este proceso se repite iterativamente hasta que no pueda encontrarse una solución que mejore la solución actual. Una mejor descripción del algoritmo *hill climbing* está dada por el Alg. 2.1. La ventaja principal de este algoritmo es su simplicidad por lo que implementarlo cuesta poco tiempo y es capaz de dar soluciones aceptables en un tiempo muy corto siempre y cuando se utilice una buena solución inicial, y un criterio adecuado para construir los vecindarios de búsqueda. Las características de la función heurística y la solución inicial determinan el éxito y rapidez de la búsqueda. La principal desventaja del algoritmo *hill climbing* es que puede quedarse estancado con facilidad en una solución local no muy buena, si se escoge una inadecuada solución inicial. La estrategia del algoritmo hace que la búsqueda pueda acabar en un punto donde la solución sólo sea óptimo local. El algoritmo *hill-climbing* encuentra problemas al enfrentarse con espacios de búsqueda que tienen valles, mesetas y crestas (ver Fig. 2.15).

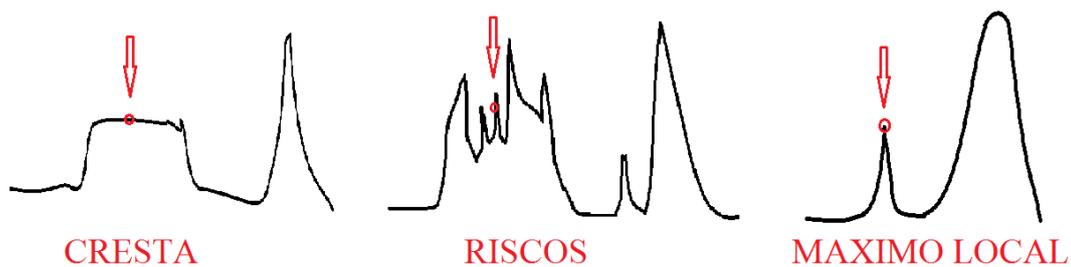


Figura 2.15: Problemas de *hill climbing*

Debido a que se trata de un algoritmo que se mueve entre vecinos solo mejora una solución a la vez, si las soluciones con casi idénticas, tomará un tiempo considerable en dar con la solución optima o tal vez no dará. Por el contrario, si las soluciones no presentan ninguna tendencia positiva o negativa, también genera este problema (riscos). Si el algoritmo *hill climbing* se utiliza en combinación con otros algoritmos heurísticos es posible aminorar parte de los problemas que este algoritmo presenta. Entre los algoritmos que puede considerarse variantes de hill climbing están el temple simulado (Simulated annealing) [40] y los algoritmos geneticos [41].

---

**Algoritmo 2.1** Pseudo algoritmo *hill climbing*

---

```
Actual ← Estado inicial
lin ← falso
mientras no fin hacer
  Hijos ← generar_sucesores(Actual)
  Hijos ← ordenar y eliminar peores (Hijos, actual)
  si no vacio?(Hijos) entonces
    Actual ← Escoger mejor (Hijos)
  sino
    fin ← cierto
  fin
fin
```

---

El algoritmo *hill climbing* estocástico es una variante de los algoritmos *hill climbing* determinísticos [42] tal como el simple *hill climbing* (Mejor 1er. vecino), el paso descendiente *hill climbing* (mejor vecino), y de las mejores aproximaciones *hill climbing* paralelo y *hill climbing* aleatorio con reinicio. El algoritmo *hill climbing* probabilístico surgió de la necesidad de hacerle frente a los problemas de estancamiento que tienen los algoritmos *hill climbing* determinísticos.

---

**Algoritmo 2.2** Pseudocódigo de *hill climbing* estocástico

---

```
Input  $Iter_{max}$  ProblemSize
Output: Current
Current ← RandomSolution(ProblemSize);
foreach  $iter \in Iter_{max}$  do
  Candidate ← RandomNeighbor(Current);
  if Cost(Candidate)  $\geq$  Cost(Current) then
    Current ← Candidate;
  end
end
return Current;
```

---

La estrategia de esta variante se enfoca en elegir una solución candidata aleatoria al comienzo del algoritmo, y a partir de la solución inicial elegida busca otra solución aleatoriamente y si es mejor se queda como actual hasta que no encuentra mejor solución o termina con cierto número iteraciones. Es decir no examina todos los vecinos antes de decidir cómo moverse. En vez de eso, selecciona a un vecino al azar, y decide (en base a la cantidad de mejora en ese vecino) si se debe seguir por ese vecino o examinar otro. A pesar de la modificación estocástica aún se puede atorar o ciclar en óptimos locales. El procedimiento puede ser aplicado a múltiples soluciones candidatas actuales, permitiendo ejecutar múltiples ejecuciones del algoritmo, para obtener diversas realizaciones al mismo tiempo (*hill climbing* paralelo).

## 2.9.2. Búsqueda global

La brecha de las matemáticas aplicadas y el análisis numérico que estudia el desarrollo de algoritmos determinísticos que son capaces de garantizar la convergencia en un tiempo finito a la solución óptima de un problema no convexo es llamada optimización global. La búsqueda global se enfoca entonces en encontrar la mejor solución de todas. En esta sección se revisarán algunos de los principales algoritmos de optimización global, como son el temple simulado (*simulated annealing*) y los algoritmos genéticos. Ambos algoritmos son llamados meta heurísticos (son algoritmos heurísticos optimizados y que incorporan un guía en la búsqueda) y se aplican a problemas reales con bastante éxito. Este tipo de algoritmos, utilizan en una u otra manera además alguna de las variantes del algoritmo *hill climbing* para llevar a cabo la adaptación evolutiva (o física) en la cual están inspirados.

### 2.9.2.1. Temple simulado

Este algoritmo de tipo *hill climbing*, está inspirado en el proceso de enfriamiento de los metales [40]. Aquí, el proceso físico en el cual se basa esta estrategia dice que un metal calentado a altas temperaturas, si se deja enfriar lentamente al solidificarse o cristalizarse se hará más sólido y por lo tanto más resistente, debido a que sus átomos optimizan su estructura cristalina. La relación que guarda con la optimización es que los átomos en su estado de alta excitación se mueven aleatoriamente y al enfriarse lentamente encuentra un estado donde la energía es la mínima (un óptimo). Este algoritmo de optimización lo que tiene como esencia es la distribución de probabilidad dada de la misma manera que en el temple simulado (distribución de *Bolzmán*). Consideremos la expresión

$$P(\text{aceptación}) = \exp\left(\frac{e - e_i}{T_i}\right), \quad (2.9.1)$$

cuyos parámetros que se extrapolan a un problema de optimización son; la temperatura  $T_i$ , la cual va decrementándose con el tiempo.  $e$  es la energía (o costo de la solución), y  $e_i$  es la energía de la solución candidata a ser considerada. Básicamente lo que la Ec. (2.9.1) hace es que por medio del parámetro  $T_i$  cambia de manera aleatoria por la función objetivo y si se encuentra una buena solución la toma y le asigna una probabilidad de aceptación. Después, vuelve a cambiar aleatoriamente, y si la solución es mala también la considera pero con una probabilidad muy baja. Esto evita que la solución se estanque en una solución mediocre y entre a mejores máximos locales, ya que el valor inicial de  $T_i$  impide que al principio el algoritmo se estanque. Esto es debido, a que todas las probabilidades asignadas a estos puntos son bajas, lo que ocasiona que todos los valores sean considerados. Sin embargo, al decrementar  $T_i$ , el

valor de la probabilidad se hace mas grande, lo que le impide cambiar aleatoriamente si encuentra un valor peor. Lo anterior se traduce en que el algoritmo se comporte primero como un algoritmo de caminata aleatoria (busca el mejor máximo local) y al descender la temperatura como el algoritmo *hill climbing* busca el mejor valor del mejor máximo local. Una comparación del algoritmo hill-climbing con el temple simulado se muestra en la Fig. 2.16.

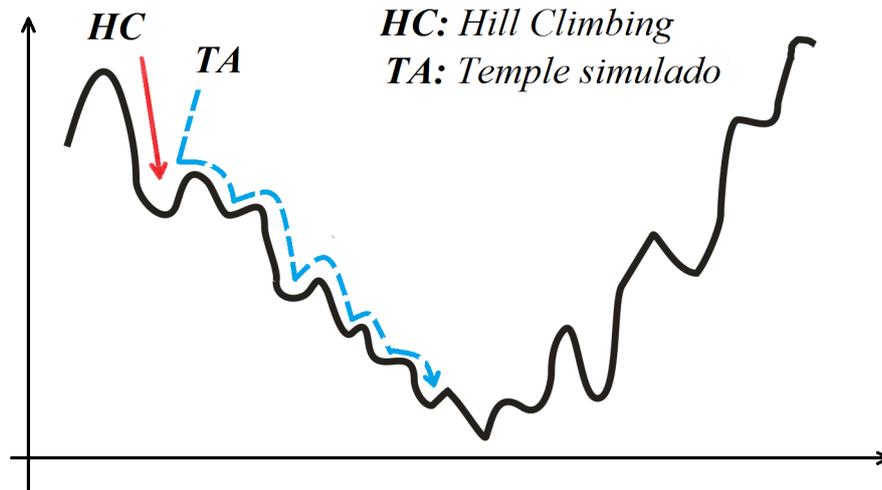


Figura 2.16: Trayectorias de búsqueda de los algoritmos hill-climbing y temple simulado.

### 2.9.2.2. Algoritmos genéticos

Los Algoritmos genéticos (denominados originalmente “planes reproductivos genéticos”) fueron desarrollados por John H. Holland a principios de los 1970s [43, 44], motivado por resolver los problemas de aprendizaje en maquinas. Holland, desarrolla la idea de la adaptación en sistemas artificiales y naturales. Los algoritmos genéticos son considerados como un algoritmo de optimización global, su principal ventaja respecto a otros algoritmos es que trabaja de manera paralela y la búsqueda es guiada de manera determinística y estocástica a la vez. Todas estas características están fundamentadas en el uso de población como soluciones, los procesos de recombinación y mutación de las soluciones. Por lo cual es capaz de encontrar una de la mejores soluciones del espacio de búsqueda en poco tiempo, resolver problemas de búsqueda multi-objetivo y resolver los problemas de funciones objetivo excesivamente complejas. Un algoritmo de su funcionamiento se presenta a continuación:

---

**Algoritmo 2.3** Seudocódigo para el algoritmo genético

---

**Input:**  $Population_{size}$ ,  $Problem_{size}$ ,  $P_{crossover}$ ,  $P_{mutacion}$ **Output:**  $S_{best}$  $Population \leftarrow \text{INITIALIZEPOPULATION}(Population_{size}, Problem_{size});$  $\text{EVALUATEPOPULATION}(Population);$  $S_{best} \leftarrow \text{GETBESTSOLUTION}(Population);$ **while**  $\text{STOPBESTCODITION}()$  **do**     $Parents \leftarrow \text{SELECTPARENTS}(Population, Population_{size});$      $Children \leftarrow \emptyset;$     **foreach**  $Parents1, Parents2 \in Parents$  **do**         $Child1, Child2 \leftarrow \text{CROSSOVER}(Parent1, Parent2, P_{crossover});$          $Children \leftarrow \text{MUTATE}(Child1, P_{mutacion});$          $Children \leftarrow \text{MUTATE}(Child2, P_{mutacion});$     **end**     $\text{EVALUATEPOPULATION}(Children);$      $S_{best} \leftarrow \text{GETBESTSOLUTION}(Children);$      $Population \leftarrow \text{REPLACE}(Population, Children);$ **end****return**  $S_{best}$  ;

---

# 3

## FILTROS DE CORRELACIÓN

### 3.1. ARQUITECTURAS DEL PROCESADORES ÓPTICOS

Una de las ventajas más notables del uso de filtros de correlación es que pueden implementarse en el dominio de la frecuencia a través del cálculo de la TF. La transformada de Fourier de una imagen observada, puede calcularse ópticamente empleando elementos difractivos convencionales, o bien, a través del uso de algoritmos rápidos que se ejecutan en una computadora. Los sistemas ópticos para el procesamiento de imágenes proporcionan una alta velocidad de operación respecto a aquella alcanzada por los sistemas digitales. Sin embargo, los sistemas digitales tienen la ventaja de ser más flexibles ya que comúnmente pueden ser reprogramados con facilidad. En esta sección se verá una breve descripción de los procesadores ópticos analógicos. Es bien conocido, que el uso de la luz para el procesamiento de información tiene un gran potencial, ya que los sistemas ópticos poseen la propiedad de superparalelismo, que es lo que hace posible el procesamiento de grandes volúmenes de información a alta velocidad. La propiedad de superparalelismo consiste en que el cálculo de cada pixel de la imagen a ser procesada se realiza de forma simultánea siguiendo una trayectoria óptica impuesta por el arreglo físico de los componentes difractivos.

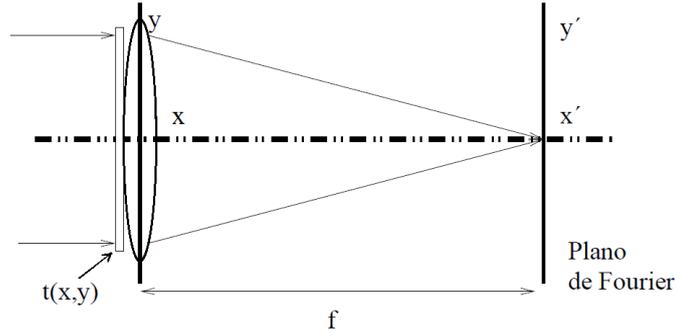


Figura 3.1: Difracción de Fraunhofer producida por una transparencia.

El cálculo de la transformada de Fourier por medios ópticos está basada en la teoría escalar de la difracción de Fraunhofer [45]. Con la aparición del láser [46, 47], es decir luz coherente, se aprovecha la capacidad que posee una lente convergente para implementar de manera práctica la TF, la cual puede ser unidimensional o bidimensional. La Fig. 3.1 muestra un arreglo óptico para el cálculo de la transformada de Fourier en base a la teoría de la difracción de Fraunhofer. Se puede demostrar que la intensidad  $I(x', y')$  observada sobre el plano focal  $(x', y')$  de una lente convergente con distancia focal  $f$  (Plano de Fourier), está dada por:

$$I(x', y') = \left\| \iint_{-\infty}^{\infty} t(x, y) \exp\left[-\frac{ik}{f}(xx' + yy')\right] dx dy \right\|^2, \quad (3.1.1)$$

donde  $k = \frac{2\pi}{\lambda}$  es el módulo del vector de onda, y  $\lambda$  es la longitud de onda de la luz. Nótese que la Ec. (3.1.1) es la TF de la función de transmisión  $t(x, y)$ , calculada para las frecuencias espaciales  $f_x = x'/\lambda f$  y  $f_y = y'/\lambda f$ , por lo tanto se puede decir que el plano  $I(x', y')$ , puede reescribirse de forma compacta, como se ve en la ecuación siguiente:

$$I(x', y') = \|F\{t(x, y)\}\|^2, \quad (3.1.2)$$

donde  $F\{\}$  es el operador que genera la transformada de Fourier. En este contexto, las frecuencias espaciales bajas son aquéllas que se localizan a una distancia cercana al origen (eje óptico) del plano de Fourier  $(x', y')$ . Las frecuencias espaciales bajas dan información sobre los detalles gruesos (texturas) de la imagen  $t(x, y)$ . Las frecuencias altas sirven para definir los detalles finos (bordes) de  $t(x, y)$ . Estos conceptos descritos anteriormente han sido ampliamente explotados por los sistemas de procesamiento de imágenes, donde hoy en día tales procesadores ópticos hacen uso de componentes electrónicos convirtiéndolos en procesadores de tipo opto electrónico híbrido que fusionan lo mejor de la óptica y la electrónica. La estructura y concepto de la TF en sistemas ópticos es la base para crear esquemas mucho más complejos como los correladores

opto-digitales. Los correladores opto-digitales mas importantes son el correlador  $4f$  y el correlador de transformación conjunta (JTC, por sus siglas en inglés). La aplicación mas trascendente de un correlador opto-digital es en el reconocimiento de patrones en imágenes, es decir, la identificación de un patrón en una escena observada y la estimación exacta de su localización dentro de la escena.

### 3.1.1. Procesador óptico $4F$

Este procesador fue el primero en ser usado para el reconocimiento de objetos y fue propuesto por Vanderlug [48] en 1964. Su diagrama a bloques se ilustra en la Fig. 3.2 donde  $f(x, y)$  representa a la escena de entrada al sistema.

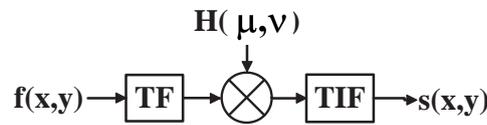


Figura 3.2: Diagrama a bloques del procesador  $4f$ .

La transformada de Fourier está dada por  $F(\mu, \nu)$ , la cual se obtiene a la salida del bloque TF (ver Fig. 3.2), posteriormente la función  $F(\mu, \nu)$  es multiplicada por la respuesta en frecuencia del filtro  $H(\mu, \nu)$ , obteniendo  $S(\mu, \nu) = H(\mu, \nu)F(\mu, \nu)$ , que es la convolución en el dominio frecuencial. Entonces por medio del bloque TIF (Transformada Inversa de Fourier) se obtiene la salida del correlador  $s(x, y)$  en el dominio del espacial.

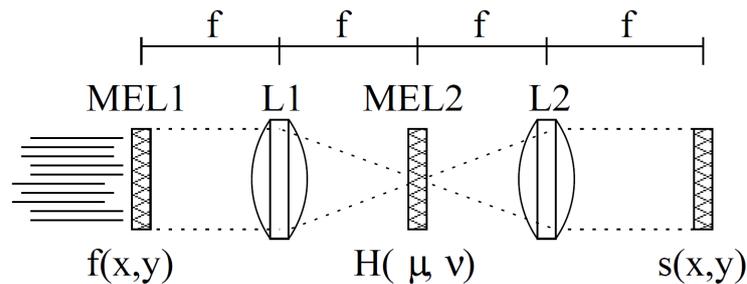


Figura 3.3: Implementación de un correlador óptico  $4f$ .

En óptica, el procesador  $4f$  puede implementarse como muestra en la Fig. 3.3 en donde se observa un frente de onda plano (luz coherente) ilumina al Modulador Espacial de Luz (MEL) 1, el cual despliega la función  $f(x, y)$ . A una distancia  $f$  (distancia focal), se encuentra colocada la lente L1, encargada de generar la transformada de Fourier de  $f(x, y)$  a una distancia focal posterior a la lente. En esta misma posición se encuentra el MEL 2, desplegando la función de transferencia del filtro  $H(\mu, \nu)$ .

Como resultado se realiza la multiplicación entre las funciones  $F(\mu, \nu)$  ( transformada de Fourier de  $f(x, y)$  ) y  $H(\mu, \nu)$ , lo cual es la proyección de  $F(\mu, \nu)$  contra el plano  $H(\mu, \nu)$ . Posteriormente, a una distancia  $f$  esta colocada la lente L2, encarga de generar la transformada inversa de Fourier del producto  $S(\mu, \nu) = H(\mu, \nu) F(\mu, \nu)$  en el plano de salida, es decir, se obtendrá la convolución entre  $f(x, y)$  y  $h(x, y)$  (respuesta al impulso del filtro) que es representada por  $s(x, y)$ . Para realizar el reconocimiento óptico de patrones mediante el procesador 4f, debe diseñarse la función de transferencia de un filtro  $H(\mu, \nu)$ , que produzca como salida la correlación entre la entrada y la señal de referencia [24]. Posteriormente las coordenadas del valor de intensidad máxima en el plano de correlación indicara la posición del objeto deseado.

### 3.1.2. Procesador óptico JTC

El correlador de transformación conjunta (JTC) fue propuesto por Weaver y Goodman [49] en 1966. Este correlador representa una técnica más novedosa aunque con un desempeño similar de aquel del correlador 4f.

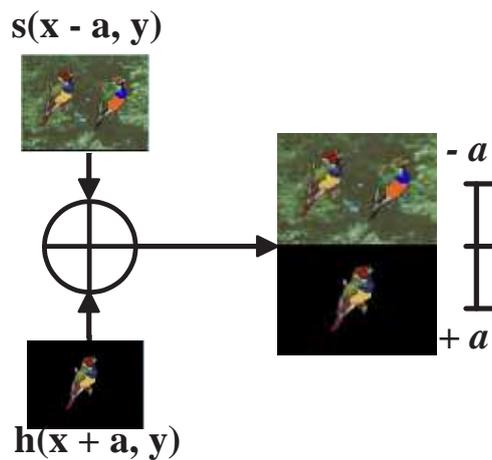


Figura 3.4: Imagen conjunta para un correlador JTC.

El procesador JTC considera el uso de dispositivos opto electrónicos como son cámaras digitales y MELs. Los MELs son dispositivos para el despliegue de imágenes y operan con la tecnología del cristal liquido. Básicamente, los MELs son pantallas semitransparentes que permiten el despliegue de imágenes cuando se utiliza una configuración adecuada de la polarización de la luz. Comúnmente, estos dispositivos son utilizados comercialmente en proyectores para computadora.

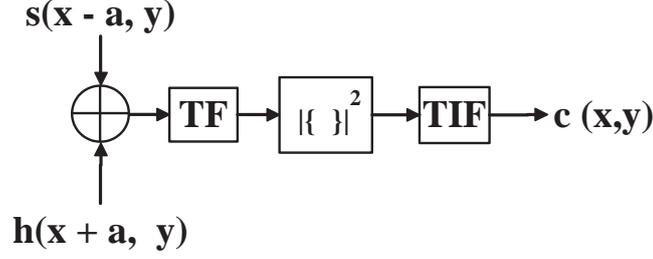


Figura 3.5: Diagrama a bloques del correlador de transformada conjunta.

La idea principal del JTC es obtener la TF de manera conjunta (al mismo tiempo) tanto para la imagen de entrada como para el patrón objetivo que se quiere localizar en la escena (véase Fig. 3.5). La arquitectura del JTC, tiene la ventaja de que es menos sensible a desalineaciones físicas (inclinación, rotación, y escala) de los dispositivos en el arreglo óptico, en comparación en el correlador 4F. Esto se debe principalmente, a que en el correlador 4f se tienen que alinear de manera precisa los planos (de la escena y de la respuesta impulso), además de que deben conservar la misma escala, para que pueda operar adecuadamente. Sin embargo, el correlador JTC posee la desventaja, de que comparte el espacio disponible del modulador de entrada entre la imagen de la escena y la respuesta al impulso del filtro de correlación, por lo que la resolución de ambas imágenes debe ser menor que si se utilizara el mismo modulador en el correlador 4f. El funcionamiento básico del correlador JTC está dado por el diagrama a bloques mostrado en la Fig. 3.5 en donde la imagen de entrada (imagen conjunta)  $f(x, y)$ , está compuesta por la imagen de la escena  $s(x, y)$  como por la respuesta al impulso del filtro  $h(x, y)$ , separados por una distancia  $2a$  respecto al origen, tal como se muestra a continuación:

$$f(x, y) = s(x + a, y) + h(x - a, y). \quad (3.1.3)$$

Es fácil darse cuenta, que la transformada de Fourier de  $f(x, y)$ , que se obtiene a la salida del bloque TF (ver Fig. 3.5), está dada por:

$$F(\mu, \nu) = S(\mu, \nu)\exp(ia\mu) + H(\mu, \nu)\exp(-ia\mu). \quad (3.1.4)$$

La intensidad del espectro conjunto:  $\|F(\mu, \nu)\|^2$ , está dado por:

$$\begin{aligned} \|F(\mu, \nu)\|^2 &= \|S(\mu, \nu)\|^2 + \|H(\mu, \nu)\|^2 + \\ &S(\mu, \nu)H^*(\mu, \nu)\exp(i2a\mu) + \\ &H(\mu, \nu)S^*(\mu, \nu)\exp(-i2a\mu). \end{aligned} \quad (3.1.5)$$

Finalmente, la salida del sistema se obtiene aplicando la transformada inversa de Fouri-

er a la Ec. (3.1.5), como a continuación:

$$c(x, y) = c_{ss}(x, y) + c_{hh}(x, y) + c_{sh}(x + 2a, y) + c_{hs}(x - 2a, y), \quad (3.1.6)$$

donde la función  $c_{sh}(x, y)$  representa la correlación cruzada entre las funciones  $s(x, y)$  y  $h(x, y)$ . A partir de la Ec. (3.1.6) podemos ver como los términos de interés (términos de correlación cruzada), aparecen en las distancias  $\pm 2a$  respecto al origen, en el plano de salida.

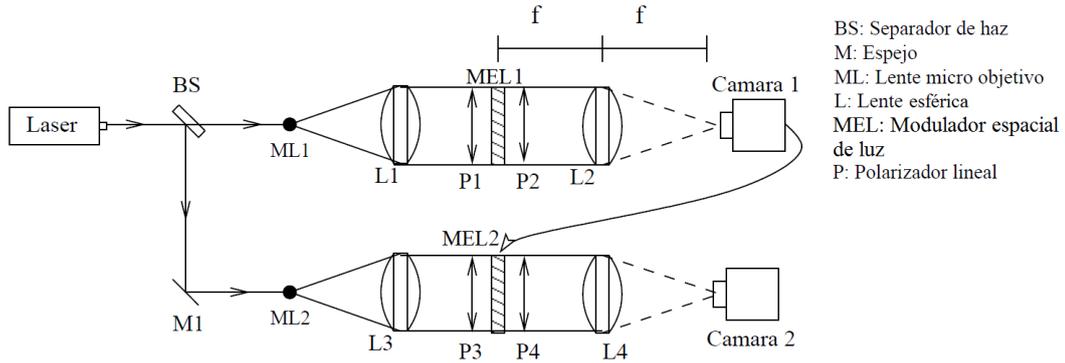


Figura 3.6: Implementación del correlador JTC.

La implementación física del JTC, se muestra en la Fig. 3.6. La arquitectura física del correlador JTC se compone principalmente de dos sistemas básicos para obtener la transformada óptica de Fourier. El haz de un láser es separado en dos, por un cubo separador (*splitter*). Cada haz es colimado independientemente al pasar por un filtro espacial y una lente esférica (ML1 y L1, para el primer sistema, y ML2 y L3, en el segundo sistema), para posteriormente iluminar cada uno de los sistemas ópticos. El MEL 1 del primer sistema, despliega la imagen conjunta de entrada. La luz modulada, a la salida del MEL1 es enfocada por la lente L2 formando la transformada de Fourier de la imagen conjunta, en el plano de la cámara 1. El espectro conjunto (intensidad de la transformada de Fourier de la imagen conjunta) es registrado por la cámara 1. La señal de salida de esta cámara, es desplegada por el MEL 2. Finalmente, de manera similar que el primer sistema, la cámara 2 registra la intensidad del plano de correlación de salida.

## 3.2. RECONOCIMIENTO DE PATRONES CON FILTROS DE CORRELACIÓN

El reconocimiento de patrones es una área de la inteligencia artificial que se dedica al estudio de algoritmos capaces de identificar y clasificar patrones. Primero, un algoritmo

mo que ejecuta una maquina debe aprender a distinguir formas interesantes de lo que se observa y posteriormente, debe llegar a categorizar en clases las diferentes formas vistas en relación con las similitudes de las características de cada forma de interés. Nótese que este es un proceso muy difícil para una maquina, ya que éstas solo cuentan con una lógica estructurada y lograr un modelo que pueda inducir, pensar, y deducir como lo hace el cerebro humano es una meta difícil de alcanzar. En este contexto, diferentes investigadores del área coinciden en que el mejor sistema de reconocimiento de objetos que existe es el cerebro humano [50, 51]. En la actualidad, existen diversas técnicas computacionales para el RP de señales multidimensionales como audio, imágenes y video, las cuales han sido extensamente estudiadas [52, 53]. Uno de los enfoques más utilizados para atacar el problema del reconocimiento de objetos, está basado en extracción de características [54]. En este enfoque, primero se extraen las características principales de los objetos de la escena de entrada, y después, las características extraídas son comparadas con aquellas características de los patrones de interés, para determinar si estos objetos pertenecen o no, a alguna de las categorías de los patrones objetivo. Este proceso se describe con más detalle en la Fig. 3.7.

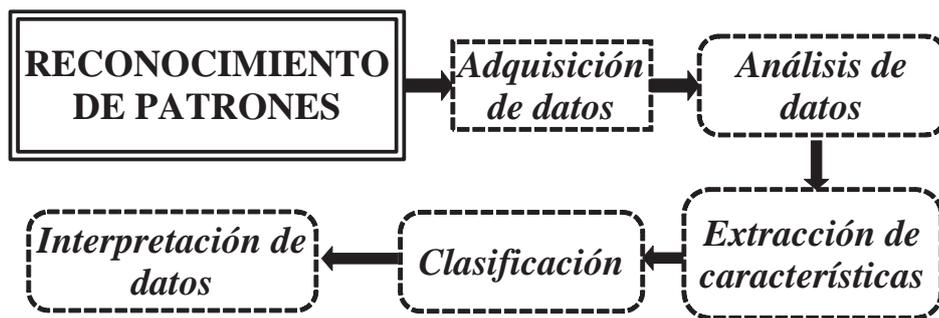


Figura 3.7: Proceso general del reconocimiento de patrones basado en características.

Dentro del procesamiento de imágenes, el RP ha sido ampliamente estudiado, donde se han desarrollado diversos algoritmos confiables, basados en el proceso descrito en la Fig. 3.7. Los algoritmos de reconocimiento de objetos basados en características, han demostrado ser efectivos en diversas aplicaciones de visión por computadora, sin embargo, estos sistemas pueden ser computacionalmente costosos, y su operación adecuada depende en cierta forma de la toma de decisiones subjetivas para cada aplicación, se basan en la propia experiencia del diseñador. Por lo tanto, hay un creciente interés por diseñar métodos efectivos de reconocimiento de objetos, para procesar imágenes de alta resolución cada vez más complejas, en sistemas automáticos de alta velocidad. Por ello, los filtros de correlación son una excelente alternativa para el reconocimiento de tiempo real. Los filtros de correlación, poseen sólidas bases matemáticas y es posible implementarlos en sistemas embebidos y en correladores opto-digitales a gran velocidad. Esto es posible ya que para su implementación se requieren solo dos operaciones

básicas, éstas son: convolución y localización del valor máximo en el plano de correlación de salida. La correlación es un proceso altamente paralelizable y la localización del pico de correlación, es una operación simple y rápida que sustituye en gran medida parte de las etapas de comparación de características de los algoritmos basados en características. Sin embargo aún existen retos por superar en el diseño de filtros de correlación.

### 3.3. MÉTRICAS DE DESEMPEÑO

Las métricas de desempeño en filtros de correlación son ampliamente usadas para el diseño y construcción de estos filtros. En esta sección se presentarán algunas de las métricas más importantes. En esta sección y posteriores, las matrices son señaladas con letras mayúsculas en negritas y los vectores por letras minúsculas en negritas, en cualquier otro caso se estará refiriendo a variables de una dimensión.

#### 3.3.1. Relación señal a ruido

La relación señal a ruido (SNR, por sus siglas en inglés) es una métrica ampliamente usada en procesamiento digital de señales, en el caso de los filtros de correlación es empleada para medir la sensibilidad del pico de correlación del filtro cuando la escena de entrada está afectada por la presencia de ruido aditivo. La SNR, puede describirse formalmente como a continuación:

$$SNR = \frac{|E\{y(0)\}|^2}{var\{y(0)\}}. \quad (3.3.1)$$

Donde  $E\{y(0)\}$  es el valor esperado del pico de correlación de salida en el origen, y  $var\{y(0)\}$  es la varianza del pico de correlación en el origen. Donde la Ec. (3.3.1) registra un valor alto, indica que el ruido tiene poca influencia sobre la señal, es decir, la señal domina al ruido. Los parámetros de promedio y varianza del pico de correlación se obtienen mediante varias realizaciones de las correlaciones por lo que la métrica no puede ser estimada usando solo un experimento de correlación.

#### 3.3.2. Error Cuadrático Promedio

El objetivo primario de los filtros de correlación es la tolerancia a la distorsión del reconocimiento de objetos en desorden. Este problema, es fácilmente resuelto por rotaciones dentro del plano y cambios de escala. El método prevaleciente para manejar distorsiones de rotación es usar un conjunto representativo de la vista del objeto.

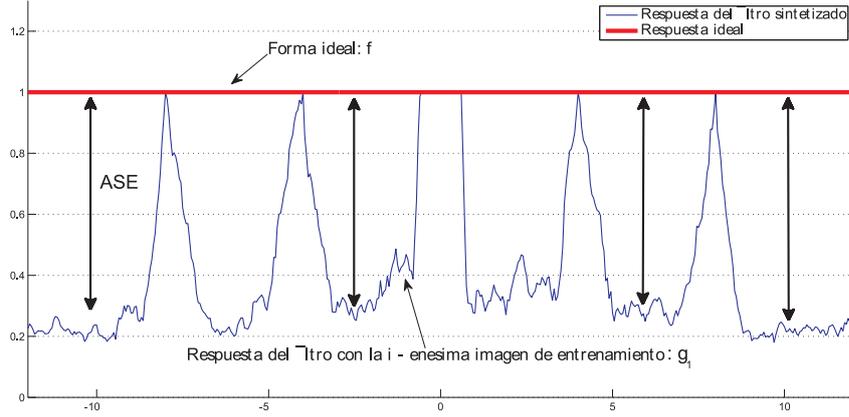


Figura 3.8: Métrica de promedio de error cuadrático.

Tradicionalmente en el diseño de filtros de correlación de tipo compuesto, las restricciones lineales son impuestas en las imágenes de entrenamiento para producir un valor conocido en un lugar específico del plano de correlación. Sin embargo, tales restricciones en el plano de correlación satisfacen condiciones solo en puntos aislados en la imagen. Pero no se menciona explícitamente el control de la capacidad del filtro para generalizar sobre todo el dominio entero de las imágenes de entrenamiento. Varios filtros muestran diferentes niveles de tolerancia a la distorsión incluso con la mismo conjunto de entrenamiento y restricciones. La razón se encuentra en el hecho de que no tratamos a las imágenes de entrenamiento como representaciones determinísticas del objetivo, si no como muestras de una clase cuyos parámetros característicos deben ser usados en la codificación del filtro. Una de las métricas desarrolladas para este fin es el error cuadrático promedio [55] (ASE, por sus siglas en inglés), la cual es una medida de la distorsión de las imágenes de entrenamiento en el dominio de frecuencial con respecto a una figura de referencia  $\mathbf{f}$ , el cual es un parámetro libre a elegir como se desee. Sea  $\mathbf{g}_i = \mathbf{X}_i \mathbf{h}$ , el  $i$ -ésimo plano de correlacion de salida (en el dominio de la frecuencia) entre  $\mathbf{X}_i$  y  $\mathbf{h}$ , donde  $\mathbf{X}_i$  en una matriz diagonal que contiene los elementos de la  $i$ -ésima imagen de entrenamiento, y  $\mathbf{h}$  es la respuesta al impulso del sistema. La forma de correlación ideal  $\mathbf{f}$  será aquella que de una respuesta ideal como la vista en la Fig. 3.8, en donde la línea verde representa la respuesta ideal de filtro ante imágenes con distorsionadas con rotación, y la línea  $\mathbf{f}$  indica que debe responder de manera uniforme ante cualquier rotación en los rangos de  $\pm 90$ . Dentro de este contexto, el error cuadrático promedio (ASE), puede definirse como:

$$ASE = \frac{1}{N} \sum_{i=1}^N (\mathbf{g}_i - \mathbf{f})^+ (\mathbf{g}_i - \mathbf{f}). \quad (3.3.2)$$

La Ec. (3.3.2) mide que tanto difieren los planos de correlación de las diferentes imágenes de entrenamiento representadas por  $\mathbf{g}_i$ , con la forma ideal  $\mathbf{f}$  que deseamos que tenga la respuesta de los diferentes planos. En estadística, a esta métrica se le conoce como error cuadrático medio (MSE, por sus siglas en inglés) y sirve para cuantificar la diferencia entre un estimador y el valor verdadero de una cantidad siendo estimada. La ASE o MSE es una función de riesgo correspondiente al valor esperado de la pérdida del error cuadrático [56].

### 3.3.3. Medida de similaridad promedio

Notemos que en la métrica ASE, se requiere de la figura ideal  $\mathbf{f}$ . Sin embargo, es difícil definir una figura ideal para todos los casos. Una opción comúnmente utilizada está dada por la figura del plano de correlación de salida que genera la imagen de entrenamiento promedio, es decir,

$$\mathbf{f}_{opt} = \frac{1}{N} \sum_{i=1}^N \mathbf{g}_i = \bar{\mathbf{g}}. \quad (3.3.3)$$

La medida de similaridad promedio (ASM), está dada por el error promedio entre los planos de correlación producidos por las imágenes de entrenamiento respecto a la figura definida en la Ec. 3.3.3, esto es,

$$ASM = \frac{1}{N} \sum_{i=1}^N (\mathbf{g}_i - \bar{\mathbf{g}})^+ (\mathbf{g}_i - \bar{\mathbf{g}}). \quad (3.3.4)$$

Es fácil darse cuenta, que un buen desempeño del filtro requiere que todas la  $i$ -ésimas correlaciones den un valor similar a la correlación de la imagen de entrenamiento promedio  $\bar{\mathbf{g}}$ .

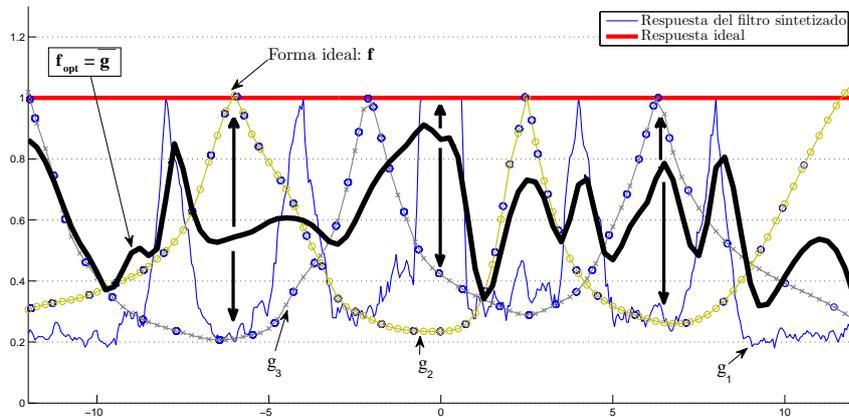


Figura 3.9: Representación gráfica de la métrica ASM.

En terminos generales, la ASM es una medida del desempeño de la tolerancia a las distorsiones geométricas de las imágenes.

### 3.3.4. Capacidad de discriminabilidad

La Capacidad de discriminabilidad (DC) mide la habilidad de un filtro para reconocer los objetos de la clase verdadera y rechazar los objetos de la clase falsa. Formalmente, la DC puede definirse como:

$$DC = 1 - \frac{|c^b(0, 0)|^2}{|c^t(0, 0)|^2}, \quad (3.3.5)$$

donde  $c^b(0, 0)$  representa el valor máximo en el plano de correlación sobre el área del fondo a ser rechazado y  $c^t(0, 0)$  es el valor máximo en el plano de correlación sobre el área del objeto. Si se obtienen valores negativos de la DC indica que el filtro probado falla al reconocer cualquier objeto.

### 3.3.5. Relación pico - lóbulo

La relación pico-lóbulo (PSR, por sus siglas en inglés) [57] permite cuantificar el realce del pico a nivel local. Básicamente, esta medida consiste en comparar el pico de correlación con sus valores adyacentes. La métrica PSR es ampliamente utilizada debido a dos razones principales:

- 1) Logra invariancia local y global a variaciones del brillo de la imagen.
- 2) Puede utilizarse para diseñar filtros de correlación con tolerancia al fondo y a objetos falsos.

Además el PSR representa una especie de métrica normalizada sobre la probabilidad de detección del patrón objetivo. Por ejemplo, un valor de 4 en la PSR indica que el pico de correlación es de 4 veces la desviación estándar por arriba de la media de los valores circundantes, con lo cual es altamente improbable que sean fluctuaciones aleatorias. La PSR, esta definida como:

$$PSR = \frac{pico - \mu}{\sigma}, \quad (3.3.6)$$

donde  $\mu$  y  $\sigma$  son los valores de la media y desviación de correlación dentro de la ventana alrededor del pico de correlación. Una segunda definición que se tiene del PSR [58], es, como una razón del pico de correlación a la desviación estándar de valores de correlación en una región centrada en el pico (usualmente de 5 x 5 píxeles), pero excluyendo una pequeña región alrededor del pico (la cual comprende un área de 20 x

20 píxeles). De esta forma, la pequeña región centrada en el pico es excluida para la determinación de los lóbulos laterales. Esto se puede ver en la figura 3.10.

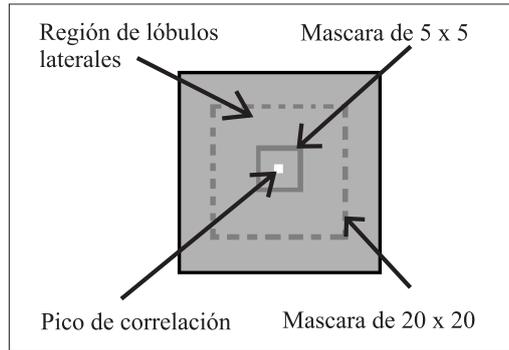


Figura 3.10: Región para calcular la PSR.

### 3.4. TAXONOMÍA DE LOS FILTROS DE CORRELACIÓN

En la presente sección se verá una clasificación general de los filtros de correlación existentes. Como es sabido no existe filtro de correlación perfecto, lo que deriva en una gran cantidad de técnicas de diseño. El desarrollo de un filtro trata siempre de solventar o mejorar un problema asociado a ellos, como los mencionados en secciones anteriores. En el área de filtros de correlación para el reconocimiento de patrones existen tres enfoques de diseño principales: analítico, compuesto, y adaptativo. El diseño analítico se basa en la optimización matemática de algún criterio estadístico [59].

Tabla 3.1: Taxonomía de los filtros de correlación.

Filtros de correlación lineales							
Restrictivos				No restrictivos			
MACE	x	x		MACH	x		
SDF	x	x		DCCF	x		
CMF		x		UMACE	x		
ASDF	x	x	x	UOTSDF	x		
OTSDF	x						
	Compuesto	Analítico	Adaptativo		C	An	Ad

El enfoque compuesto trata de solventar el problema de la tolerancia a distorsiones de los objetos, estos filtros se diseñan como una combinación de imágenes de entrenamiento [60], que deben representar a los objetos y a sus distorsiones esperadas. En el enfoque adaptativo, los filtros se diseñan con un algoritmo de adaptación que, sintetiza la respuesta al impulso del filtro, usando información del entorno [61]. La forma

más común de diseñar filtros de correlación es mediante la optimización de diferentes métricas de desempeño. Los filtros también se diseñan para que den un valor específico a la salida para los diferentes objetos que se usan como imágenes de entrenamiento.

A esta clase de filtros se le llaman filtros restrictivos, ya que tienen que cumplir con ciertas restricciones impuestas a la salida del sistema, y que son definidas en la etapa de diseño. La vertiente opuesta a este tipo de filtros son los filtros no restrictivos [55], los cuales dan un valor esperado a la salida que se define de forma probabilística. También hay otras clases de filtros, los lineales y los no lineales. En los primeros es posible realizar las operaciones en el dominio de la frecuencia, mientras que para los no lineales se tiene que hacer en el dominio del espacio.

### 3.4.1. Filtro de comparación clásico

El filtro de comparación clásico (CMF) es óptimo con respecto a la SNR para el modelo aditivo. Sea  $f(x, y)$  la imagen observada, dada por

$$f(x, y) = t(x - x_0, y - y_0) + n_a(x, y), \quad (3.4.1)$$

donde  $t(x - x_0, y - y_0)$  es el objeto deseado (en la posición arbitraria  $(x_0, y_0)$ ), y  $n_a(x, y)$  es ruido aditivo Gaussiano con media cero y desviación estándar  $\sigma_n$ . El objetivo es diseñar un filtro de correlación  $H(\mu, \nu)$ , que sea capaz de generar como salida, un valor de intensidad muy alto “1” en la posición  $(x_0, y_0)$ , y un valor muy bajo “0”, en cualquier otra parte fuera del área del objeto, es decir,

$$s(x, y) = \iint_{-\infty}^{\infty} f(x, y)h(x + \tau_x, y + \tau_y)d\tau_x d\tau_y = \left\{ \begin{array}{ll} 1; & \text{Si } (\tau_x = x_0, \tau_y = y_0) \\ 0; & \text{Otra manera} \end{array} \right\}. \quad (3.4.2)$$

Aquí,  $h(x, y)$  es la respuesta al impulso del filtro. Si la señal observada Ec. (3.4.1) estuviera libre de ruido, entonces, el filtro óptimo estaría dado por  $h(x, y) = t(-x, -y)$ , (asumiendo que  $f(x, y) \in \Re$ ). Sin embargo, debido al ruido aditivo, es inevitable que se presenten fluctuaciones en el valor de intensidad del pico de correlación, en la posición  $(x_0, y_0)$ . Este hecho motivó a Vanderlugt a definir la SNR como se muestra en la Ec. (3.3.1), ya que con esta medida es posible caracterizar las fluctuaciones del valor del pico de correlación ocasionadas por el ruido. Notemos que el numerador en la Ec. (3.3.1), se refiere a la intensidad del valor promedio del pico de correlación, tomado de un conjunto de muestras. El denominador de la Ec. (3.3.1), representa la varianza del valor del pico de correlación, tomado también de un conjunto de muestras. De esta forma, mientras mayor sea la SNR, menor será la variación del valor del pico de correlación. Asumiendo que el valor esperado de la función de ruido aditivo es igual a

cero y usando la relación de Parseval's (ver tabla 2.5), la SNR puede reescribirse como

$$SNR = \frac{\left| \int_{-\infty}^{\infty} S(\mu, \nu) H(\mu, \nu) d\mu d\nu \right|^2}{\int_{-\infty}^{\infty} P_n(\mu, \nu) |H(\mu, \nu)|^2 d\mu d\nu}. \quad (3.4.3)$$

Aquí  $S(\mu, \nu)$  y  $H(\mu, \nu)$  son las transformadas de Fourier de  $s(x, y)$  y  $h(x, y)$ , respectivamente, y  $P_n(\mu, \nu)$  es la densidad espectral de potencia del ruido  $n_a(x, y)$ . Para determinar el filtro  $H(\mu, \nu)$  que maximice la SNR, podemos aplicar la desigualdad de Cauchy - Schwarz a la Ec. (3.4.3), como a continuación:

$$SNR \leq \frac{\int_{-\infty}^{\infty} \frac{|S(\mu, \nu)|}{P_n(\mu, \nu)} d\mu d\nu \int_{-\infty}^{\infty} P_n(\mu, \nu) |H(\mu, \nu)|^2 d\mu d\nu}{\int_{-\infty}^{\infty} P_n(\mu, \nu) |H(\mu, \nu)|^2 d\mu d\nu} = \int_{-\infty}^{\infty} \frac{|S(\mu, \nu)|^2}{P_n(\mu, \nu)} d\mu d\nu. \quad (3.4.4)$$

En base a la Ec. (3.4.4), el valor máximo de la SNR se obtiene únicamente cuando

$$\sqrt{P_n(\mu, \nu)} H(\mu, \nu) = \alpha \frac{S^*(\mu, \nu)}{\sqrt{P_n(\mu, \nu)}}, \quad (3.4.5)$$

donde  $\alpha$  es una constante, resolviendo la Ec. (3.4.5) para  $H(\mu, \nu)$ , se obtiene

$$H(\mu, \nu) = \alpha \frac{S^*(\mu, \nu)}{P_n(\mu, \nu)}. \quad (3.4.6)$$

La Ec. (3.4.6) representa la respuesta en frecuencia del filtro clásico de comparación. Es fácil notar que si el ruido es blanco entonces su densidad espectral de potencia está dada por un valor constante,  $P_n(\mu, \nu) = c$ . De esta forma, el filtro estaría dado por  $H(\mu, \nu) = \alpha S^*(\mu, \nu)$ , es decir, por el complejo conjugado de la transformada de Fourier del objeto deseado, y multiplicado por una constante.

### 3.4.2. Filtro de funciones de discriminación sintética

Definamos la clase verdadera como las diferentes versiones del objeto a ser reconocido. Definamos la clase falsa de objetos, como cualquier objeto no deseado para ser rechazado por el filtro. En este caso, el filtro de correlación puede construirse como una combinación lineal de los filtros de comparación clásicos para los diferentes objetos. Los coeficientes en la combinación lineal, son determinados de tal manera que las correlaciones cruzadas en el origen del pico de correlación, deben ser igual para todos los objetos pertenecientes a una misma clase. Sea  $\{s_i(x, y) | i = 1, 2, \dots, N\}$  un conjunto de imágenes de entrenamiento (linealmente independientes), cada una con  $M$  píxeles. La respuesta al impulso  $h(x, y)$  del filtro SDF, puede expresarse como la combinación

lineal del conjunto de las imágenes de entrenamiento, es decir,

$$h(x, y) = \sum_{i=1}^N a_i s_i(x, y), \quad (3.4.7)$$

donde  $\{a_i | i = 1, 2, \dots, N\}$  son los coeficientes de peso, los cuales son seleccionados para satisfacer las siguientes condiciones:

$$\langle s_i, h \rangle = c_i. \quad (3.4.8)$$

En la Ec. (3.4.8), el símbolo “ $\langle, \rangle$ ” representa al producto interno, y  $\{c_i | i = 1, 2, \dots, N\}$  son los valores predefinidos de la salida de correlación en el origen, para cada imagen de entrenamiento. Sea  $\mathbf{S}$  una matriz con  $N$  columnas y  $M$  renglones (número de píxeles en cada imagen de entrenamiento), donde su  $i$ -enésima columna está dada por el vector imagen de  $s_i(x, y)$ . Sean  $\mathbf{a}$  y  $\mathbf{c}$  los vectores columna de  $a_i$  y  $c_i$ , respectivamente. Las Ecs. (3.4.7) y (3.4.8) pueden reescribirse en su notación vectorial como

$$\mathbf{h} = \mathbf{S}\mathbf{a}, \quad (3.4.9)$$

y

$$\mathbf{c} = \mathbf{S}^+\mathbf{h}, \quad (3.4.10)$$

donde el superíndice “+” indica transpuesto - conjugado. El  $(i, j)$ -ésimo elemento de la matriz

$$\mathbf{X} = (\mathbf{S}^+\mathbf{S}), \quad (3.4.11)$$

es el valor del origen de la correlación - cruzada entre las imágenes de entrenamiento  $s_i(x, y)$  y  $s_j(x, y)$ . Si la matriz  $\mathbf{X}$  es no singular, entonces la solución del sistema de ecuaciones lineales está dado por

$$\mathbf{a} = (\mathbf{S}^+\mathbf{S})^{-1}\mathbf{c}, \quad (3.4.12)$$

finalmente, el vector del filtro es

$$\mathbf{h} = \mathbf{S}(\mathbf{S}^+\mathbf{S})^{-1}\mathbf{c}. \quad (3.4.13)$$

El filtro SDF de la Ec. (3.4.13), puede usarse para reconocimiento de patrones con invariancia a distorsiones intraclase, es decir, para la detección de patrones con distorsiones pertenecientes a la clase verdadera de objetos. Esto puede lograrse al colocar “1” en todos los elementos del vector  $\mathbf{c}$ , es decir,

$$\mathbf{c} = [1, 1, \dots, 1]^T. \quad (3.4.14)$$

Ahora, consideraremos el problema de reconocimiento de dos clases. Supongamos que hay  $N$  imágenes de entrenamiento de la clase verdadera y  $K$  imágenes de entrenamiento de la clase falsa. Para reconocer las imágenes de entrenamiento de la clase verdadera y rechazar las imágenes de entrenamiento de la clase falsa, podemos fijar la salida del filtro como  $\{c_i = 1 | i = 1, 2, \dots, N\}$  para los objetos de la clase verdadera, y para los objetos de la clase falsa  $\{c_i = 0 | i = N + 1, N + 2, \dots, N + K\}$ , de tal forma que

$$\mathbf{c} = [1, 1, \dots, 1, 0, 0, \dots, 0]^T. \quad (3.4.15)$$

Usando el filtro de la Ec. (3.4.13) para el reconocimiento de objetos, se espera que el origen del pico de correlación sea muy cercano a “1” para los objetos de la clase verdadera, y muy cercano a “0” para los objetos de la clase falsa. Obviamente, este procedimiento puede extenderse teóricamente para cualquier número de clases. Es interesante notar que este enfoque solo puede controlar un solo valor en el plano de correlación a la salida del filtro. En consecuencia diferentes lóbulos (picos falsos) podrían aparecer en cualquier lugar, dentro del plano de correlación.

### 3.4.3. Filtro de promedio mínimo de energía de correlación

Este filtro es considerado un filtro restringido y fue diseñado para aminorar los lóbulos laterales y producir un realce del pico de correlación cuando la imagen de prueba es de la misma clase que la imágenes usadas en el diseño del filtro. Para lograr esto, el filtro MACE en su diseño permite minimizar el promedio de la energía de la correlación (*Average Correlation Energy*, ACE) [16] de las imágenes de entrenamiento mientras restringe la salida de correlación en el origen a un valor pre - especificado. La solución del filtro MACE viene dada por:

$$\mathbf{h}_{MACE} = \mathbf{D}_x^{-1} \mathbf{X} (\mathbf{X}^+ \mathbf{D}_x^{-1} \mathbf{X}) \mathbf{u}, \quad (3.4.16)$$

donde la matriz diagonal  $\mathbf{D}$  contiene los elementos del promedio de la potencia espectral de las imágenes de entrenamiento.  $\mathbf{X}$  es una matriz donde cada columna contiene una imagen de entrenamiento (en el dominio de la frecuencia) de forma vectorial reordenada de manera lexicográfica. El parámetro  $\mathbf{u}$  es un vector columna que contiene el valor de la correlación deseada en el origen del plano de correlación de salida. Este filtro tiene un muy buen desempeño, cuando se conocen de antemano las imágenes esperadas en la escena, y éstas se encuentran dentro de las imágenes de entrenamiento. Lo anterior produce realce del pico de correlación eliminando casi por completo los lóbulos, sin embargo, ante imágenes no esperadas tiene una muy mala respuesta. En consecuencia a lo anterior, se han propuesto nuevos filtros que buscan eliminar las restricciones impuestas a la salida del filtro, con el fin de mejorar la respuesta a las distorsiones

intraclase y seguir conservando la tolerancia al ruido. Este filtro recibe el nombre de filtro MACE no restringido (UMACE) [62]. El filtro UMACE, lo que hace es eliminar las restricciones ( $\mathbf{X}^+\mathbf{H} = \mathbf{u}$ ) y trabajar solamente con el promedio de correlación de la altura del pico (ACH) [63].

#### 3.4.4. Filtro de altura máxima promedio de correlación

El filtro de altura máxima promedio de correlación (*Maximum Average Correlation Height*, MACH) maximiza el peso relativo del promedio del pico de correlación con respecto a las distorsiones esperadas. El filtro MACH produce un pico de alta correlación en respuesta a la imagen de entrenamiento promedio y asegura que otras muestras de la clase verdadera exhiban similar comportamiento. Esto es logrado minimizando en el diseño la ASM de la clase verdadera, minimizando el ACE de la clase falsa, y maximizando el pico de correlación en el origen. Todo esto puede lograrse al maximizar la siguiente función objetivo:

$$J'(\mathbf{h}) = \frac{\mathbf{h}^+\overline{\mathbf{x}\mathbf{x}\mathbf{h}}}{\mathbf{h}^+\mathbf{S}_x\mathbf{h} + \mathbf{h}^+\mathbf{D}_y\mathbf{h}}. \quad (3.4.17)$$

El filtro MACH, está dado por:

$$\mathbf{h} = c(\mathbf{D}_y + \mathbf{S}_x)^{-1}\overline{\mathbf{x}}, \quad (3.4.18)$$

donde  $\mathbf{D}_y$  es una matriz diagonal que contiene los elementos de la energía promedio de correlación (ACE) de las imágenes de entrenamiento de la clase falsa representadas por  $\{y_i\}$ ,  $\overline{\mathbf{x}}$  es un vector que contiene los elementos del promedio de todas las imágenes de entrenamiento de la clase verdadera, y el termino  $\mathbf{S}_x$  representa basicamente la ASM de la clase verdadera.

#### 3.4.5. Filtro de correlación clasificador por distancia

El filtro de correlación clasificador por distancia (*Distance Classifier Correlation Filter*, DCCF) es un filtro que es capaz de manejar multi-clasificación de objetos (véase Fig. 3.11) y usa todo el plano de correlación (no solo el pico de correlación) para calcular la distancia entre la salida de correlación del filtro para la imagen de prueba, respecto a la salida de correlación que representa las diferentes clases. Como podemos ver en la Fig. (3.12), este filtro se basa en una transformación global que maximiza la distancia entre clases y compacta las respuestas del filtro para las imágenes de las misma clase

de la siguiente manera:

$$A(\mathbf{h}) = \frac{1}{C} \sum_{k=1}^C |\mathbf{m}_k^+ \mathbf{h} - \mathbf{m}^+ \mathbf{h}|^2 = \mathbf{h}^+ \mathbf{M} \mathbf{h}, \quad (3.4.19)$$

representa la separacion entre clases, y

$$B(\mathbf{h}) = \frac{1}{C} \sum_{k=1}^C \frac{1}{N} \sum_{i=1}^N \mathbf{h}^+ [\mathbf{X}_{ik} - \mathbf{M}_k] [\mathbf{X}_{ik} - \mathbf{M}_k]^* \mathbf{h} = \mathbf{h}^+ \mathbf{S} \mathbf{h}, \quad (3.4.20)$$

representa la compactacion de las respuestas del filtro para las imagenes de una misma clase. La transformación global que plantea este enfoque, deja a un filtro de correlación para cada clase (banco de filtros). El uso de los filtros DCCF es similar al uso de los filtros de correlación convencionales a excepción del paso final (ver Fig. (3.11)). El diseño de los filtros DCCF está basado en maximizar la separación entre clases, respecto a la clase promedio.

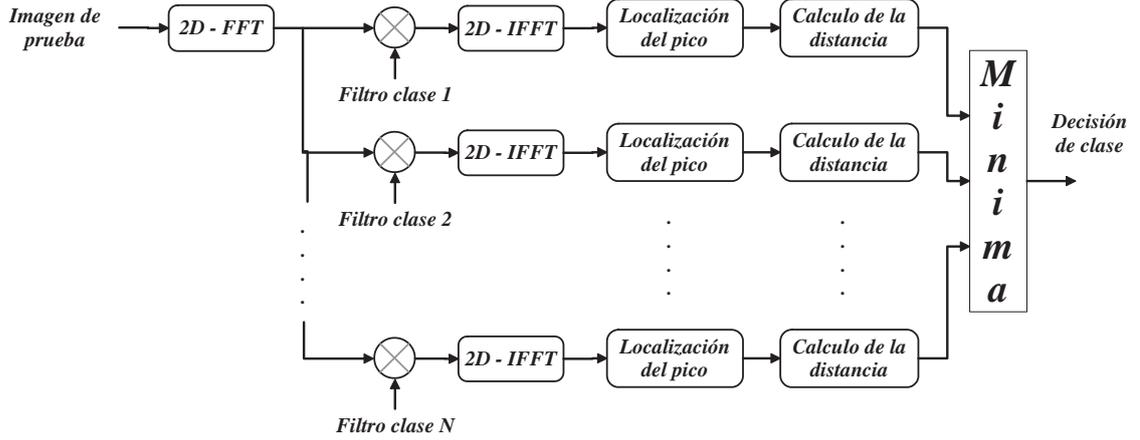


Figura 3.11: Esquema de operación de un filtro DCCF.

En términos mas formales, el filtrado DCCF debe maximizar la siguiente función objetivo:

$$J(\mathbf{h}) = \frac{A(\mathbf{h})}{B(\mathbf{h})} = \frac{\mathbf{h}^+ \mathbf{M} \mathbf{h}}{\mathbf{h}^+ \mathbf{S} \mathbf{h}}. \quad (3.4.21)$$

La solución particular del filtro DCCF para dos clases, está dada por:

$$\mathbf{h} = \mathbf{S}^{-1}(\mathbf{m}_1 - \mathbf{m}_2). \quad (3.4.22)$$

Para determinar la clase a la que pertenece un objeto hay que encontrar la clase que tenga la distancia mínima  $d_k$  entre la energía de la correlación del filtro para la escena de entrada, respecto a la energía de la correlación del filtro para la imagen de

entrenamiento promedio de una clase, es decir, hay que encontrar el valor mínimo de

$$d_k = |\mathbf{H}^* \mathbf{z} - \mathbf{H}^* \mathbf{m}_k|^2, \quad (3.4.23)$$

donde  $\mathbf{z}$  es la escena en el dominio de Fourier, y  $\mathbf{H}$  es el filtro DCCF y  $\mathbf{m}_k$  la imagen promedio de la  $k$ -ésima clase.

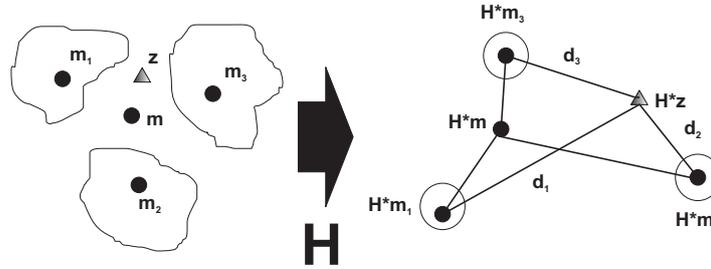


Figura 3.12: Distancias entre clases por medio de la transformación  $\mathbf{H}$  del filtro DCCF.

## 3.5. PROBLEMAS EN FILTROS DE CORRELACIÓN

### 3.5.1. Modelo disjunto

El modelo disjunto describe a una escena capturada ante la presencia de un fondo. Consideremos el caso donde una cámara registra la imagen de un objeto de interés, en condiciones arbitrarias.

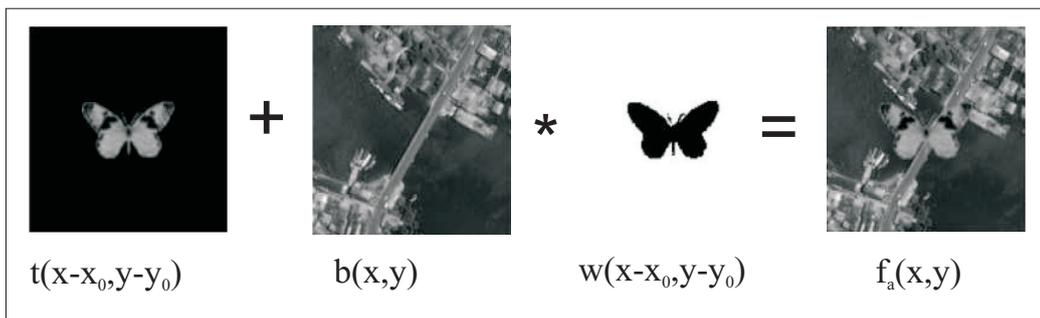


Figura 3.13: Representación del modelo disjunto.

Notemos que la imagen resultante estará formada por la imagen del objeto mas la porción del fondo en el que el objeto está incrustado (ver Fig. 3.13). Este modelo puede escribirse como se muestra a continuación:

$$f_d(x, y) = t(x - x_0, y - y_0) + w(x - x_0, y - y_0)b(x, y), \quad (3.5.1)$$

donde  $b(x, y)$  es la imagen del fondo, y  $w(x - x_0, y - y_0)$  es una función binaria definida como

$$w(x - x_0, y - y_0) = \begin{cases} 0; & \text{Dentro del \u00e1rea de } t(x, y) \\ 1; & \text{Otra manera} \end{cases} \quad (3.5.2)$$

En el modelo disjunto, la similaridad de la escena con las im\u00e1genes de entrenamiento, ocasiona l\u00f3bulos a la salida de la correlaci\u00f3n del filtro con la escena. Es de suponer que al haber mayor n\u00famero de im\u00e1genes de entrenamiento de la clase verdadera ser\u00e1 m\u00e1s probable que se presente este efecto, tal como se muestra en la Fig. 3.14.

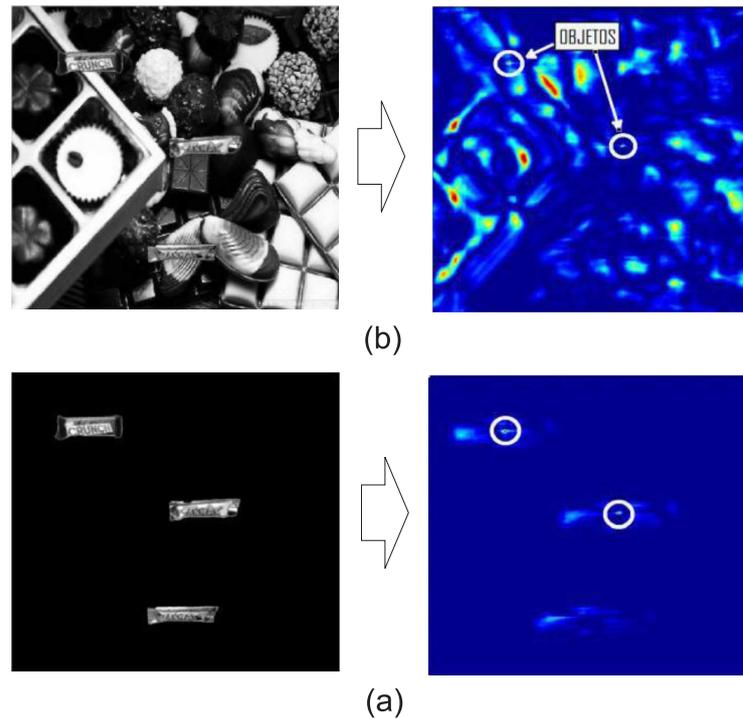


Figura 3.14: Efecto en la salida de correlaci\u00f3n de una escena bajo el modelo disjunto.

### 3.5.2. Ruido aditivo en filtros de correlaci\u00f3n

El ruido aditivo se presenta en el sensor opto-electr\u00f3nico de una c\u00e1mara por el flujo aleatorio de electrones. Generalmente, la distribuci\u00f3n del ruido aditivo es gaussiana y es llamado en la literatura AWGN (por sus siglas en ingl\u00e9s). El ruido aditivo, afecta de manera directa el desempe\u00f1o de los filtros de correlaci\u00f3n, debido a que el ruido aditivo no solo contamina la escena con peque\u00f1as perturbaciones si no tambi\u00e9n a los objetos presentes en ella. La incorporaci\u00f3n de este ruido en la imagen  $f_a(x, y)$  se expresa de manera matem\u00e1tica como se muestra a continuaci\u00f3n:

$$f_a(x, y) = t(x - x_0, y - y_0) + n_a(x, y), \quad (3.5.3)$$

donde  $n_a(x, y)$  es el ruido del sensor, y  $(x_0, y_0)$  son coordenadas arbitrarias del objeto en la escena. En la Fig. 3.15 se puede ver como la imagen original pierde calidad al incluir esta clase de ruido de manera aditiva.

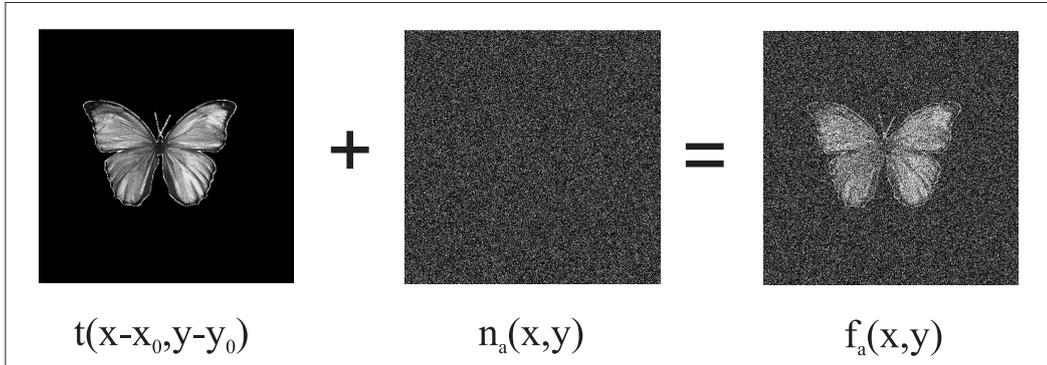


Figura 3.15: Representación gráfica del ruido aditivo.

### 3.5.3. Problema de la clasificación en filtros de correlación

Como se ha mencionado anteriormente los filtros de correlación son muy buenos para reconocer y localizar los objetos, pero tienen problemas para clasificar los objetos identificados. Esto se debe a las mismas características intrínsecas de estos filtros, ya que no son capaces de asignar a una categoría correspondiente a las localizaciones de los diversos picos de correlación generados. El problema clásico que se explica para ilustrar este hecho es con dos clases. Si bien los filtros convencionales necesitan de  $N$  correlaciones para reconocer  $N$  objetos en una imagen. En el caso de dos clases se necesitan dos correlaciones en la imagen. Dada una versión distorsionada de cualquiera de las dos clases. Es de suponerse que posiblemente sea detectada o muestre un pico de correlación en el filtro correspondiente a la clase a la que pertenece. También consideremos el caso de un objeto que no pertenezca a ninguna clase. Tal efecto se muestra en la Fig. 3.16.

En la Fig. 3.16, el pico de correlación de mayor amplitud pertenece a la clase A, sin embargo no es fácil asignar este pico a la categoría mencionada de manera automática. Debido a que para "B" también da un pequeño pico, para el mismo filtro. Y esto hace que no sea fácil discriminar la letra "B".

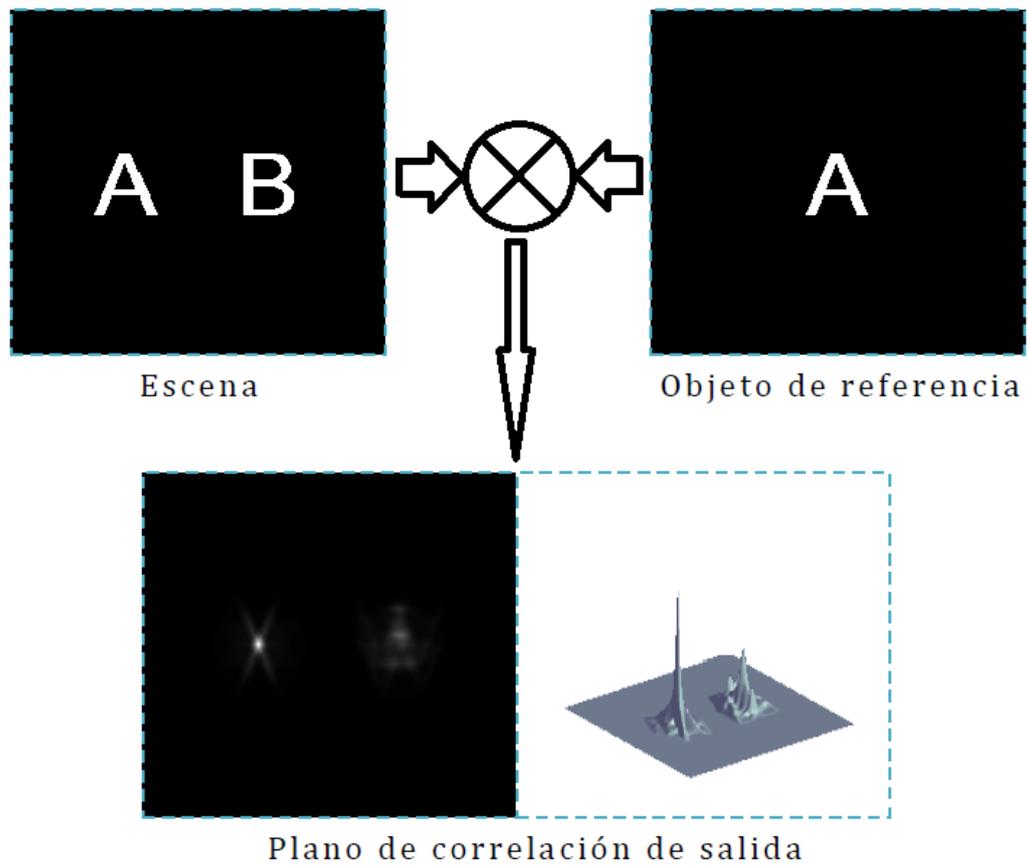


Figura 3.16: El problema de clasificación clásico usando el filtro de correlación MSF.

# 4

## FILTRO PROPUESTO DE MULTICLASIFICACIÓN

### 4.1. FILTROS DE CORRELACIÓN ADAPTATIVOS PARA LA MULTICLASIFICACIÓN EN EL RECONOCIMIENTO DE PATRONES

Suponga que  $(K + 1)N$  imágenes de entrenamiento son distribuidas en  $K + 1$  clases diferentes de objetos. Por facilidad, consideraremos que cada clase contiene  $N$  imágenes de entrenamiento y cada imagen de entrenamiento tiene  $d$  píxeles. Definamos un conjunto de objetos a ser rechazados como:

$$\{S^F\} = \{s_1^0(x, y), s_2^0(x, y), \dots, s_N^0(x, y)\}, \quad (4.1.1)$$

donde  $s_i^0(x, y)$  es la  $i$ -ésima imagen de la clase falsa; la imagen que se puede dar por cualquier objeto falso o por un fondo a ser rechazado. De manera similar, se define una clase verdadera de objetos, como todos los objetos a ser reconocidos y clasificados. Matemáticamente, la clase verdadera de objetos se define como

$$\{S^T\} = \bigcup_{k=1}^K \{S_k^T\}, \quad (4.1.2)$$

donde  $\{S_k^T\}$  es un conjunto de objetos que pertenecen a la  $K$ -ésima clase verdadera, que es,  $\{S_k^T\} = \{s_1^k(x, y), \dots, s_N^k(x, y)\}$ . La imagen  $s_i^k(x, y)$  representa al  $i$ -ésimo objeto de entrenamiento de la  $k$ -ésima clase. De acuerdo al enfoque del filtro SDF [64] un filtro compuesto puede ser construido como una combinación lineal de todas las imágenes de entrenamiento, tal como:

$$h(x, y) = \sum_{k=0}^K \sum_{i=1}^N a_{Nk+i} s_i^k(x, y), \quad (4.1.3)$$

donde  $\{a_j\}$  son coeficientes de peso que deben ser encontrados, para satisfacer las restricciones,

$$c_i^k = \sum_{i=1}^N s_i^k(x, y) h^*(x, y), \quad (4.1.4)$$

donde  $*$  significa el complejo conjugado. Las restricciones  $\{c_i^k\}$  son valores predefinidos de la correlación de salida en el origen para las imágenes de entrenamiento. Sea reescrita la Ec. (4.1.4) en un vector de la forma

$$\hat{\mathbf{c}} = [c_0, c_1, \dots, c_k]^T, \quad (4.1.5)$$

donde  $\mathbf{c}_k$  es un vector  $N \times 1$  de restricciones impuestas a el filtro, en respuesta a las imágenes de entrenamiento de la  $k$ -ésima clase. Nótese que se puede ajustar la intensidad de salida del filtro para que sea igual a la unidad cualquier objeto perteneciente a  $\{S^T\}$ , o a cero para cualquier objeto perteneciente a  $\{S^F\}$ ,

$$|\mathbf{c}_0|^2 = [0, 0, \dots, 0]^T, \quad (4.1.6)$$

donde  $|\mathbf{c}|^2$  es el operador de intensidad elemento a elemento sobre el vector  $\mathbf{c}$ . De esta manera, podemos sintetizar un filtro de correlación que identifica cualquier objeto perteneciente a la clase verdadera  $\{S^T\}$  y rechazar cualquier objeto falso perteneciente a la clase falsa  $\{S^F\}$ . Para diseñar un filtro compuesto que sea capaz de llevar a cabo la clasificación, se propone utilizar un vector complejo  $\hat{\mathbf{c}}$ . Las magnitudes de los componentes del vector contienen los valores convencionales predefinidos de la salida de la correlación en el origen de las imágenes de entrenamiento mientras que la información acerca de las clases de las imágenes de entrenamiento es codificada en las fases de los componentes del vector. Esto puede lograrse a través del uso de los siguientes vectores

complejos:

$$c_0 = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, c_1 = \begin{bmatrix} \exp(i\phi_1) \\ \exp(i\phi_1) \\ \vdots \\ \exp(i\phi_1) \end{bmatrix}, c_2 = \begin{bmatrix} \exp(i\phi_2) \\ \exp(i\phi_2) \\ \vdots \\ \exp(i\phi_2) \end{bmatrix}, \dots, c_K = \begin{bmatrix} \exp(i\phi_K) \\ \exp(i\phi_K) \\ \vdots \\ \exp(i\phi_K) \end{bmatrix}. \quad (4.1.7)$$

Aquí,  $\{\phi_k | k = 1, \dots, K\}$  son valores de fase pre-especificados para los  $k$ -ésimos objetos de la clase verdadera  $\{S_k^T\}$ . Para sintetizar un filtro SDF con restricciones complejas para el reconocimiento y clasificación de objetos, las Ecs. (4.1.3) y (4.1.4) pueden ser reescritas en una notación vector matriz, como la siguiente,

$$\mathbf{h}_{csdf} = \mathbf{R}\mathbf{a}, \quad (4.1.8)$$

y

$$\hat{\mathbf{c}}^* = \mathbf{R}^+\mathbf{h}_{csdf}, \quad (4.1.9)$$

donde  $\mathbf{R}$  es una matriz  $d \times (K+1)N$ , donde cada columna está dada por una versión vectorial de  $s_i^k(x, y)$  usando un orden lexicográfico, y  $\mathbf{a} = [a_1, \dots, a_{(K+1)N}]^T$  es el vector desconocido de coeficientes. Sustituyendo la Ec. (4.1.8) en la Ec. (4.1.9) tenemos  $\mathbf{a} = (\mathbf{R}^+\mathbf{R})^{-1}\hat{\mathbf{c}}^*$ , y el vector del filtro con restricciones complejas (si la matriz  $(\mathbf{R}^+\mathbf{R})^{-1}$  es no singular), esta dada por

$$\mathbf{h}_{csdf} = \mathbf{R}(\mathbf{R}^+\mathbf{R})^{-1}\hat{\mathbf{c}}^*. \quad (4.1.10)$$

Observe que usando el filtro SDF con restricciones complejas, se satisface la igualdad de las restricciones impuestas de la intensidad de salida por las Ecs. (4.1.5) y (4.1.6), y al mismo tiempo, podemos clasificar cualquier patrón desconocido de la escena de entrada mediante la comparación de los valores de fase obtenidos  $\hat{\phi}_k$  en las coordenadas de la intensidad máxima (pico de correlación), con los valores pre-establecidos  $\phi_k$  previamente definidos en las restricciones del filtro.

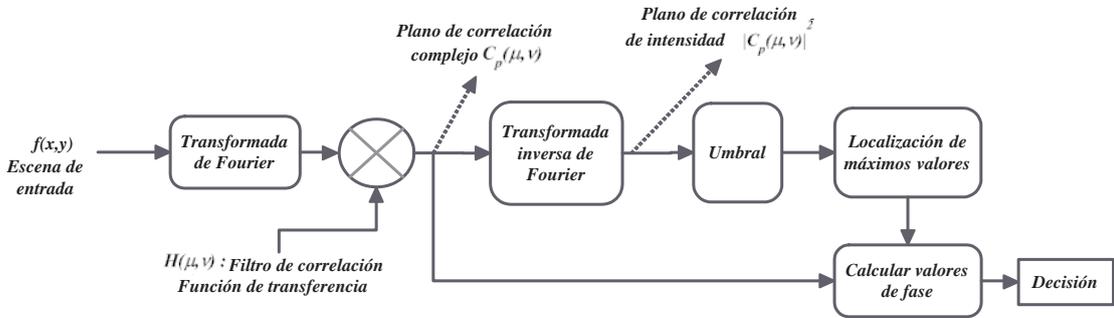


Figura 4.1: Diagrama a bloques del algoritmo propuesto para el reconocimiento y clasificación de objetos.

El diagrama de bloques del algoritmo propuesto para el reconocimiento y clasificación de los objetos se presenta en la Fig. 4.1.

Aunque el filtro SDF propuesto con restricciones complejas puede ser usado para el reconocimiento y clasificación de varios objetos, y rechazar diferentes objetos de la clase falsa con solo una operación de correlación, el filtro es incapaz de controlar el plano de salida de toda correlación. Esto se debe a que el filtro sólo puede controlar los puntos de correlación de salida en el origen de las imágenes de entrenamiento, donde las restricciones complejas deben cumplirse. Por lo tanto, la aparición de altos lóbulos laterales de correlación en el área de fondo es muy probable. Para lograr un alto desempeño en el reconocimiento y clasificación, es necesario reducir los valores de correlación de salida en todas partes, excepto los valores en el origen de las imágenes de entrenamiento donde las restricciones complejas tienen lugar.

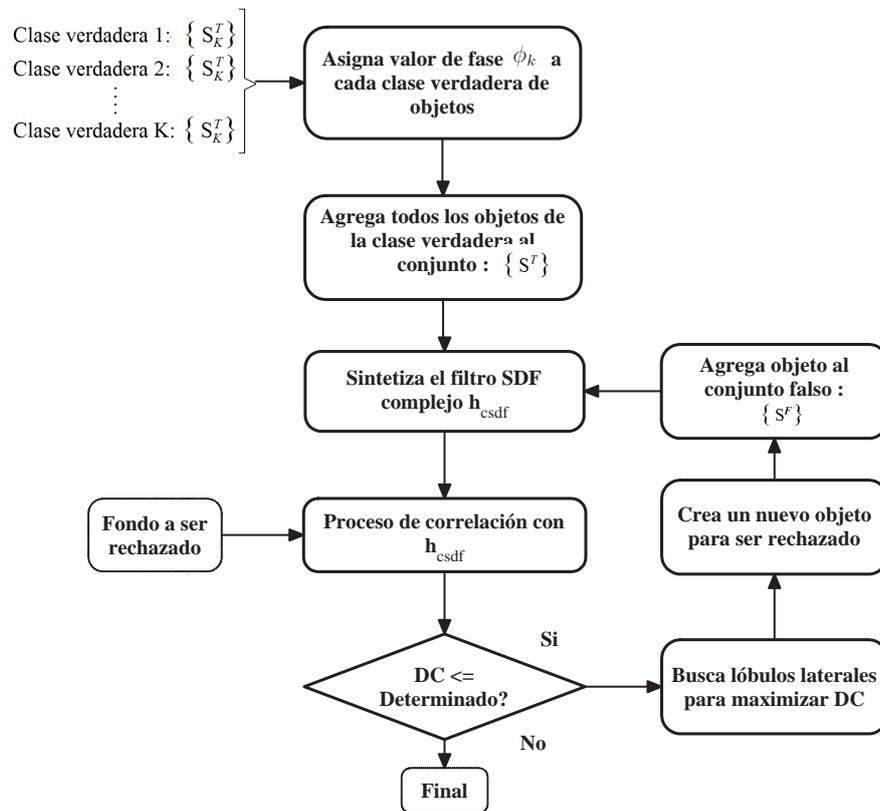


Figura 4.2: Algoritmo iterativo para el diseño de filtros adaptativos compuestos con restricciones complejas.

Esto se puede hacer mediante la explotación de un enfoque adaptativo [17, 18]. De acuerdo con este enfoque, estamos interesados en la síntesis de un filtro con un rendimiento optimizado para un conjunto fijo de patrones y de un fondo a ser rechazado, en lugar de un filtro con parámetros de rendimiento promedio para un conjunto de imágenes. Para sintetizar un filtro compuesto adaptativo se propone un algoritmo iterativo. En

cada iteración el algoritmo reduce el lóbulo mayor en el plano de intensidad de la correlación y, por tanto, monótonamente aumenta el valor de la capacidad de discriminación (DC) hasta que un valor deseado es alcanzado. Un filtro con un valor de DC cercano a la unidad posee una buena capacidad de distinción entre objetos de interés y objetos no deseados. Los valores negativos de la DC indican que el filtro no es capaz de reconocer los objetos de interés. El primer paso del algoritmo de adaptación es asignar  $\phi_k$  a cada uno de los  $K$  valores predefinidos para los objetos de la clase verdadera (véase Eq. 4.1.7). A continuación, se lleva a cabo la correlación entre el fondo y un filtro SDF básico con restricciones complejas, inicialmente entrenado con todas las versiones disponibles de cada objeto y con objetos de la clase falsa conocidos.

La función del fondo puede ser descrita determinísticamente como una fotografía, o estocásticamente por la realización de un proceso aleatorio. A continuación, buscamos un lóbulo con un valor máximo en el plano de correlación de salida que nos ayude a maximizar la DC del filtro. La posición del lóbulo es usada como origen de un nuevo objeto falso artificial con la forma similar a la del objeto de interés. El objeto falso es incluido dentro de la clase falsa  $\{S^F\}$  y un nuevo filtro SDF es sintetizado. Este procedimiento continua hasta alcanzar un valor de DC deseado. Un diagrama a bloques del algoritmo de adaptación es mostrado en la Fig. 4.2, y los pasos del algoritmo son:

**PASO 1:** Asignar el valor  $\phi_k$  para cada clase verdadera de los objetos:  $\phi_k \rightarrow \{S_k^T\}$ ; para  $k = 1, \dots, K$ .

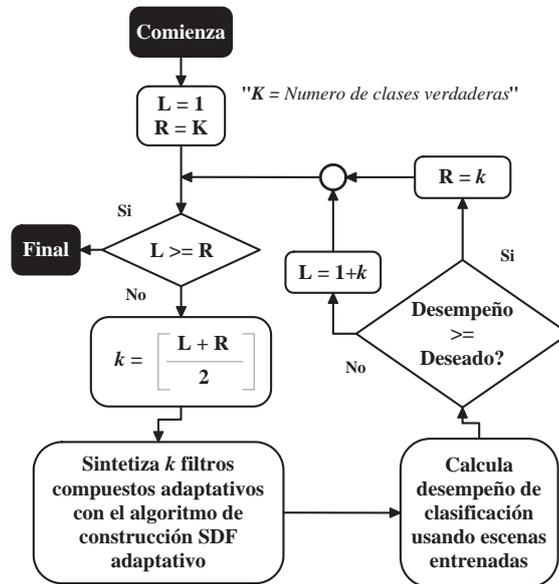


Figura 4.3: Algoritmo propuesto para la construcción de un banco de filtros con restricciones complejas.

**PASO 2:** Incluir todas las imágenes de entrenamiento disponibles en el subconjunto

correspondiente  $\{S_k^T\}$  o  $\{S^F\}$ , y construir el conjunto

$$\{S\} = \left\{ \bigcup_{k=1}^K \{S_k^T\} \right\} \cup \{S^F\}. \quad (4.1.11)$$

**PASO 3:** Sintetizar un filtro SDF con restricciones complejas entrenado para  $\{S\}$ , con la ayuda de Ec. (4.1.10) y Ec. (4.1.7).

**PASO 4:** Llevar a cabo la correlación entre el filtro adaptativo y la imagen de fondo.

**PASO 5:** Calcular la DC por cada imagen de entrenamiento de  $\{S^T\}$ . Entonces, se establece el valor mínimo de DC como el actual valor de DC del filtro, si la DC del filtro es mayor que el valor pre-especificado, el procedimiento es termina. De otra manera, ir al siguiente paso.

**PASO 6:** Encontrar el valor máximo de intensidad en el plano de correlación de salida, y alrededor de este punto construir una nueva imagen de entrenamiento para ser rechazada, con información del fondo. La región de soporte de esta nueva imagen de entrenamiento es similar a la región de soporte del objeto.

**PASO 7:** Incluir la nueva imagen falsa al conjunto  $\{S^F\}$  y actualizar el conjunto  $\{S\}$ . Después, ir a el PASO 3.



Figura 4.4: Imágenes de referencia de las clases de los objetivos: (a) Clase 1; (b) Clase 2; (c) Clase 3; (d) Clase 4.

Se puede observar, que mediante el uso del procedimiento iterativo propuesto se puede sintetizar un filtro adaptativo compuesto como el filtro SDF adaptativo con restricciones complejas (ACSDF), que es capaz de controlar todo el plano de correlación de salida asegurando picos agudos de correlación en las localizaciones de los objetos, y manteniendo bajos los niveles de correlación en toda el área del fondo. Por otra parte, observemos que una de las características más importantes de la propuesta del filtro adaptativo es que puede reconocer y clasificar varios objetos pertenecientes a diferentes clases en una operación de correlación única.

Por otra parte, estamos interesados en la construcción de un banco de filtros de correlación para el reconocimiento de patrones que pueda garantizar el reconocimiento deseado y el rendimiento de clasificación con el mínimo número posible de filtros de correlación. Para construir un sistema con los filtros adaptativos propuestos, se sugiere el uso de un procedimiento iterativo. Este procedimiento se basa en un algoritmo de

búsqueda binaria [65], y su diagrama de flujo se presenta en la Fig. 4.3. Vemos que mediante el algoritmo propuesto basado en búsqueda binaria se puede construir un banco de filtros de correlación con el desempeño deseado y con una baja complejidad computacional.

#### 4.1.1. Búsqueda en el proceso de adaptación

Parte del proceso de adaptación incluye un proceso de búsqueda y construcción de las imágenes de entrenamiento de la clase falsa, debido al uso de los filtros restrictivos, ya que con la ventaja de usar restricciones en tales filtros viene un costo en cuanto al control del plano de correlación, entonces al no tener un control sobre el plano de correlación se deben elegir las mejores imágenes de rechazo del fondo, para esto se utiliza un algoritmo de tipo *hill climbing*, combinado con otro algoritmo que produce aleatoriedad en el proceso de búsqueda evitando que las soluciones se estanquen en solo un área de la imagen. Sin embargo las soluciones caerán en las áreas que presentan mayores similitudes a los objetivos, con los cuales está construido el filtro, esto puede verse en la Fig. 4.5. Donde los puntos rojos representan los puntos elegidos por el algoritmo que se recortan con la región de soporte de la imágenes de entrenamiento de la clase verdadera, para formar los objetos a ser rechazados que pertenecerán al conjunto  $\{S^F\}$  de la clase falsa.



Figura 4.5: Puntos seleccionados de la escena para la construcción de filtro.

Notar que las áreas con bajo contraste reciben poca atención por parte del algoritmo de búsqueda. Lo anterior se traduce a que solo es necesario hacer unas cuantas correlaciones para identificar las imágenes de entrenamiento de la clase falsa que ayudarán a suprimir los lóbulos laterales causados por la escena. Por otra parte de haber objetos no deseados se tendrán que incorporar al filtro antes del proceso de adaptación.

# 5

## RESULTADOS

### 5.1. DESCRIPCIÓN DE PRUEBAS DE DESEMPEÑO DEL FILTRO PROPUESTO

Una de las consideraciones a tomar en cuenta en el diseño de un filtro, es medir el desempeño ante las situaciones para las cuales fue diseñado, comparar la mejora con respecto a los filtros existentes. Es por esto que se han realizado simulaciones por computadora, bajo distintos tipos de ruido aditivo, posiciones aleatorias de los objetivos dentro de la escena, y un análisis comparativo de la discriminabilidad y confiabilidad del filtro respecto a las clases utilizadas en las pruebas es desarrollado por medio de métodos estadísticos.

#### 5.1.1. Algoritmo de las pruebas para el desempeño del ruido en escenas reales

Parte de las pruebas a realizar en la presente tesis consiste una serie de ejecuciones del filtro propuesto bajo un determinado nivel ruido aditivo de distribución Gaussiana. El objetivo de esta prueba es medir de manera estadística como afecta el ruido de la imagen a la detección y confiabilidad del reconocimiento al filtro propuesto. Como un objetivo secundario también se busca obtener las métricas para los filtros tradicionales con el fin de comparar de manera objetiva (es decir, bajo las mismas circunstancias). El

algoritmo para realizar las pruebas bajo ruido en la imagen se muestra a continuación en la Fig. 5.1.

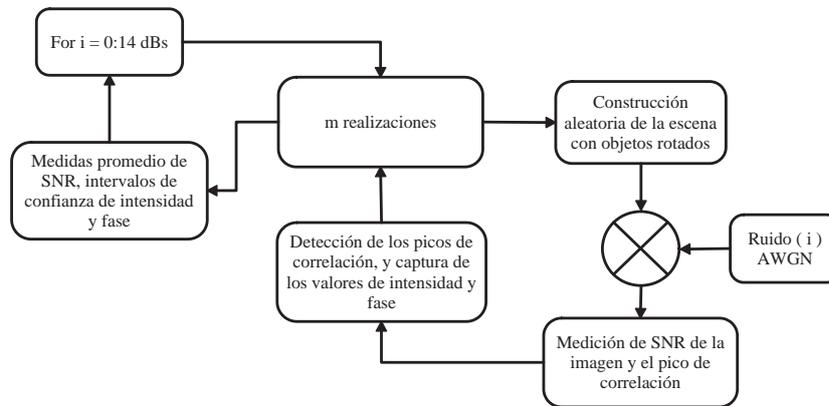


Figura 5.1: Simulación del desempeño del filtro propuesto bajo distintos niveles de ruido aditivo.

### 5.1.2. Algoritmo para medir el comportamiento de la fase ante imágenes geoméricamente distorsionadas por rotación

El filtro propuesto bajo el enfoque de la adaptación, también ha sido diseñado para mostrar tolerancia modesta a distorsiones geométricas como es el caso de imágenes del objetivo rotadas. Para medir el comportamiento del filtro ante imágenes rotadas se propone el algoritmo mostrado en la Fig. 5.2, en el cual se mide el nivel del pico de correlación de salida, para un rango en el cual, los objetivos aparecen rotados.

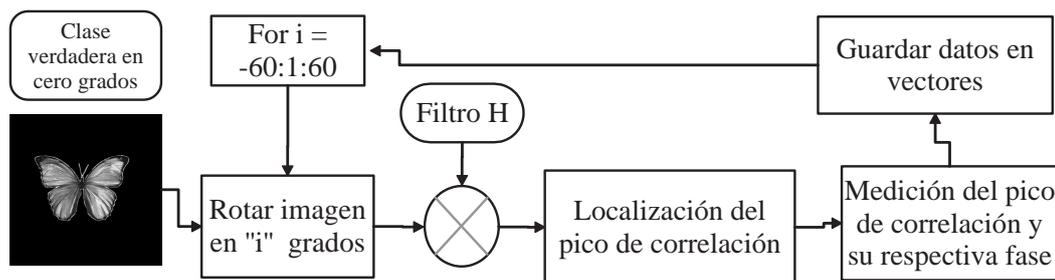


Figura 5.2: Algoritmo del desempeño del filtro ante distorsiones de rotación.

### 5.1.3. Algoritmo para la construcción de matriz de confusión

La discriminabilidad en un filtro multi-clasificador es de suma importancia, dado que el estudio de esta característica fundamenta la confiabilidad del filtro durante su operación. Las tablas de confusión son una medida estadística sobre el desempeño del clasificador. Estas tablas se componen por una relación entre los datos esperados que

debe de arrojar el clasificador (renglones), respecto a los resultados que realmente obtiene (columnas). La dimensión de la matriz de confusión dependerá del número de clases que se pretendan reconocer. En la Tabla 5.1 puede observarse que la matriz es de  $3 \times 3$ , lo que significa que el clasificador tiene la capacidad de reconocer tres clases.

Tabla 5.1: Ejemplo de matriz de confusión de tres clases.

		Clases asignadas		
		A	B	C
Clases reales	A	10	2	1
	B	0	6	1
	C	0	3	8

La fila de la primera de la matriz indica que hay 13 objetos que pertenecen a la clase A y que diez de ellos están correctamente clasificados como pertenecientes a la clase A. Observemos que dos de los objetos han sido confundidos con la clase B y C. La construcción de la matriz se puede ver en el algoritmo en donde las entradas corresponden a dos vectores que contiene los valores de clasificación del sistema, posteriormente en el algoritmo se lleva a cabo un conteo, en el cual se estiman los aciertos del clasificador, los datos confundidos con alguna de las clases de asignación o como dato perdido, en el caso de que no lo reconozca el clasificador.

---

**Algoritmo 5.1** Pseudocódigo para la construcción de la matriz de confusión.

---

**Input:** *Vector<sub>real</sub>*, *Vector<sub>obtenido</sub>*, *Numero<sub>clases</sub>*

**Output:** *Matriz<sub>confusion</sub>*

**for** tamaño(*Vector<sub>real</sub>*)

**for** tamaño(*Numero<sub>clases</sub>*)

        Comparar elementos (*Vector<sub>real</sub>*, *Vector<sub>obtenido</sub>*)

**if** elemento de *Vector<sub>real</sub>* = = *Vector<sub>obtenido</sub>*

                Contar elemento como reconocido

**if** elemento de *Vector<sub>real</sub>* = = A cualquier *valor de Numero<sub>clases</sub>*

                Contar elemento como confundido

**Otherwise**

                Contar elemento como perdido

**end**

**end**

**return** *Matriz<sub>confusion</sub>* ;

---

#### 5.1.4. Prueba de error de localización

En esta sección, los resultados de simulaciones por computadora son obtenidos con el enfoque propuesto para el reconocimiento y clasificación de objetos. Tales resultados son analizados y discutidos. Estos resultados son una comparación con bancos de filtros basados en los siguientes filtros compuestos: MACE [14], MACH [15], y DCCF [16]. El desempeño de los bancos de filtros es evaluado en términos del desempeño del reconocimiento, eficiencia de clasificación y complejidad computacional. El desempeño del reconocimiento está caracterizado por la capacidad de discriminación y error de localización (LE).

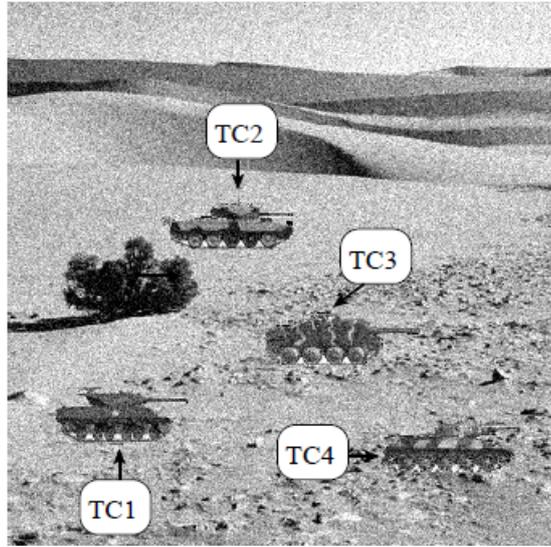


Figura 5.3: Ejemplo de la escena de entrada de prueba con ruido aditivo de varianza  $\sigma_n^2 = 4/256$ . TC1: Objeto de la clase 1; TC2: Objeto de la clase 2; TC3: Objeto de la clase 3; TC4: Objeto de la clase 4.

La capacidad de discriminación está dada por la Eq. (3.3.5), mientras que el error de localización está definido por:

$$LE = [(\tau_x - \hat{\tau}_x)^2 + (\tau_y - \hat{\tau}_y)^2]^2, \quad (5.1.1)$$

donde  $\tau_x, \tau_y$  y  $\hat{\tau}_x, \hat{\tau}_y$  son coordenadas exactas y estimadas del objeto, respectivamente.  $\tau_x, \tau_y$  se asume que son conocidas mientras que  $\hat{\tau}_x, \hat{\tau}_y$  son tomadas de la localización del pico de correlación. La eficiencia de clasificación está caracterizada por la mínima tasa de error de clasificación (ERC), y la complejidad computacional es el número de correlaciones requeridas en el proceso de reconocimiento y clasificación. En nuestros experimentos hemos utilizado cuatro diferentes clases de objetos. Estos son vehículos

tipo tanque que son mostrados en la Fig. 4.4. La escena de entrada está definida como

$$f(x, y) = \sum_{k=1}^K t_k(x - \tau_{xk}, y - \tau_{yk}) + \left[ 1 - \sum_{k=1}^K w_k(x - \tau_{xk}, y - \tau_{yk}) \right] b(x, y) + n(x, y), \quad (5.1.2)$$

donde  $t_k(x, y)$  representa un objeto de la  $k$ -ésima clase verdadera,  $\tau_{xk}, \tau_{yk}$  son variables aleatorias representando coordenadas desconocidas de la  $k$ -ésimo objeto en la escena,  $b(x, y)$  es el fondo,  $n(x, y)$  es la media cero del ruido aditivo con varianza  $\sigma_n^2$ , y  $w_k(x, y)$  es la region de soporte de  $t_k(x, y)$ .

## 5.2. RESULTADOS DE LA SÍNTESIS DE BANCOS DE FILTROS

La entrada de la escena puede ser interpretada como un conjunto de objetos embebidos en un fondo con coordenadas desconocidas, y corrompidas con ruido aditivo. Asumiremos que no ocurren oclusiones entre los objetos. En nuestros experimentos, usamos imágenes monocromáticas con  $400 \times 400$  pixeles. El rango de la señal es de  $[0,1]$  con 256 niveles de cuantización. El tamaño de los objetos es de aproximadamente  $95 \times 40$  pixeles, con un valor de media y desviación estándar de  $\mu_{t1} = .258$ ,  $\sigma_{t1} = 0.242$  para el objeto de interés 1,  $\mu_{t2} = 0.242$  para el objeto de interés 2,  $\mu_{t3} = 0.263$ ,  $\sigma_{t3} = 0.236$  para el objeto de interés 3, y  $\mu_{t4} = 0.272$ ,  $\sigma_{t4} = 0.252$  para el objeto de interés 4. La imagen de fondo tiene un valor medio de  $\mu_b = 0.68$  y una desviación estándar de  $\sigma_b = 0.18$ .

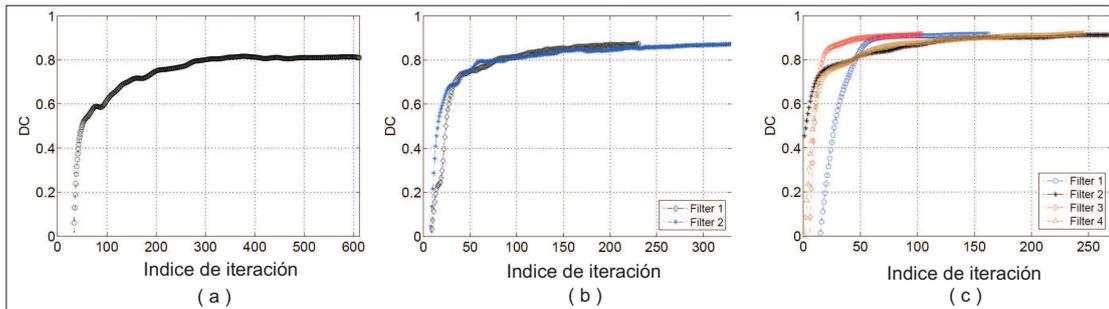


Figura 5.4: Desempeño del banco de filtros propuesto en el proceso de diseño: (a)ACSDF1, (b)ACSDF2, (c)ACSDF4

Un ejemplo de la escena de prueba en la entrada se muestra en la Fig. 5.3. Diseñamos filtros compuestos para el reconocimiento y clasificación de cuatro clases de objetos con distorsiones en rotación dentro del rango de  $[-8,8]$  grados. Notemos que los bancos de filtros MACE, MACH y DCC requieren un filtro compuesto por cada clase de objetos, mientras que un banco con filtros adaptativos puede ser construido con un número de

filtros de correlación menor que el número de clases. Diseñamos tres bancos de filtros basado en filtros adaptativos: ACSDF4, ACSDF2 y ACSDF1. El banco ACSDF4 usa cuatro filtros adaptativos y cada filtro está entrenado para reconocer solo una clase verdadera de objetos. El banco ACSDF2 consiste en dos filtros adaptativos y cada filtro está entrenado para reconocer dos clases verdaderas de objetos. El banco ACSDF1 consiste en un filtro adaptativo sintetizado para reconocer y clasificar las 4 clases. Notar que la complejidad computacional del banco filtros es de 4 correlaciones para el ACSDF4, MACE, MACH y DCC. Para el ACSDF2 son 2 correlaciones, y solo una correlación para el ACSDF1.

### 5.3. RESULTADOS DEL DESEMPEÑO DEL FILTRO ANTE LA PRESENCIA DE RUIDO ADITIVO Y RUIDO DISJUNTO

El conjunto de entrenamiento que representa a las  $k$ -ésimas clases de objetos esta dada por 9 imágenes de entrenamiento:  $\{S_k^T\} = \{s_i^k(x, y; \theta_i) | \theta_i = -8, -6, \dots, 0, 2, \dots, 8\}$ , donde  $s_i^k(x, y; \theta_i)$  es la  $i$ -ésima imagen de referencia de la  $k$ -ésima clase, rotadas por un ángulo  $\theta_i$ . De está manera, el filtro ACSDF1 contiene 36 imágenes de la clase verdadera, cada uno de los dos filtros del banco ACSDF2 contiene 18 imágenes de entrenamiento de la clase verdadera, y cada uno de los cuatro filtros del banco ACSDF4, MACE, MACH y DCC contiene 9 imágenes de la clase verdadera. Los filtros adaptativos propuestos son diseñados con ayuda de un algoritmo iterativo mostrado en la Fig. 4.2. Los valores de fase para las diferentes clases son asignadas como sigue:  $\phi_1 = \frac{\pi}{8}$  (22.5 grados) para la clase 1,  $\phi_2 = \frac{\pi}{4}$  (45 grados) para la clase 2,  $\phi_3 = \frac{3\pi}{8}$  (67.5 grados) para la clase 3 y  $\phi_4 = \frac{\pi}{2}$  (90 grados) para la clase 4.

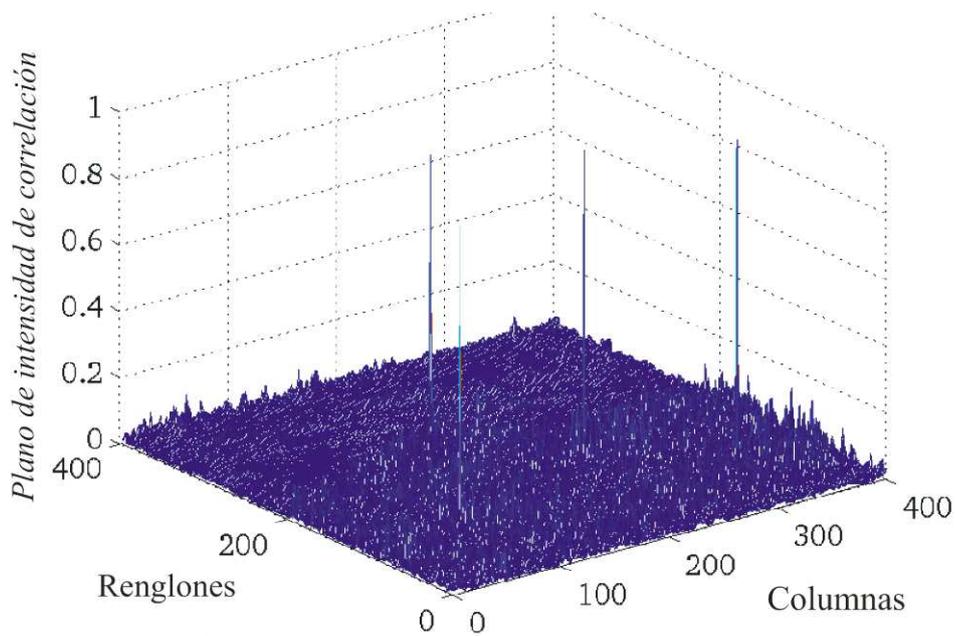


Figura 5.5: Plano de intensidad de la correlación obtenida con el filtro ACSDF1 para la escena de entrada de la Fig. 5.3.

En el diseño de los filtros adaptativos se utilizó un fondo que es diferente pero similar en términos estadísticos a la de los utilizados en los experimentos de reconocimiento. Antes de la primera iteración el valor de DC para todos los filtros es negativo. Después de 612 iteraciones en el proceso adaptativo el ACSDF1 alcanza un  $DC=0.80$ . El banco de filtros ACSDF2 alcanza una  $DC=0.85$  después de 350 iteraciones, y el banco de filtros ACSDF4 alcanza una  $DC=0.91$  después de 271 iteraciones. Esto significa que al usar el algoritmo propuesto, podemos lograr un alto nivel de control sobre el plano de correlación de la escena de entrada. La Fig. 5.4 muestra el desempeño del banco de filtros adaptativo en el proceso de diseño en términos de la DC versus el índice iterativo. Para ilustrar el desempeño del método propuesto, el plano de intensidad de salida obtenido con el filtro ACSDF1 para la escena de entrada en la Fig. 5.3, es presentada en la Fig. 5.5. Podemos ver 4 picos de correlación indicando la presencia de los objetos en las posiciones correctas. Uno puede observar, que los valores de intensidad de salida de correlación sobre el área del fondo son muy bajos. En la tabla 5.2 se muestra el desempeño del reconocimiento del filtro ACSDF1. Se puede observar que todos los objetivos son identificados con valores de DC cercanos a los obtenidos durante el proceso de entrenamiento. Por otra parte, los valores de fase  $\hat{\phi}_k$  contenidos en la ubicación de los picos de correlación son muy similares a los definidos en las restricciones del filtro. Esto significa que todos los objetivos pueden ser reconocidos y clasificados correctamente con una operación de correlación única. A continuación, ponemos a prueba el rendimiento de reconocimiento y clasificación de filtro ACSDF1 cuando los objetivos de las diferentes clases se rotan en un rango de  $[-20,20]$  grados.

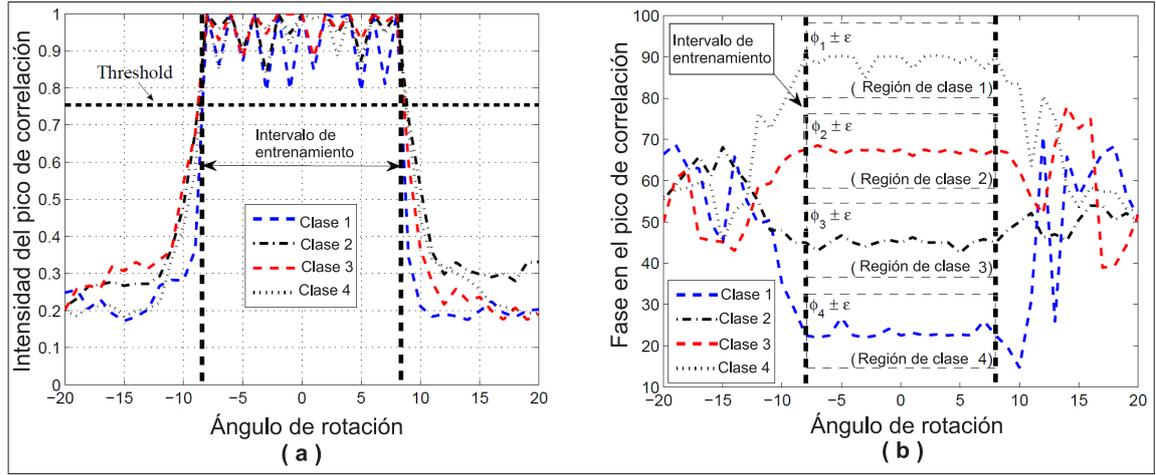


Figura 5.6: Desempeño del reconocimiento y clasificación del filtro ACSDF1 con versiones rotadas de objetos de todas las clases.

Los resultados se presentan en la Fig. 5.6. Vemos que en el interior del intervalo de entrenamiento, todos los objetivos rotados producen picos de intensidad por encima de 0.8. Por lo tanto, los objetivos pueden ser reconocidos con el valor de umbral de 0.8, para la instancia 0.75. De otra manera, la Fig. 5.6(b) muestra que en el intervalo de entrenamiento, los valores de fase de los picos de correlación de salida, jamás se traslapan, y estos son muy cercanos a aquellos definidos en las restricciones del filtro. Ahora, probaremos el desempeño del reconocimiento de todos los bancos de filtros considerados cuando las versiones rotadas de los objetos pertenecientes a diferentes clases están inmersas en el fondo en las coordenadas desconocidas, y cuando la varianza del ruido aditivo  $\sigma_n^2$  es modificada.

## 5.4. RESULTADOS DE LA CAPACIDAD DE DISCRIMINALIDAD DEL FILTRO PROPUESTO

En este experimento, los objetivos pueden aparecer rotados dentro de un rango de  $[-10,10]$  grados. Para garantizar la obtención de correctos resultados estadísticos, se realizaron 50 intentos estadísticos de cada experimento para diferentes posiciones y rotaciones de cada objetivo y realizaciones de procesos aleatorios son llevados a cabo. Con el 95% de confianza los resultados obtenidos en términos de la DC y LE son presentados en las Fig. 5.7 y 5.10, respectivamente. Adicionalmente, se calcularon los valores de fases estimadas  $\hat{\phi}_k$  en los picos de correlación para los bancos de filtros propuestos ACSDF4, ACSDF2 y ACSDF1 son calculados. Con el 95% de confianza estos resultados son presentados en la Fig. 5.8. Puede observarse que los bancos de

filtros propuestos ACSDF4 y ACSDF2 producen el mejor resultado en términos de la DC y LE.

Tabla 5.2: Desempeño del ACSDF1 en términos de la DC y de los valores de fase  $\phi^k$  estimados en el pico de correlación para la escena de entrada mostrada en la Fig.5.3.

	TC1	TC2	TC3	TC4
DC	0.79	0.79	0.80	0.81
$\hat{\phi}_k$	22.9°	45.8°	68.7°	89.3°

Con estos resultados puede ser visto que el filtro simple ACSDF1 se ejecuta un poco mejor en términos de DC que los bancos DCC y una mejora significativa en términos de LE que los bancos de filtros MACE, MACH, y DCC, mientras que su complejidad computacional es cuatro veces menor. Se debe tomar en cuenta que los bancos de filtros MACE y DCC tienen una mala tolerancia al ruido aditivo.

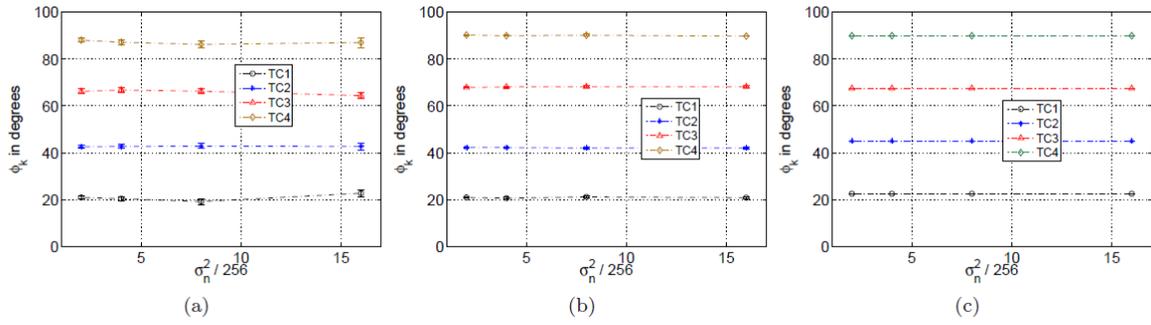


Figura 5.8: Valores de fase  $\phi_k$  estimados del pico de correlación con un intervalo de confianza del 95% mientras la varianza del ruido  $\sigma_n$  es cambiada. (a) ACSDF1, (b) ACSDF2, (c) ACSDF4.

Es importante mencionar que, dado que la complejidad computacional de los bancos propuestos puede ser variable, un equilibrio entre el rendimiento del reconocimiento y la complejidad computacional debe ser establecida. En la Fig. 5.8, se observa que los valores de fase obtenidos  $\hat{\phi}_k$  en los picos de correlación para los bancos ACSDF4, ACSDF2 y ACSDF1 producen un comportamiento estable en presencia de ruido aditivo y no hay traslapes entre las clases.

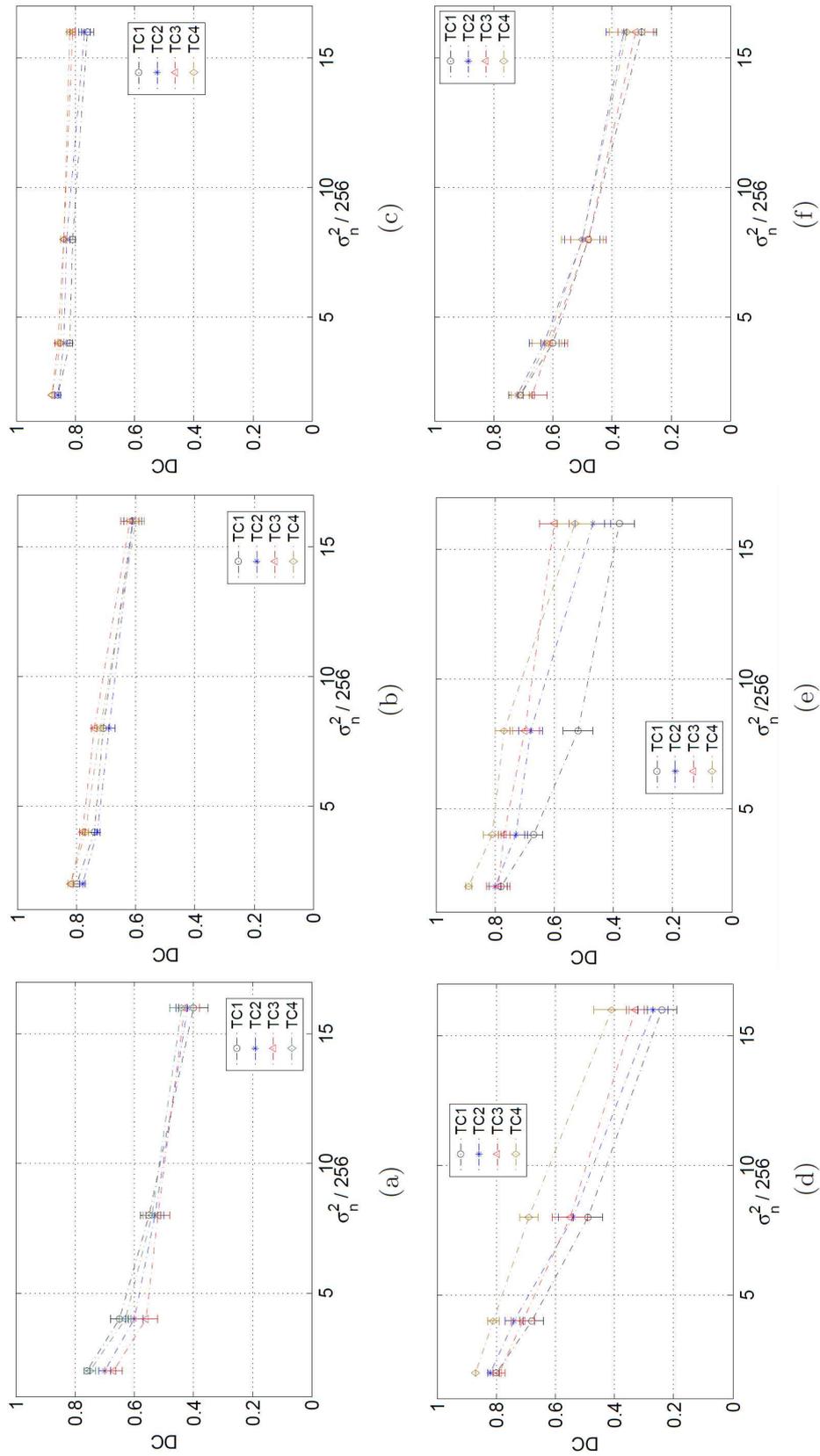


Figura 5.7: Desempeño de la DC con 95% de confianza para los diferentes bancos de filtros mientras la varianza del ruido  $\sigma_n$  es cambiada. (a) ACSDf1, (b) ACSDf2, (c) ACSDf4, (d) MACE, (e) MACH, (f) DCC.

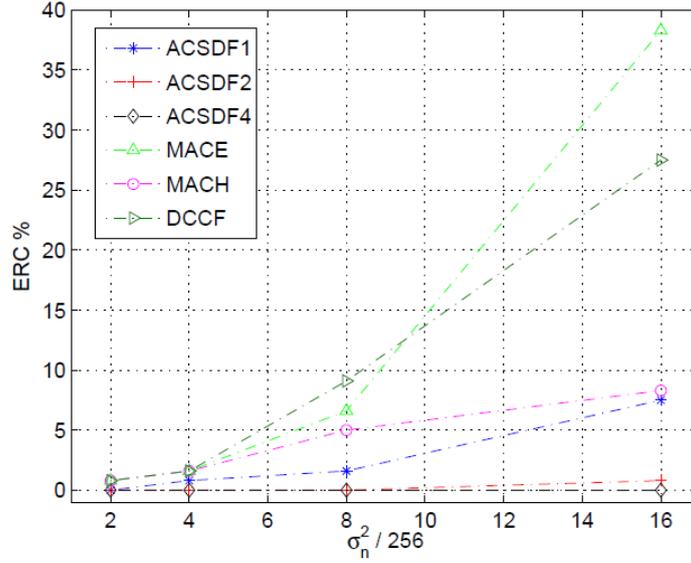


Figura 5.9: Desempeño del ERC con intervalo de confianza del 95 % para los diferentes bancos de filtros mientras la varianza del ruido  $\sigma_n$  se incrementa.

A continuación, ponemos a prueba el desempeño de la clasificación de los bancos de filtros. Para garantizar los correctos resultados estadísticos, 120 pruebas estadísticas de cada experimento para diferentes posiciones y rotaciones de los objetivos y las realizaciones de procesos aleatorios se han llevado a cabo. En cada ensayo, se escoge aleatoriamente una versión girada de un objetivo en el rango de  $[-10,10]$  grados de diferentes clases. Este objetivo está incorporado a un fondo en coordenadas desconocidas y la escena está contaminada con ruido aditivo.

La escena construida es correlacionada con cada banco de filtros y la intensidad del plano de salida con la más alta DC elegida. Si el LE es menor que la mitad del tamaño del objetivo (20 pixeles), entonces el objeto identificado es asignado a su correspondiente categoría, de otra manera, el objeto es rechazado. Los resultados son resumidos en las Tablas 5.3-5.8 y la Fig. 5.9. Las entradas en las Tablas 5.3-5.8 muestran el número de objetos de entradas de cada clase (denominadores) contra la decisiones hechas por los bancos de filtros (numeradores). Esto también muestra el porcentaje de la tasa de error de la clasificación (ERC) por cada experimento. EL ERC está definido como

$$ERC = \frac{\text{error de clasificación} + \text{rechazos}}{\text{intentos}} \times 100$$

Observe que el propósito de el banco ACSDF4 no produce errores en todas las pruebas. El banco ACSDF2 produce una alta ERC de 1.6 % solo cuando el ruido es fuerte,  $\sigma_n^2 = 16/256$ . Finalmente, notar que el banco de filtros MACE produce los peores valores de clasificación.

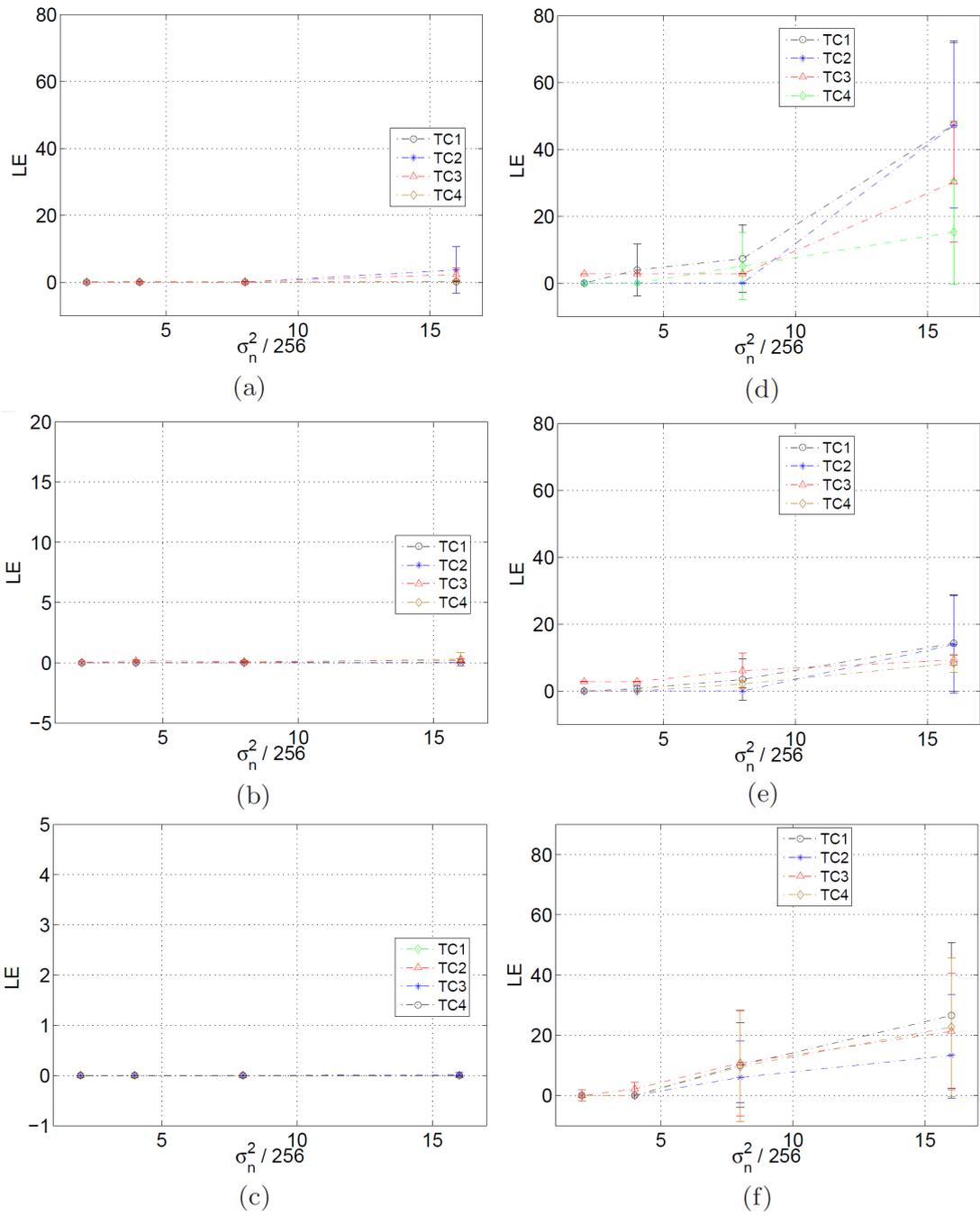


Figura 5.10: Desempeño de LE con intervalos de confianza del 95 % para los diferentes bancos de filtros mientras la varianza del ruido  $\sigma_n$  crece. (a) ACSDF1, (b) ACSDF2, (c) ACSDF4, (d) MACE, (e) MACH, (f) DCC.

Tabla 5.3: Desempeño de clasificación del ACSDF1.

Decisión	$\sigma_n^2 = 4$				$\sigma_n^2 = 8$				$\sigma_n^2 = 16$			
	Clases de entrada				Clases de entrada				Clases de entrada			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
C1	29/29	0	0	0	34/34	1	0	0	30/33	0	0	0
C2	0	28/28	1	0	0	28/29	1	0	1	28/28	1	0
C3	0	0	29/30	0	0	0	31/32	0	0	0	31/35	0
C4	0	0	0	33/33	0	0	0	25/25	0	0	1	22/24
Rechazado	0	0	0	0	0	0	0	0	0	0	0	0
<b>ERC</b>	<b>0.8 %</b>				<b>1.6 %</b>				<b>7.5 %</b>			

Tabla 5.4: Desempeño de clasificación del ACSDF2.

Decisión	$\sigma_n^2 = 4$				$\sigma_n^2 = 8$				$\sigma_n^2 = 16$			
	Clases de entrada				Clases de entrada				Clases de entrada			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
C1	35/35	0	0	0	30/30	0	0	0	30/31	0	0	0
C2	0	34/34	0	0	0	29/29	1	0	0	26/26	0	0
C3	0	0	26/26	0	0	0	32/33	0	0	0	28/29	0
C4	0	0	0	25/25	0	0	0	28/28	0	0	0	34/34
Rechazado	0	0	0	0	0	0	0	0	1	0	1	0
<b>ERC</b>	<b>0.0 %</b>				<b>.8 %</b>				<b>1.6 %</b>			

Tabla 5.5: Desempeño de clasificación del ACSDF4.

Decisión	$\sigma_n^2 = 4$				$\sigma_n^2 = 8$				$\sigma_n^2 = 16$			
	Clases de entrada				Clases de entrada				Clases de entrada			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
C1	23/23	0	0	0	24/24	0	0	0	24/24	0	0	0
C2	0	30/30	0	0	0	34/34	0	0	0	42/42	0	0
C3	0	0	46/46	0	0	0	46/46	0	0	0	31/31	0
C4	0	0	0	21/21	0	0	0	16/16	0	0	0	23/23
Rechazado	0	0	0	0	0	0	0	0	0	0	0	0
<b>ERC</b>	<b>0.0 %</b>				<b>0.0 %</b>				<b>0.0 %</b>			

Tabla 5.6: Desempeño de clasificación del MACE.

Decisión	$\sigma_n^2 = 4$				$\sigma_n^2 = 8$				$\sigma_n^2 = 16$			
	Clases de entrada				Clases de entrada				Clases de entrada			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
C1	21/21	0	0	0	16/19	0	0	0	10/17	0	0	0
C2	0	34/34	0	0	0	35/38	0	0	0	21/33	1	0
C3	0	0	43/45	0	1	0	42/44	0	0	0	25/46	0
C4	0	0	0	20/20	0	1	0	19/19	0	0	0	18/24
Rechazado	0	0	2	0	2	2	2	0	7	12	20	6
<b>ERC</b>	<b>1.6 %</b>				<b>6.6 %</b>				<b>38.3 %</b>			

Tabla 5.7: Desempeño de clasificación del MACH.

Decisión	$\sigma_n^2 = 4$				$\sigma_n^2 = 8$				$\sigma_n^2 = 16$			
	Clases de entrada				Clases de entrada				Clases de entrada			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
C1	23/23	0	0	0	18/22	0	0	0	9/11	0	0	0
C2	0	38/38	0	0	0	34/35	0	0	0	39/42	1	0
C3	0	0	44/45	0	1	0	43/44	0	0	0	48/51	0
C4	0	0	0	14/14	0	0	0	19/19	0	0	0	14/16
Rechazado	0	0	1	0	3	1	1	0	2	3	2	2
<b>ERC</b>	<b>0.8 %</b>				<b>5.0 %</b>				<b>8.3 %</b>			

Tabla 5.8: Desempeño de clasificación del DCCF.

Decisión	$\sigma_n^2 = 4$				$\sigma_n^2 = 8$				$\sigma_n^2 = 16$			
	Clases de entrada				Clases de entrada				Clases de entrada			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
C1	31/31	0	0	0	25/28	0	0	0	19/32	1	0	0
C2	0	33/33	0	0	0	23/25	0	0	0	22/30	0	0
C3	0	0	25/27	0	0	0	32/35	1	0	0	22/27	1
C4	0	0	1	29/29	0	0	0	29/32	0	0	0	24/31
Rechazado	0	0	1	0	3	2	3	2	13	7	5	6
<b>ERC</b>	<b>1.6 %</b>				<b>9.1 %</b>				<b>27.5 %</b>			

# 6

## CONCLUSIONES

En esta tesis, se ha presentado un método eficiente para un confiable reconocimiento y clasificación de objetos usando un banco de filtros adaptativos de correlación. Hemos comprobado que el uso de ambas distribuciones de intensidad y fase de la salida compleja del plano de correlación permiten reconocer y clasificar múltiples objetos pertenecientes a diferentes clases con una sola operación de correlación. Los filtros adaptativos propuestos, fueron construidos con la ayuda de un algoritmo iterativo que optimiza la capacidad de discriminación. Se ha comprobado, que a través del uso del método propuesto podemos establecer un compromiso entre la complejidad computacional del sistema en términos del número de correlaciones y el desempeño de la multclasificación en el reconocimiento de patrones.

Se realizaron numerosos experimentos en simulación por computadora que muestran que el enfoque propuesto puede producir resultados significativamente mejores en términos de capacidad de discriminación, errores de localización, y de la tasa de error de la clasificación en comparación a los obtenidos con los bancos de filtros construidos con filtros compuestos de correlación tradicionales. Por lo que, los experimentos demuestran que el uso de la fase como característica, adicional del plano de correlación, permite una mayor discriminación entre las clases. Por otro lado, el uso de adaptabilidad en los filtros de correlación, nos permite mejorar el desempeño en el reconocimiento de los objetos ante el ruido disjunto y el ruido aditivo. Además, se corrobora, que las técnicas de optimización permiten converger a los mejores resultados. En cuanto a

la síntesis del filtro se observó que el número aceptado de imágenes falsas que puede tomar en cuenta la propuesta es bastante alta (alrededor de 700 imágenes), sin embargo, muestra mejoría significativa hasta las 80 a 200 imágenes de entrenamiento de la clase falsa. Y que el número de imágenes verdaderas con que se puede construir el filtro propuesto es limitado (aproximadamente hasta 32 imágenes de entrenamiento). Por lo que no se puede prescindir de un banco de filtros, sin embargo, es posible aminorar su uso con la aproximación propuesta en comparación con los demás filtros.

## 6.1. APORTACIONES

Las aportaciones realizadas en la presente tesis son:

- El desarrollo de herramientas de software para la construcción y evaluación de los filtros, mencionados en la tesis.
- Un amplio estudio de la respuesta compleja de la correlación.
- El diseño de un filtro adaptativo complejo capaz de clasificar simultáneamente, varios objetos en una misma escena, con solo una operación de correlación.
- Publicaciones en congresos nacionales e internacionales.

## 6.2. TRABAJOS FUTUROS

Como parte de las investigaciones realizadas en la presente tesis, se derivan ciertas interrogantes que pueden ser tomadas como trabajo futuro, y que son necesarias responder para mejorar la aproximación propuesta. Entonces, basado en lo anterior se propone los siguientes temas:

- *Es posible modificar la propuesta, para operar con otros filtros analíticos?* Es necesario, realizar filtros compuestos con otros filtros analíticos, con el fin de evaluar de una manera más amplia el estudio de la fase, y así determinar cuál de los filtros existentes es el más adecuado para controlar el plano de correlación y converja más rápidamente en la síntesis.
- *Es posible realizar, una aproximación que no utilice restricciones para controlar la fase?* Uno de los problemas más comunes de los filtros de correlación restrictivos es que no pueden controlar de manera amplia el plano de correlación. Por lo tanto, sería deseable determinar si es posible desarrollar una aproximación que puede de cierta manera controlar más de un solo punto en el plano complejo,

tanto en la intensidad como en la fase. En resumen, realizar una aproximación que no restrictiva o semi-restrictiva para intensidad y fase.

- *Cual es el alcance de la aproximación propuesta en una implementación óptica?* Sin duda, una aplicación física puede llevar a resultados diferentes a los mostrados en la presente tesis, ya que posee diferentes limitaciones y ventajas en comparación con los sistemas digitales.

## REFERENCIAS BIBLIOGRAFICAS

- [1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification*. John Wiley and Sons, Inc., New York, 2001.
- [2] V. H. Diaz-Ramirez and V. Kober. Target recognition under nonuniform illumination conditions. *Applied Optics*, 48:1408–1418, 2009.
- [3] V. H. Diaz-Ramirez and V. Kober. Adaptive phase-input joint transform correlator. *Applied Optics*, 46(26):6543–6551, 2007.
- [4] M. S. Alam, A. Bal, El-Houssine Horache, S. Goh, C. H. Loo, S. P. Regula, and A. Sharma. Metrics for evaluating the performance of joint-transform-correlation-based target recognition and tracking algorithms. *Optical Engineering*, 44(6):067005, 2005.
- [5] V. Kober, Y. Seong, and T. Choi. Trade-off filters for optical pattern recognition with nonoverlapping target and scene noise. *Progress in pattern recognition, image analysis and applications*, 10(1):149–151, 2000.
- [6] R. A. Kerekes and B. V. K. Vijaya-Kumar. Correlation filters with controlled scale response. *IEEE Transactions on Image Processing*, 15(7):1794–1802, 2006.
- [7] B. V. K. Vijaya-Kumar, A. Mahalanobis, and A. Takessian. Optimal tradeoff circular harmonic function correlation filter methods providing in-plane rotation response. *Applied Optics*, 9(6):1025–1034, 2000.
- [8] P. Refregier. Bayesian theory for target location in noise with unknown spectral density. *Journal of the Optical Society of America*, 16:276–283, 1999.
- [9] S. M. A. Bhuiyan, M. S. Alam, and M. Alkanhal. New two-stage correlation-based approach for target detection and tracking in forward-looking infrared imagery using filters based on extended maximum average correlation height and polynomial distance classifier correlation. *Optical Engineering*, 46(8):086401, 2007.

- [10] S. M. A. Bhuiyan, M. S. Alam, and S. R. F. Sims. Target detection, classification, and tracking using a maximum average correlation height and polynomial distance classification correlation filter combination. *Optical Engineering*, 45(116401):xxx, 2006.
- [11] M. Alkanhal and V. B. K. Vijaya-Kumar. Polynomial distance classifier correlation filter for pattern recognition. *Applied Optics*, 42(23):4688–4708, 2003.
- [12] V. H. Diaz-Ramirez. Constrained composite filter for intraclass distortion invariant object recognition. *Optics and Lasers in Engineering*, 48(12):1153–1160, 2010.
- [13] R. A. Kerekes and B. V. K. Vijaya-Kumar. Selecting a composite correlation filter design: a survey and comparative study. *Optical Engineering*, 47(6):067202, 2008.
- [14] A. Mahalanobis, B. V. K. Vijaya-Kumar, and D. Casasent. Minimum average correlation energy filters. *Applied Optics*, 26:3633–3640, 1987.
- [15] A. V. Nevel and A. Mahalanobis. Comparative study of maximum average correlation height filter variants using ladar imagery. *Optical Engineering*, 42:541–550, 2003.
- [16] A. Mahalanobis, B. V. K. Vijaya-Kumar, and S. R. F. Sims. Distance-classifier correlation filters for multiclass target recognition. *Applied Optics*, 35(17):3127–3133, 1996.
- [17] V. H. Diaz-Ramirez, V. Kober, and J. Alvarez-Borrego. Pattern recognition with an adaptive joint transform correlator. *Applied Optics*, 45(23):5929–5941, 2006.
- [18] V. Kober J. A. Gonzalez-Fraga and J. Alvarez-Borrego. Adaptive synthetic discriminant function filters for pattern recognition. *Optical Engineering*, 45:0570051–05700510, 2006.
- [19] T. Acharya and Ajoy K. Ray. *Image processing principles and applications*, pages 7 – 11. Wiley - interscience, 2005.
- [20] T. Acharya and Ajoy K. Ray. *Image processing principles and applications*, pages 22 – 24. Wiley - interscience, 2005.
- [21] Real academia española. *Diccionario de la real academia española*. Espasa, 21 edition, 2008.
- [22] S. H. Nawab A. V. Oppenheim, A. S. Willsky. *Señales y sistemas*, pages 38 – 56. Prentice Hall, 2da. edition, 1998.

- [23] D.G. Manolakis J.G. Proakis. *Tratamiento Digital de Señales: Principios, Algoritmos y Ap.*, chapter 3, page 76. Prentice Hall, 2004.
- [24] L. Hassebrook B. V. K. Vijaya Kumar. *Correlation pattern recognition*, chapter 2, pages 59, 60. cambridge, 1era. edition, 2005.
- [25] Frank Y. Shih. *Image processing and pattern recognition fundamentals and techniques*, chapter 1, pages 5,8. John Wiley - IEEE Press, 2010.
- [26] Anil K. Jain. *Fundamentals of digital image processing*, chapter 1, page 13. Prentice Hall, 2da. edition, 1989.
- [27] S. H. Nawab A. V. Oppenheim, A. S. Willsky. *Señales y sistemas*, chapter 3, pages 179 – 202. Prentice Hall, 2da edición edition, 1998.
- [28] D. B. Williams V. K. Madisetti. *Digital signal processing handbook*, chapter 1, pages 29 – 31. CRC Press, 1999.
- [29] Douglas B. Williams V. K. Madisetti. *Digital signal processing handbook*, chapter 4,7, pages 144 –180. CRC Press, 1999.
- [30] C. Britton Rorabaugh. *Digital Filter designer's handbook featuring C routines*, chapter 2, page 53. McGraw-Hill, 1993.
- [31] Andreas Antoniou. *Digital signal processing: Signals, Systems and Filters*, chapter 16, pages 779 – 766. McGraw-Hill, 2006.
- [32] T. Acharya and Ajoy K. Ray. *Image processing principles and applications*, chapter 5, pages 79 – 82. Wiley interscience, 2005.
- [33] F. Gudail and Ph. Refregier. *Statistical image processing techniques for noisy images: An application - oriente approach*. KA/PP, 2004.
- [34] Jason Brownlee. *Clever Algorithms: Nature-Inspired Programming Recipes*, chapter 7, pages 169 – 174. Creative Commons, 2011.
- [35] Stanislaw H. Zak E. K.P. Chong. *An introduction to optimization*, chapter 6, page pag. 73. Wiley interscience, 2001.
- [36] Carlos A. Coello Coello. Introduccion a la computacion evolutiva (notas de curso). Technical report, CINESTAV, Mayo 2010. ccoello@cs.cinvestav.mx.
- [37] Stuar J Russell and N. Peter. *Artificial intelligence: A Modern Approach*, chapter 5, pages 121 –125. Prentice Hall, 2009.

- [38] Stanislaw H. Zak E. K.P. Chong. *An introduction to optimization*, chapter 6, page pag. 115. Wiley interscience, 2001.
- [39] Jason Brownlee. *Clever Algorithms: Nature-Inpired Programming Recipes*, chapter 5, page 73. Creative Commons, 2011.
- [40] O. Martin and S.W. Otto. Combining simulated annealing with local search heuristics. *Annals of Operations Research*, 63:1 – 17, 1996.
- [41] Runwei Cheng Mitsuo Gen. *Genetic Algorithms & Engineering optimization*, chapter 1, pages 1 – 40. John Wiley & Sons,, 2000.
- [42] Jason Brownlee. *Clever Algorithms: Nature-Inpired Programming Recipes*, chapter 5, page 39. Creative Commons, 2011.
- [43] J. H. Holland. *Concerning efficient adaptive systems*, chapter 7, pages 215, 230. Spartan books, 1ra edition, 1962.
- [44] J. H. Holland. Outline for a logical theory of adaptive systems. *Journal of the Association for Computing Machinery*, 9:297, 314, 1962.
- [45] J. W. Goodman. *Introduction to Fourier Optics*, chapter 6, pages 96 –120. McGraw-Hill, 1996.
- [46] T. H. Maiman. Simulated optical radiation in ruby. *Nature*, 187:493–494, 1960.
- [47] W. R. Bennet Javan, A. and D. R. Herrior. Population inversion and continuous optical maser oscillation in a gas discharge containing a he-ne mixture. *Physical Review Letters*, 6:106–110, 1961.
- [48] A. Vanderlugt. Signal detection by complex filtering. *IEEE Transactions on Information Theory*, IT-10:139–145, 1964.
- [49] C. S. Weaver and J. L. Goodman. Technique for optically convolving two functions. *Applied Optics*, 5:1248–1249, 1966.
- [50] P. E. Ross. Flash of genius. *Forbes*, 16:98 – 104, 1998.
- [51] Jianchang M. Anil K. Jain, Robert P. W. Duin. Statical pattern recognition: A review. *IEEE Transactions*, x:2, 1999.
- [52] Robert P.W. Duin A. K. Jain and Jianchang Mao. Statistical pattern recognition: A review. *Transactions on pattern analysis and machine intelligence*, 22:4 – 37, 2000.

- [53] G. Nagy. State of the art in pattern recognition. *Proceedings of the IEEE*, 56:836 – 357, 1968.
- [54] L. Hassebrook B. V. K. Vijaya Kumar. *Correlation pattern recognition*, chapter 1, pages 1 – 20. cambridge, 1era. edition, 2005.
- [55] S. Song S. R. F. Sims A. Mahalanobis, B. V. K. Vijaya Kumar and J. F. Epperson. Unconstrained correlation filters. *Applied Optics*, 33:3752 – 3755, 1993.
- [56] G. C. Runger Douglas C. Montgomery. *Applied Statistics and Probability for Engineers*, chapter 7, pages 226 – 227. John Wiley & sons, 2003.
- [57] L. Hassebrook B. V. K. Vijaya Kumar. *Correlation pattern recognition*, chapter 2, pages 149 – 150. cambridge, 1era. edition, 2005.
- [58] Ryan A. Kerekes. *Combining correlation outputs for Enhanced distortion-tolerant target detection*. PhD thesis, Carnegie Mellon Univ., May 2007. pags. 9 -10.
- [59] F. Gudail and Ph. Refregier. *Statical image processing techniques for noisy images: An application - oriente approach*. KA/PP, 2004.
- [60] B. V. K. Vijaya Kumar. Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, 31:4773 – 4801, 1992.
- [61] A. Diaz-Ramirez O. Castro, V. H. Diaz-Ramirez and Vitaly Kober. Improvement of pattern recognition with a heuristic design of correlation filters. *Proceedings of SPIE*, 7442:744217–(1–9), 2009.
- [62] B. V. K. Vijaya-Kumar. Tutorial survey of composite filter designs for optical correlators. *Applied Optics*, 31:4773–4801, 1992.
- [63] B. V. K. Vijaya-Kumar and L. Hassebrook. Performance measures for correlation filters. *Applied Optics*, 29:2997–3006, 1990.
- [64] C. F. Hester and D. Casasent. Multivariant technique for multiclass pattern recognition. *Applied Optics*, 19:1758–1761, 1980.
- [65] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein. *Introduction to Algorithms, Third edition*. The MIT Press, 2009.