

**UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA**



FACULTAD DE CIENCIAS

---

LICENCIATURA EN CIENCIAS COMPUTACIONALES

---

**RECONOCIMIENTO DE PATRONES POR CORRELACIÓN APLICADO A UN  
SISTEMA DE VISIÓN ARTIFICIAL PARA LA SUPERVISIÓN DEL ARMADO**

**DE UN TANGRAMA**

TESIS PROFESIONAL

QUE PARA OBTENER EL TÍTULO DE

LICENCIADO EN CIENCIAS COMPUTACIONALES

Presenta:

EVERARDO SANTIAGO RAMÍREZ

ENSENADA, BAJA CALIFORNIA, MÉXICO, MAYO DE 2010.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE CIENCIAS

RECONOCIMIENTO DE PATRONES POR CORRELACIÓN APLICADO A UN  
SISTEMA DE VISIÓN ARTIFICIAL PARA LA SUPERVISIÓN DEL ARMADO DE  
UN TANGRAMA

TESIS PROFESIONAL

QUE PRESENTA

EVERARDO SANTIAGO RAMÍREZ

APROBADO POR:



DR. JOSÉ ÁNGEL GONZÁLEZ FRAGA

PRESIDENTE DEL JURADO



M. C. MARÍA VICTORIA MEZA KUBO  
SECRETARIO



M. C. JOSÉ IGNACIO ASCENCIO LÓPEZ  
MIEMBRO VOCAL

**Resumen** de la tesis de **Everardo Santiago Ramírez**, presentada como requisito parcial para obtener el grado de LICENCIADO EN CIENCIAS COMPUTACIONALES. Ensenada, Baja California, México. Mayo del 2010.

**RECONOCIMIENTO DE PATRONES POR CORRELACIÓN APLICADO A UN SISTEMA DE VISIÓN ARTIFICIAL PARA LA SUPERVISIÓN DEL ARMADO DE UN TANGRAMA**

El gran avance tecnológico que hemos experimentado en los últimos años nos inspira a desarrollar nuevas tecnologías que acompañen al hombre en el quehacer de su vida diaria. Tecnologías capaces de realizar las mismas tareas que un ser humano, de una forma más eficiente y barata, o de apoyo en el mejoramiento de su calidad de vida. En este aspecto, el reconocimiento de patrones por correlación juega un papel importante en el desarrollo de sistemas de visión artificial.

En este trabajo de tesis se presenta el análisis de diferentes tipos de filtros de correlación, de los cuales se seleccionaron tres para evaluar su desempeño. El primer filtro que se evalúa es el filtro de acoplamiento. La segunda opción es el filtro compuesto SDF de sólo fase el cual requiere de un conjunto de entrenamiento que describa las distorsiones esperadas para la imagen de referencia, este filtro muestra resultados mejores que el filtro de acoplamiento al ofrecer invariancia a la iluminación y a pequeñas rotaciones. La tercera opción es el filtro compuesto no lineal k-law que al igual que la segunda opción también requiere de un conjunto de entrenamiento. Este filtro tiene un mejor desempeño en términos de la métrica "Capacidad de Discriminación" que las primeras dos opciones.

El filtro que mostró mejor desempeño en la etapa de evaluación se usó para desarrollar un sistema de visión artificial capaz de evaluar el progreso del armado de un tangrama tangible. Dicho sistema puede proveer soporte tecnológico a terapias de estimulación cognitiva.

Resumen aprobado por:



---

Dr. José Ángel González Fraga

**Palabras Clave:** reconocimiento de patrones, filtros de correlación, tangrama.

## Dedicatorias

Dedico este trabajo como símbolo de agradecimiento a mis padres, Marciano y Lucía Florencia, a todas mis hermanas y hermanos, a la señora Shirley por todo su apoyo y cariño.

Everardo Santiago Ramírez

Jehová es mi pastor, nada me faltará.  
En lugares de delicados pastos me hará descansar,  
junto a aguas de reposo me pastoreará.  
Confortará mi alma; me guiará por sendas de justicia  
por amor de su nombre.

Aunque ande en valle de sombra de muerte, no temeré mal alguno,  
porque tú estarás conmigo; tu vara y tu cayado me infundirán aliento.  
Aderezas mesa delante de mí, en presencia de mis angustiadores;  
Unges mi cabeza con aceite; mi copa está rebosando.

Ciertamente el bien y la misericordia me seguirán todos los días de mi vida,  
y en la casa de Jehová moraré por varios días.

Salmo 23.

## **Agradecimientos**

Agradezco primeramente a Dios por permitirme hacer lo que me gusta, por darme la vida y por darme unos padres y hermanos excepcionales.

Al Dr. José Ángel González Fraga, por aceptar ser mi director de tesis, por ser un gran maestro, por guiarme y apoyarme en mi vida estudiantil, muchas gracias.

A mis padres por su apoyo, por enseñarme a dirigirme ante la vida con dignidad, respeto y esmero. También a mis hermanas por su incondicional apoyo, gracias por estar conmigo siempre que las necesitaba. A mis hermanos por soportarme y acompañarme siempre. Esto es por y gracias a ustedes.

A la M. C. María Victoria Meza Kubo y al M. C. José Ignacio Ascencio López por aceptar ser mis sinodales y por todos sus comentarios realizadas durante la etapa de revisión de la tesis.

A Francisco E. Martínez Pérez por sus aportaciones y comentarios para la presentación de este trabajo. A mis amigos Cecilia y Omar Aparicio por sus comentarios y aportes a este trabajo.

## Contenido

Resumen .....	I
Dedicatorias.....	II
Agradecimientos .....	III
Contenido .....	IV
Lista de figuras.....	VIII
Lista de tablas.....	X
I. Introducción.....	1
I.1. Objetivos.....	3
I.1.1. Objetivo general.....	3
I.1.2. Objetivos específicos .....	3
I.2. Organización de la tesis .....	4
II. Antecedentes.....	5
II.1 Reconocimiento de patrones.....	6
II.2. Filtros de correlación en el reconocimiento automático de objetos .....	7
II.3. Evolución histórica de los filtros de correlación .....	9
II.4. Visión artificial .....	12
III. Fundamento matemático .....	15

III.1. Representación de imágenes digitales .....	15
III.2. La Transformada de Fourier.....	16
III.2.1. Propiedades de la transformada de Fourier .....	18
III.3. Sistemas lineales .....	20
III.3.1. Sistemas básicos .....	21
III.3.2. Sistemas lineales invariantes a desplazamientos.....	22
III.3.3. Sistemas LSI bidimensionales.....	22
III.4. Desarrollo de sistemas de visión artificial.....	23
III.4.1. Componentes de un sistema de visión artificial .....	23
III.4.2. Metodologías para el desarrollo de sistemas de visión artificial..	26
III.4.3. Herramientas para el desarrollo de aplicaciones para la visión artificial.....	30
IV. Filtros de correlación .....	33
IV.1. Filtros básicos .....	33
IV.1.1. Filtros de acoplamiento .....	34
IV.1.2. Filtros de sólo fase.....	36
IV.1.3. Filtros de sólo fase binarios.....	37
IV.2. Filtros avanzados.....	38
IV.2.1. Invariancia a distorsiones en el plano .....	39

IV.2.2. Funciones armónicas circulares .....	40
IV.2.3. Filtros compuestos de correlación.....	42
IV.2.4. Funciones sintéticas discriminantes (SDF) .....	42
IV.2.5. Filtro del Mínimo Promedio de la Energía de Correlación (MACE).....	45
IV.2.6 Filtro SDF de mínima varianza (MVSDF).....	47
IV.2.7. Filtros tolerantes a distorsiones .....	48
IV.2.8. Filtro compuesto no lineal k-law.....	49
IV.2.9. Filtros de correlación clasificadores de distancia.....	50
IV.2.10. Filtros polinomiales de correlación .....	52
IV.3. Medidas de desempeño de los filtros de correlación .....	53
V. Evaluación del desempeño de algunos filtros de correlación .....	55
V.1. Filtro de acoplamiento.....	56
V.2. Filtro compuesto SDF de sólo fase.....	60
V.3. Filtro compuesto no lineal k-law .....	66
VI. Aplicación del filtro compuesto no lineal k-law en la supervisión del armado de un tangrama.....	73
VI.1. Uso de tangrama como herramienta de estimulación cognitiva .....	73

VI.2 Escenario de uso de la visión artificial durante el desarrollo de las actividades de estimulación cognitiva.....	74
VI.3. Selección de valor para el factor de no linealidad k.....	75
VI.4. Sistema de visión artificial para la supervisión del armado de un tangrama .....	83
VII. Conclusiones y trabajo futuro.....	90
VII.1. Conclusiones.....	91
VII.2. Trabajo futuro .....	93
Referencias.....	94
A. Tablas comparativas de los filtros.....	98
B. Diseño del sistema de visión artificial.....	100
B.1. Descripción del hardware .....	102
B.2. Implementación del filtro de correlación.....	104

## Lista de figuras

Figura 1. Imagen vista como una función bidimensional. ....	16
Figura 2. Sistema LSI básico .....	21
Figura 3.- Componentes de un sistema de visión artificial .....	25
Figura 4. Diagrama de bloque mostrando el proceso completo de la visión artificial.....	29
Figura 5. Metodología para un sistema de visión artificial con aprendizaje máquina.....	30
Figura 6. Filtro lineal para el problema de detección de señales binarias.....	35
Figura 7.- Descomposición de la imagen en fase y magnitud: a) Imagen original, b) Fase de la imagen original, c) Fase de la imagen original pero con magnitud pseudo-aleatoria, d) la fase domina a la magnitud en sus efectos relativos para representar la información de la imagen .....	37
Figura 8. Conversión de una imagen bidimensional a su representación vectorial y su inclusión en la matriz que generará el filtro compuesto. ....	43
Figura 9. Operaciones necesarias para obtener el plano de correlación con el filtro de acoplamiento. ....	57
Figura 10.-Imagen de referencia .....	58
Figura 11.- Resultado producido por el filtro de acoplamiento .....	58
Figura 12.- Gráfica de resultados del filtro de acoplamiento.....	59
Figura 13. Ejemplo de conjunto de imágenes de entrenamiento. ....	61
Figura 14. Imagen compuesta formada por el conjunto de entrenamiento representado por la figura 13 .....	61

Figura 15. Operaciones necesarias para obtener el plano de correlación con el filtro compuesto SDF de sólo fase. ....	63
Figura 16.- Reconocimiento de objetos con el filtro SDF de sólo fase .....	64
Figura 17.- Resultados de aplicar el filtro SDF de sólo fase a configuraciones parciales del Tangrama.....	65
Figura 18. Operaciones necesarias para obtener el plano de correlación con el filtro compuesto no lineal k-law. ....	68
Figura 19.- Reconocimiento automático de objetos con el filtro compuesto no lineal SDF del tipo k-law.....	69
Figura 20.- Resultado de aplicar el filtro compuesto no lineal k-law con $k=0.1, \dots, 0.4$ . .	70
Figura 21.- Resultados de aplicar el filtro compuesto no lineal k-law con $k=0.5, \dots, 0.9$ .	71
Figura 22.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con $k=0.1$ .....	77
Figura 23.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con $k=0.2$ .....	78
Figura 24.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con $k=0.3$ .....	79
Figura 25.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con $k=0.4$ .....	80
Figura 26.- Desempeño del filtro compuesto no-lineal k-law ante configuraciones parciales de una figura del tangrama.....	81

Figura 27.- Desempeño del filtro compuesto no lineal k-law con un factor de no linealidad $k=0.3$ ante el filtro de sólo fase y el filtro de acoplamiento .....	82
Figura 28.- Imágenes de referencia a probar con SISVASAT .....	84
Figura 29.- Figura completa de tangrama en SISVIASAT .....	85
Figura 30.- Plano de correlación de una figura realizada en SISVIASAT con un DC = 0.983.....	86
Figura 31.- Resultado de una configuración parcial de 6 piezas del tangrama con SISVIASAT. ....	87
Figura 32.- Resultado de una configuración parcial 5 piezas del tangrama con SISVIASAT. ....	87
Figura 33.- Resultado al armar las figuras 25(b) y 25(c) en SISVIASAT.....	88
Figura 34.- Desempeño de SISVIASAT con el filtro compuesto no lineal k-law .....	89
Figura 35. Metodología para SISVIASAT.....	101
Figura 36.- Procesamiento de imágenes en SISVIASAT .....	102
Figura 37.-Vista esquemática de SISVIASAT.....	103
Figura 38.- Arquitectura MVC para SISVIASAT .....	105
Figura 39.- Diagrama de casos de usos para SISVIASAT.....	106
Figura 40.- Diagrama de flujo para SISVIASAT.....	107

## **Lista de tablas**

Tabla I.- Filtros básicos de correlación: criterios de maximización y desventajas.....	98
Tabla II.- Filtros avanzados de correlación: criterios de maximización y desventajas....	99

# Capítulo I

## Introducción

El reconocimiento de patrones es la disciplina científica que se encarga de la clasificación de objetos en clases o categorías de acuerdo a una propiedad específica. Uno de los enfoques del reconocimiento de patrones se basa en el uso de filtros de correlación para detectar cualquier similitud entre una señal de referencia y una señal de prueba. Este enfoque es conocido como reconocimiento de patrones por correlación, el cual cuenta con aplicaciones en áreas como el reconocimiento automático de objetos, reconocimiento biométrico, reconocimiento óptico de caracteres, entre otros.

Un problema importante en el enfoque de reconocimiento de patrones por correlación se presenta al momento de implementar los filtros en sistemas de reconocimiento automático de objetos, es decir, sistemas de visión artificial. Los problemas se presentan en estos sistemas cuando el objeto a reconocer presenta cambios de escala, rotación, desplazamiento y efectos causados por la iluminación. Actualmente, en la literatura se han propuesto muchos filtros de correlación que intentan resolver los problemas presentados, algunos de los cuales hacen uso de un conjunto de imágenes de entrenamiento, utilizan la fase de la transformada de Fourier de la imagen, se componen de filtros no lineales y muchas otras características con el fin de ofrecer invariancia a los cambios presentados por los objetos.

En este trabajo de tesis se presenta una descripción de los diferentes tipos de filtros de correlación, de los cuales se seleccionaron tres para evaluar su desempeño. El primer filtro que se evalúa es el filtro de acoplamiento, posteriormente el filtro compuesto SDF de sólo fase el cual requiere de un conjunto de entrenamiento que describa las distorsiones esperadas para la imagen de referencia y por último el filtro compuesto no lineal k-law que también requiere de un conjunto de imágenes de entrenamiento, pero además la aplicación de un factor de no linealidad tanto a la transformada de Fourier de la imagen de referencia como al de la imagen de prueba.

Con el fin de proveer soporte tecnológico a terapias de estimulación cognitiva y probar el funcionamiento de los filtros de correlación en un ambiente real, se desarrolló un sistema de visión artificial para evaluar el progreso del armado de un tangrama tangible, entiéndase por tangrama tangible aquél que puede ser manipulado físicamente.

El objetivo de la estimulación cognitiva consiste en mejorar el funcionamiento cognitivo de las personas, tales como memoria, lenguaje, atención, concentración, razonamiento, abstracción y operaciones aritméticas. La estimulación cognitiva es una terapia de apoyo en el envejecimiento y complementaria al tratamiento farmacológico de la demencia.

Si estimulamos estas capacidades de una forma tal que el individuo lo tome como un juego o una actividad de recreación, por ejemplo la resolución de un tangrama, lograremos un mayor avance que si usamos medicamentos. Esto es posible debido a que el individuo al armar el tangrama aprovecha las capacidades que aún posee,

manteniéndose ocupado, estimulando sus habilidades cognitivas y mejorando su conducta social (Rodríguez, 2009).

## **I.1. Objetivos**

### **I.1.1. Objetivo general**

Mostrar la funcionalidad del reconocimiento automático de objetos mediante filtros de correlación, tanto filtros lineales como filtros no lineales con invariancia a la rotación, traslación y desplazamiento para la supervisión del armado de un Tangrama tangible.

### **I.1.2. Objetivos específicos**

1. Analizar los fundamentos matemáticos del reconocimiento de patrones por correlación y sistemas de visión artificial.
2. Analizar y comparar los diferentes filtros de correlación para el reconocimiento de objeto invariante a la rotación, desplazamiento e iluminación.
3. Diseñar un método para un sistema de visión artificial.
4. Usar un filtro de correlación invariante a desplazamiento y el método diseñado en el punto tres de los objetivos específicos en un sistema de visión artificial para supervisión automática del armado de un tangrama tangible.

## **I.2. Organización de la tesis**

La presente tesis se divide en siete capítulos. En el capítulo I se presenta una breve descripción acerca del tema tratado en este documento. En el capítulo II se presenta una descripción de trabajos que se han reportado en la literatura sobre el reconocimiento de patrones por correlación y la visión artificial. El capítulo III contiene los fundamentos teóricos, es decir, la información matemática necesaria para adentrarse al mundo del reconocimiento de patrones por correlación. En el capítulo IV, se presentan los filtros básicos y avanzados de correlación, así como las medidas de desempeño para dichos filtros. El capítulo V describe la evaluación de tres filtros de correlación aplicados a la problemática del reconocimiento del tangrama. En el capítulo VI se presenta la aplicación del filtro compuesto no lineal k-law en la supervisión del armado de un tangrama tangible. Por último, en la sección VII se discuten las conclusiones a las que se llegaron con la realización de este proyecto y el trabajo futuro.

## **Capítulo II**

### **Antecedentes**

Existe un gran desarrollo en el área de reconocimiento de patrones por correlación en los últimos años, no sólo por su gran capacidad de reconocimiento de objetos o personas, sino también por su capacidad de recibir como entrada una escena completa y encontrar la ubicación de los objetos dentro de la misma, así como el grado de similitud entre los objetos de la escena con el objeto de referencia.

El reconocimiento de patrones forma parte de muchos sistemas de visión artificial. Actualmente, se aplica en la industria para la automatización de procesos; tales como control de calidad y manufactura, logrando con ello un aumento significativo en la producción y una mejoría en la calidad de productos. Otra área importante en el cual la visión artificial está teniendo una gran aceptación es el cómputo ubicuo, en donde se usa para inferir el contexto de un ambiente mediante el seguimiento de objetos, detección de cambios, entre otros.

En el área de seguridad también se utiliza el reconocimiento de patrones, pues ofrece la capacidad de detectar eventos específicos y tomar decisiones en base a ellos. El uso del reconocimiento de patrones en un sistema de visión artificial hace posible una optimización de recursos. Por ejemplo, de qué sirve que en una ciudad se instalen cientos de cámaras de vigilancia si nadie las ve o simplemente no es posible monitorear todas. Porque no mejor instalar un sistema de visión artificial que detecte ciertos eventos y emitir alguna alarma.

En este capítulo se presenta una breve mención de trabajos que se han reportado en la literatura especializada acerca del reconocimiento de patrones, filtros de correlación y su evolución histórica, y por último, trabajos relacionados con la visión artificial.

## **II.1 Reconocimiento de patrones**

El reconocimiento de patrones es la disciplina científica cuyo objetivo es la clasificación de objetos en un cierto número de categorías o clases. Dependiendo de la aplicación, esos objetos pueden ser imágenes, ondas de señales o cualquier tipo de señales que necesitan ser clasificadas.

Históricamente los dos enfoques en el reconocimiento de patrones han sido el estadístico (o teoría de la decisión) y el sintáctico (o estructural). Recientemente, el desarrollo de las redes neuronales ha propiciado un nuevo enfoque. El aumento en la capacidad de almacenamiento ofrecida por los más modernos sistemas ha propiciado el reconocimiento de patrones basado en la apariencia (Pajares et al. 2008). De estos enfoques, el reconocimiento de patrones por correlación, el cual es el tema que aborda este documento, pertenece al enfoque de reconocimiento de patrones basado en la apariencia, pero debido a su base matemática y a que también hace uso del elemento estadístico para encontrar el valor máximo de correlación, también pertenece al enfoque de reconocimientos estadísticos.

Entre los varios marcos de trabajo en el cual el reconocimiento de patrones ha sido formulado, el enfoque estadístico ha sido estudiado intensivamente y usado en la

práctica. Recientemente, las técnicas de redes neuronales y métodos importados de teoría de aprendizaje estadístico han recibido una gran atención por parte de los investigadores. El diseño de un sistema de reconocimiento requiere una atención cuidadosa para los siguientes aspectos: definición de clases de patrones, reconocimiento de entorno, representación de patrones, extracción y selección de características, análisis, diseño del clasificador, selección de ejemplares de entrenamiento, evaluación de desempeño (Jain et al. 2000). Además, se debe de considerar la calidad de la lente del sensor de visión y el objetivo del sistema, ya que un mal diseño conduce a un fallo inesperado del sistema.

En (Feng et al. 2008) presentan un método de reconocimiento de patrones tridimensionales tolerante a la rotación usando exposiciones de holografías digitales. Las imágenes fueron reconstruidas de siete hologramas digitales usando el método de convolución. Esas imágenes fueron usadas para construir un filtro no lineal. El resultado del experimento mostró que un objeto verdadero puede ser diferenciado de un objeto falso por el valor del pico de correlación y que es factible usarlo en el reconocimiento de patrones tridimensionales tolerante a la rotación.

## **II.2. Filtros de correlación en el reconocimiento automático de objetos**

La correlación es una técnica robusta para el reconocimiento de patrones y es usado en muchas aplicaciones computacionales, tales como reconocimiento automático de objetos, reconocimiento biométrico, seguridad nacional y antiterrorismo, seguimiento

de armas guiadas, detección/reconocimiento de fonemas del habla para su conversión a texto, radares, aplicaciones médicas y reconocimiento óptico de caracteres (Vijaya Kumar et al. 2006). El diseño, análisis y uso de algoritmos de reconocimiento de patrones por correlación requiere de teoría de sistemas lineales, procesos y variables aleatorios, métodos matriz/vector, teoría de detección y estimación, procesamiento digital de señales y procesamiento óptico.

El reconocimiento de patrones por correlación es un subconjunto del reconocimiento de patrones con enfoque estadístico, está basado en la selección o creación de una señal de referencia para luego determinar el grado de similitud entre la señal que se está examinando y la señal de referencia. El grado de similitud es una simple estadística sobre el cual se basa la decisión sobre el objeto (Vijaya Kumar et al. 2005).

El primer filtro de correlación lo presentó (Vanderlugth, 1964). Estos filtros pueden usarse para identificación de personas mediante datos biométricos (rostro, huellas, iris, etc.) y utilizarse en aplicaciones computacionales de seguridad; tales como el control de acceso a edificios importantes y a computadoras. En los métodos de filtros de correlación, las imágenes de entrenamiento pueden ser acomodadas en un arreglo en el dominio de la frecuencia (llamado filtro de correlación). El reconocimiento de objetos se realiza por correlación cruzada entre una imagen de prueba con una imagen de referencia usando la transformada rápida de Fourier.

Aunque existen muchos tipos de filtros de correlación, incluyendo el Filtro del Mínimo Promedio de la Energía de Correlación (MACE), Función Discriminante Sintética (SDF), Filtros de Equilibrio Óptimo (OTF), y MACE sin restricciones (UMACE), esos filtros son generalmente mejores en reconocimiento de objetos que en detección de objetos (Bolme et al. 2009).

En (Shaotong, 2008) usan un filtro lineal de correlación SDF para el reconocimiento y reconstrucción de objetos en tercera dimensión invariante a la rotación. El mismo filtro se usa en (Gudmundsson, 2007) para el reconocimiento de objetos con ruido aleatorio.

En (Walvoord et al. 2008) utilizan filtros de correlación para la transcripción digital del Palimpsesto de Arquímedes. Ellos utilizan técnicas de imagen multi-espectral y métodos de reconocimiento óptico de caracteres para digitalizar los datos en un marco de trabajo de análisis y clasificación.

### **II.3. Evolución histórica de los filtros de correlación**

En (Caulfield et al.1969) sugirieron que la discriminación entre pares similares de objetos puede mejorarse procesando las salidas de correlación. Ellos sugirieron usar combinaciones lineales de las salidas. Por ejemplo, podemos encontrar el peso de esas combinaciones lineales de tal forma que sólo una combinación resultará uno y es a la que corresponde al objeto que recibimos como entrada, el resto de las otras combinaciones resultarán cero. Este algoritmo asegura que una entrada que no corresponda al objeto a clasificar produce una salida de correlación cero con todos y

cada uno de los filtros. Sin embargo, requiere la ejecución de  $N$  correlaciones para reconocer una imagen de entrada de un conjunto de  $N$  imágenes de entrenamiento. (Breuneker et al. 1979) sugirieron que este proceso es redundante y que solo necesitamos ejecutar  $K$  correlaciones, donde  $k = \log_2 N$ .

(Hester et al. 1980) sugieren que este proceso de reconocimiento puede ser simplificado usando una combinación lineal de imágenes de referencia para crear una imagen compuesta, para posteriormente realizar la correlación cruzada entre la imagen de entrada con la imagen compuesta. Los pesos de la combinación lineal son seleccionados de tal manera que la salida de correlación en el origen es el mismo para todas las imágenes pertenecientes a una clase. De este enfoque surgió el término Función Sintética Discriminante (SDF, Synthetic Discriminant Function). El enfoque SDF requiere solo una correlación cruzada por imagen de entrada. Los primeros filtros SDF's no eran adecuados porque no consideraban entradas con ruido, es decir, no fueron optimizados para ser tolerantes a ruido.

En (Leger et al. 1982) propusieron un método de reconocimiento y clasificación usando un conjunto de funciones patrón que son diseñados en base a una distorsión esperada.

(Vijaya Kumar, 1986) introdujo los filtros SDF de mínima varianza, el cual maximiza la tolerancia a ruidos de los SDF's. El filtro SDF original tenía dificultad para localizar imágenes de referencia cambiadas por factores desconocidos en la imagen de entrada. Para resolver este problema, (Mahalanobis et al. 1987) introdujeron los Filtros

del Mínimo Promedio de la Energía de Correlación (MACE) capaces de producir picos de correlación bien formados en el origen.

(Kallman, 1987) propuso que dado un conjunto de imágenes de entrenamiento, se define una figura base sobre la más baja correlación cruzada con patrones de la clase deseada o verdadera y la correlación más alta con patrones de la clase no deseada o falsa.

En (Jared et al. 1989) sugirieron un algoritmo que se conoce como Algoritmo de Relajación, el cual incluye explícitamente restricciones de modulación (por ejemplo, Filtro de Solo Fase y Filtro Binario de Solo Fase) dentro del proceso de diseño.

Considere el caso especial en el cual la distorsión es la rotación en el plano y se tienen  $N$  imágenes de entrenamiento para representar esta distorsión. Suponga que esas imágenes fueron obtenidas rotando la imagen objetivo por  $0, 2\pi/N, 4\pi/N, \dots, 2\pi(N-1)/N$  radianes, respectivamente. Este conjunto de imágenes tiene una estructura de correlación única. Es fácil verificar que la matriz  $N \times N$  que almacena el producto interno de varios pares de las imágenes de arriba es del tipo Toeplitz, es decir, todas las entradas a lo largo de la diagonal son los mismos. Esta propiedad especial fue usada por (Hassebrook et al. 1990) para construir el Filtro Compuesto de Coeficiente de Fase Lineal (LPCCF), el cual se obtiene de la combinación lineal de  $N$  imágenes de entrenamiento con coeficientes complejos que tienen una magnitud unitaria y fase lineal.

En (Pérez et al. 2002) proponen un sistema de correlación que realiza varias correlaciones entre una escena de entrada y diferentes imágenes de referencia. Combinan un banco de filtros compuestos de correlación no lineales para diseñar un sistema de reconocimiento de señales de tránsito. El banco de filtros permite la tolerancia a cambios en escala y tolerancia a cierto grado de rotación en el plano. La síntesis de los filtros compuestos de correlación no lineales permite la tolerancia a la rotación fuera del plano. El sistema se probó con varias imágenes reales, los cuales incluyen diferentes tipos de distorsiones. Los resultados del sistema muestran la tolerancia a variaciones de iluminación y objetos parcialmente ocultos.

En (González et al. 2006) propusieron un filtro SDF adaptativo de correlación basado en una función discriminante sintética convencional para reconocimiento de objetos con un fondo incrustado. La información sobre el objeto a rechazar, objetos falsos y fondo a rechazar es utilizada en un proceso iterativo de entrenamiento para diseñar un filtro de correlación con un valor de capacidad de discriminación dada. El algoritmo en cada iteración suprime el pico de correlación más alto, correspondiente al fondo o a los objetos no deseados incrementado la capacidad de discriminación de manera monótona hasta llegar a un valor determinado.

## **II.4. Visión artificial**

La visión artificial por computadora es la capacidad que poseen las máquinas para ver el mundo que les rodea, más precisamente para deducir la estructura y las propiedades del mundo tridimensional a partir de una o más imágenes bidimensionales.

En la visión artificial la escena tridimensional es vista por una, dos o más cámaras para producir imágenes monocromáticas o en color. Las imágenes adquiridas pueden ser segmentadas para obtener de ellas características de interés tales como bordes o regiones. Posteriormente de las características se obtienen las propiedades subyacentes mediante el correspondiente proceso de descripción. Tras lo cual se consigue la estructura de la escena tridimensional requerida por la aplicación de interés (Pajares et al. 2008).

En (Song, 2001) se menciona que la arquitectura de un sistema de inspección visual está compuesto de: una cámara que captura una imagen en un punto clave del proceso, la cual es enviada a una etapa de pre-procesamiento. En la etapa de pre-procesamiento la imagen es segmentada, normalizada, suavizada, escalada y en muchos casos convertida a una forma estandarizada. La siguiente etapa es la extracción de características, las cuales serán clasificadas en una etapa posterior. Una vez clasificadas las características sigue la etapa de toma de decisiones para aceptar o rechazar dicha imagen. Es preciso reconocer que hoy por hoy la visión por computadora a veces no es la mejor solución a un problema. Existen muchas ocasiones en las que el problema es tan complejo que la solución humana es lo mejor. Pero a veces, las soluciones humanas tienden a ser inexactas o subjetivas, en ocasiones lentas y presentan una ausencia de rigor así como una pobre percepción. No obstante, la solución humana es menos estructurada que la solución artificial y muchos problemas de visión por computadora requieren un nivel de inteligencia mucho mayor que el de la máquina pueda ofrecer. El sistema de visión humano puede describir automáticamente una textura en detalle, un

borde, un color, una representación bidimensional de una tridimensional, ya que puede distinguir entre imágenes de diferentes personas, firmas, colores, además puede vigilar ciertas zonas, diagnosticar enfermedades a partir de radiografías, etc. Sin embargo, aunque algunas de estas tareas pueden llevarse a cabo mediante la visión artificial, el software o el hardware necesario no consigue los resultados deseables (Pajares et al. 2008).

No obstante, cabe diferenciar entre las aplicaciones donde la visión artificial constituye una herramienta por sí sola y aquellas otras en las que es parte de un sistema multi-sensorial. El primer caso engloba todas aquellas aplicaciones en las que el único sensor presente es el de visión. En el segundo caso nos estamos refiriendo a la navegación en robótica donde la visión constituye una capacidad sensorial más para la percepción del entorno que rodea al robot (Pajares et al. 2008).

## Capítulo III

### Fundamento matemático

Este capítulo trata sobre la base teórica/matemática del reconocimiento de patrones por correlación y la visión artificial. Primero se describe la representación de imágenes digitales y posteriormente se hace descripción detallada de la Transformada de Fourier, la cual sirve como base para la construcción de los filtros de correlación, además de su uso en la detección de bordes, restauración de imágenes, comprensión de imágenes, entre otros.

También se aborda el tema de los sistemas lineales de correlación, el cual involucra dos imágenes, los cuales son la imagen de referencia y la imagen de prueba las cuales son correlacionadas para producir una salida de correlación. Finalizamos este capítulo con los requerimientos y componentes de un sistema de visión artificial.

#### III.1. Representación de imágenes digitales

El término imagen se refiere a una función de intensidad bidimensional  $f(x, y)$ , donde  $x$  e  $y$  son las coordenadas espaciales y el valor de  $f$  en cualquier punto  $(x, y)$  es proporcional a la intensidad o nivel de gris de la imagen en ese punto (Véase figura 2). Una imagen se representa digitalmente en forma de matriz con una resolución espacial de  $M \times N$  elementos. Cada elemento de la matriz denominada píxel (picture element, por sus siglas en inglés) tendrá un valor asignado, que se corresponde con el nivel de

luminosidad del punto correspondiente en la escena captada por un dispositivo de captura.

Con esta disposición se puede considerar que las imágenes digitales presentan una serie de propiedades tanto métricas como topológicas que requieren su consideración para el procesamiento posterior.

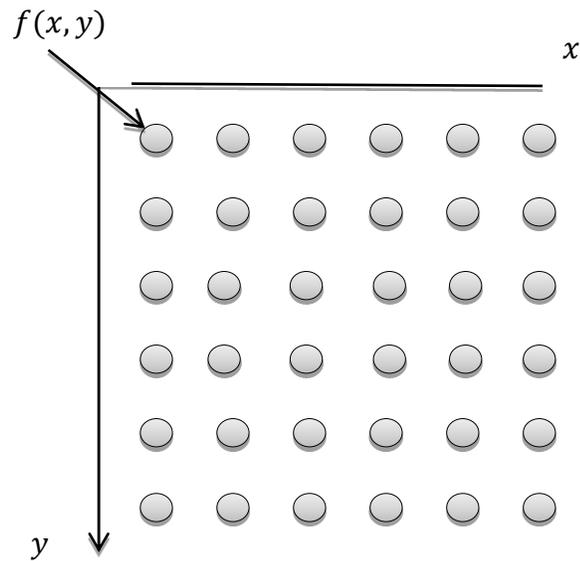


Figura 1. Imagen vista como una función bidimensional.

### III.2. La Transformada de Fourier

La transformada de Fourier  $F(u, v)$  de la imagen  $f(x, y)$  está definida por

$$F(u, v) = \mathcal{F}\{f(x, y)\} = \iint_{-\infty}^{\infty} f(x, y) \exp(-2\pi i(ux + vy)) dx dy \quad (1)$$

De manera análoga, la transformada inversa de Fourier  $f(x, y)$  de  $F(u, v)$  es:

$$f(x, y) = \mathcal{F}^{-1}\{F(u, v)\} = \frac{1}{4\pi^2} \iint_{-\infty}^{\infty} F(u, v) \exp(2\pi i(ux + vy)) \, du \, dv \quad (2)$$

Se puede ver la ecuación (2) como una expansión de la función imagen  $f(x, y)$  en forma de suma generalizada de exponenciales complejos. Para cada par de valores de las frecuencias espaciales  $u$  y  $v$  se tiene una exponencial en la suma generalizada, dicha exponencial está multiplicada por el coeficiente de peso  $F(u, v)$ , estos pesos son obtenidos a través de la ecuación (1). Por lo tanto la transformada de Fourier de  $f(x, y)$  puede verse como los coeficientes de peso de la expansión de la función de intensidad  $f$  en una suma de exponenciales.

Considere las consecuencias derivadas del hecho de que la transformada de Fourier de una imagen tiene un valor de magnitud alto en alguna frecuencia espacial dada. Supongamos por tanto que tenemos una imagen  $f(x, y)$  cuya transformada  $F(u, v)$  tiene una magnitud alta, por ejemplo en  $(u, v)$ . Puesto que  $F(u, v)$  tiene un valor alto, el término correspondiente  $F(u, v) \exp(2\pi i(ux + vy))$  contribuirá de forma importante a la suma generalizada de las exponenciales de la ecuación (2).

El análisis de Fourier nos dice que se puede construir una función imagen más compleja incluyendo más términos en la suma generalizada de exponenciales, considerando más pesos específicos de la forma  $F(u, v)$  y, por consiguiente, incluyendo más componentes de frecuencia espacial, donde el valor exacto de dichos pesos se obtiene a través de (1), es decir, de la transformada de Fourier.

$F(u, v)$  contiene componentes complejos, con la parte real correspondiendo a los términos del coseno y la parte imaginaria a los términos del seno. Se puede representar una componente espectral compleja por sus correspondientes partes, real e imaginaria:

$F(u, v) = Re(u, v) + i Im(u, v)$ , y entonces la magnitud es,

$$|F(u, v)| = \sqrt{[Re(u, v)]^2 + [Im(u, v)]^2} \quad (3)$$

Y la fase es,

$$\phi(u, v) = \tan^{-1} \left[ \frac{Im(u, v)}{Re(u, v)} \right] \quad (4)$$

### III.2.1. Propiedades de la transformada de Fourier

#### Separabilidad

Sea,

$$\begin{aligned} F(u, v) &= \frac{1}{N} \sum_{x=0}^{N-1} \exp\left(-\frac{i2\pi ux}{N}\right) \sum_{y=0}^{N-1} f(x, y) \exp\left(-\frac{i2\pi vy}{N}\right) \\ &= \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) g(x, y, u, v) \end{aligned} \quad (5)$$

Para  $u, v = 0, 1, 2, \dots, N-1$  y,

$$\begin{aligned} f(x, y) &= \frac{1}{N} \sum_{u=0}^{N-1} \exp\left(-\frac{i2\pi ux}{N}\right) \sum_{v=0}^{N-1} F(u, v) \exp\left(-\frac{i2\pi vy}{N}\right) \\ &= \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} F(u, v) h(x, y, u, v) \end{aligned} \quad (6)$$

Para  $x, y = 0, 1, 2, \dots, N-1$ .

En las ecuaciones (5) y (6), los términos  $g(x, y, u, v)$  y  $h(x, y, u, v)$  se denominan los núcleos de transformación sobre un punto directo e inverso respectivamente. El núcleo se dice que es separable si  $g(x, y, u, v) = g_1(x, u)g_2(y, v)$ , además, se dice que es simétrico si  $g_1$  es funcionalmente igual a  $g_2$ .

### Traslación

Las propiedades de traslación del par de transformadas de Fourier son,

$$f(x, y)\exp(2\pi i(u_0x + v_0y)/N) \Leftrightarrow F(u - u_0, v - v_0) \quad (7)$$

$$f(x - x_0, y - y_0) \Leftrightarrow F(u, v)\exp(-2\pi i(ux_0 + y_0)/N)$$

La ecuación (7) muestra que multiplicando  $f(x, y)$  por el término exponencial indicado y tomando la transformada del producto, resulta una traslación del origen del plano de la frecuencia al punto  $(u_0, v_0)$ . Del mismo modo, multiplicando  $F(u, v)$  por el término exponencial correspondiente y tomando la transformada inversa se mueve el origen del plano espacial a  $(x_0, y_0)$ .

### Periodicidad y simetría conjugada

La transformada de Fourier y su inversa son periódicas con periodo  $N$ , esto es

$$F(u, v) = F(u + N, v) = F(u, v + N) = F(u + N, v + N) \quad (8)$$

Como en el caso de las imágenes  $f(x, y)$  es real y por lo tanto la transformada de Fourier también exhibe simetría conjugada

$$F(u, v) = F^*(-u, -v) \text{ o bien } |F(u, v)| = |F(-u, -v)| \quad (9)$$

donde  $F^*(u, v)$  es la compleja conjugada de  $F(u, v)$ .

La propiedad de periodicidad indica que  $F(u, v)$  tiene un periodo de  $N$  y la propiedad de simetría muestra que la transformada se centra en el origen.

### Rotación

Si introducimos coordenadas polares tendremos

$$x = r \cos \theta, y = r \sin \theta, u = w \cos \phi, v = w \sin \phi$$

donde  $f(x, y)$  y  $F(u, v)$  se convierten en  $f(r, \theta)$  y  $F(w, \phi)$  respectivamente. Realizando la sustitución directa en el caso continuo o discreto produce,

$$f(r, \theta + \theta_0) \Leftrightarrow F(w, \theta + \theta_0) \quad (10)$$

Esta propiedad dice que si rotamos  $f(x, y)$  por un ángulo  $\theta_0$  entonces se rota  $F(u, v)$  por el mismo ángulo. Similarmente, rotando  $F(u, v)$  se rota  $f(x, y)$  por el mismo ángulo.

### III.3. Sistemas lineales

La correlación involucra dos señales o imágenes, en el caso de imágenes las dos imágenes involucradas son la imagen de referencia y la imagen de prueba. Una imagen de referencia es correlacionada con una imagen de prueba, conocida también como escena, para detectar y localizar la imagen de referencia dentro de la escena. De esta forma, un sistema de correlación está compuesto de una escena de entrada, una plantilla

o filtro almacenado (imagen de referencia) y una salida que es el resultado de la correlación.

Dicho sistema es lineal en el sentido de que una nueva entrada que es la suma de todos los pesos de la entrada original produce una salida que es idéntica a la suma total de los pesos de la salida original. Una propiedad importante de los sistemas lineales es su linealidad y la invariancia al tiempo, los cuales pueden ser caracterizados en términos de sus respuestas de frecuencia.

### III.3.1. Sistemas básicos

Un sistema puede ser caracterizado por producir una señal de salida  $o(x)$  en respuesta a una señal de entrada  $i(x)$ , tal y como se muestra en la figura 2.

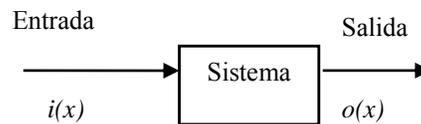


Figura 2. Sistema LSI básico

Para señales de dimensiones superiores como una imagen, el cual puede ser pensado como una señal 2-D, la señal de entrada para el sistema es  $i(x, y)$  y la salida es  $o(x, y)$ .

Una sub-clase importante de sistemas, conocidos como sistemas lineales invariantes a desplazamientos (LSI, linear shift-invariant) pueden ser caracterizados por

producir una salida solo para una entrada en particular, conocido como punto de entrada en el origen. La salida resultante se conoce como función de punto de dispersión (PSF, point spread function) en sistemas 2-D y como respuesta al impulso en sistemas 1-D.

### III.3.2. Sistemas lineales invariantes a desplazamientos

Una propiedad de un sistema LSI es que su salida se puede obtener al realizar la operación de convulsión entre la señal de entrada y su respuesta de impulso. Para que un sistema sea lineal requiere que la suma de sus pesos de la entrada pueda llevar a una señal de salida con peso idéntico. Un sistema lineal debe de satisfacer lo siguiente:

Si  $i_1(x) \rightarrow o_1(x)$  y  $i_2(x) \rightarrow o_2(x)$ , entonces  $ai_1(x) + bi_2(x) \rightarrow ao_1(x) + bo_2(x)$  para algún escalar  $a$ ,  $b$  y alguna entrada  $i_1(x)$  e  $i_2(x)$ .

Para un sistema invariante a desplazamientos, la señal de entrada desplazada por  $x_0$  produce una salida desplazada por la misma cantidad. La invariancia a desplazamientos nos dice que si conocemos la salida para una entrada en particular, entonces conocemos la salida para cada versión desplazada de una señal de entrada.

### III.3.3. Sistemas LSI bidimensionales

Consideremos  $h[n, m]$  que representa la señal de salida 2-D en respuesta a la señal de entrada  $\delta[n, m]$ . La salida  $o[n, m]$  de un sistema LSI para alguna entrada arbitraria  $i[n, m]$  se puede obtener mediante la siguiente operación de convulsión:

$$o[n, m] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} i[k, l] \cdot h[n - k, m - l] = i[n, m] * h[n, m] \quad (11)$$

donde \* indica la convolución 2-D.

De forma similar, la operación de correlación puede ser representada como sigue:

$$c[n, m] = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} i[k, l] \cdot h[k + n, l + m] = i[n, m] \otimes r[n, m] \quad (12)$$

donde  $\otimes$  indica la correlación 2-D. Cuando una señal es correlacionada consigo misma, la salida resultante es conocida como función de auto-correlación, el cual siempre produce un pico en el origen.

### **III.4. Desarrollo de sistemas de visión artificial**

La visión artificial o también conocido como visión por computadora es la capacidad que poseen las máquinas para ver el mundo que les rodea. Al ser un área muy amplia es considerada como una disciplina científica cuyo objetivo es crear teoría para la construcción de sistemas artificiales que obtienen información de imágenes. Dichas imágenes se pueden obtener de muchas formas, ya sea de una secuencia de video, una imagen fotográfica, o una imagen médica obtenida por un escáner.

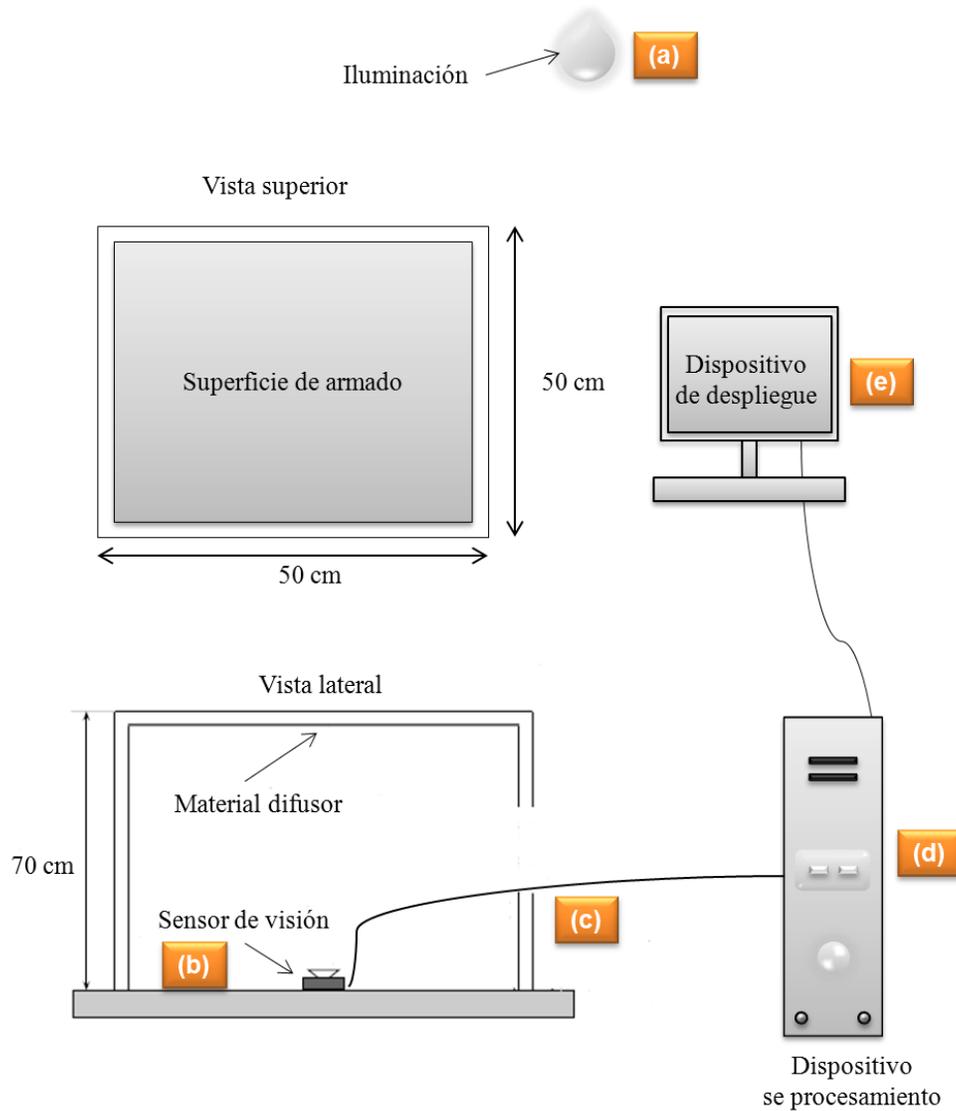
#### **III.4.1. Componentes de un sistema de visión artificial**

Actualmente se pueden distinguir dos tipos de visión artificial: visión artificial industrial y visión artificial académico. La diferencia entre ellos radica en la forma de trabajar, mientras que la visión artificial industrial obtiene imágenes del mundo real en tiempo real la visión artificial académica crea sus propias imágenes o las obtiene de una

base de datos. Independientemente de cómo y cuándo las obtienen, el objetivo de ambos enfoques es el mismo: obtener información de imágenes para dar soluciones a problemas reales.

Un sistema de visión artificial tiene varios componentes que trabajan en conjunto para resolver un problema. Los componentes de un sistema de visión artificial son descritos a continuación (ver figura 3).

- a) ***Un sistema de iluminación.*** Para un sistema de visión artificial en general es necesario un buen sistema de iluminación para la captura de imágenes de donde se obtendrá información. Si se utiliza luz ambiental, esta tiene variaciones muy grandes en cortos lapsos y son necesarios esfuerzos computacionales mayores para superar estas variaciones.
- b) ***Dispositivos de captura de imágenes.*** Para la adquisición de imágenes de un sistema de visión artificial, se necesita de dos elementos básicos. El primero es un dispositivo físico que sea sensible a una determinada banda del espectro de energía electromagnético y que produzca una señal eléctrica de salida proporcional al nivel de energía incidente en cualquier instante de tiempo. El segundo, denominado digitalizador, es un dispositivo para convertir una señal de salida continua del dispositivo físico (por ejemplo una salida eléctrica) en un conjunto discreto de localizaciones del plano de la imagen y, después en la cuantización de dicha muestra.



**Figura 3.- Componentes de un sistema de visión artificial**

c) *Un dispositivo de interfaz para transferir las imágenes a la computadora.*

Dicho dispositivo puede ser un cable USB conectado al dispositivo de captura y a la computadora, o por algún dispositivo que permite enviar las imágenes de forma inalámbrica.

- d) *Un dispositivo procesador de imágenes, una computadora o una cámara inteligente.* El procesamiento implica la manipulación de imágenes vistas como señales digitales para extraer la información más importante. Este mismo dispositivo se encarga de determinar ciertas estructuras elementales tales como bordes o regiones así como las relaciones entre ellas. Una vez hecho el procesamiento y el análisis y, en base a los resultados tratar de dar solución a ciertas situaciones del mundo real, tales como reconocimiento, movimiento y reconstrucción 3D.
- e) *Una interfaz para notificar el resultado del análisis a un operador.* El resultado se puede notificar de muchas formas, se puede hacer de forma gráfica, en forma de sonido o mediante cualquier otra señal electrónica.

### **III.4.2. Metodologías para el desarrollo de sistemas de visión artificial**

Es difícil definir una metodología estándar para los sistemas de visión artificial, pues este depende de la aplicación o problema que se quiere resolver. Pero a grandes rasgos, una metodología que propone la literatura especializada es la que presenta la figura 4, de la cual se describe cada paso a continuación:

*Adquisición de imágenes.* Para la adquisición de imágenes se requiere de una, dos o más cámaras que obtendrán imágenes del mundo 3D.

**Segmentación.** La segmentación es el proceso por el cual se extrae de la imagen cierta información subyacente para su uso posterior. La segmentación está basada en dos principios fundamentales: discontinuidad y similitud. La segmentación enfocada a la discontinuidad se orienta a la detección de bordes. Un borde es una línea que separa a dos regiones. La segmentación orientada a similitudes se orienta a la detección de regiones. Una región es un área de la imagen en la que sus píxeles poseen propiedades similares, ya sea de intensidad o color.

Con el avance tecnológico actual, se han propuesto varios operadores para la detección de bordes, siendo los más comunes los que se listan a continuación:

- Operadores primera derivada.
- Operadores segunda derivada.
- Operadores morfológicos.

Para la detección de regiones se proponen las siguientes técnicas:

- Binarización basada en el uso de umbrales.
- Crecimiento de regiones mediante la adición de píxeles.
- División de regiones.
- Similitud de textura, color o nivel de gris.

En un sistema de visión artificial, por lo general las imágenes no son de buena calidad, ya sea por efectos de la cámara o por el ambiente. Para solucionar este problema, se realiza la operación de suavizado y realzado. El suavizado nos permite

suprimir aquellos ruidos introducidos durante la captura de la imagen. El realzado se encarga de eliminar falsos reflejos y sombra que dificultan la extracción de la información para el procesamiento de la imagen.

**Descripción.** Una vez detectado los bordes o regiones, como elementos de interés, la descripción consiste en extraer propiedades o atributos para su uso en las aplicaciones. Básicamente consiste en reconocer las diferentes estructuras de la imagen para su identificación de forma inequívoca.

**Aplicaciones.** Existen cuatro áreas principales de aplicación de la visión artificial: obtención de la distancia de los objetos en la escena tridimensional y estructura tridimensional, detección de objetos en movimiento, reconocimiento de patrones y formas, y reconocimiento de objetos tridimensionales. Algunos objetivos finales en cualquiera de las áreas donde se aplique la visión artificial son: reconocer, manipular, predecir y navegar.

La metodología representada por la figura 4 es un enfoque muy general, otra metodología para desarrollar sistemas de visión artificial de aplicación industrial, el cual usa aprendizaje máquina es representado por la figura 5.

El primer paso en esta metodología es la adquisición de imágenes del mundo real, los cuales reciben un pre-procesamiento en donde las imágenes son normalizadas y estandarizadas de tal forma que facilite la extracción de características. En la etapa de extracción de características se eliminan aquellas que sean irrelevantes y redundantes de la imagen para generar las imágenes de entrenamiento que también pueden ser obtenidas

directamente de las características extraídas sin eliminar nada. Finalmente, el clasificador se basa en el conjunto de imágenes de entrenamiento para clasificar imágenes de entrada que no correspondan al conjunto de entrenamiento.

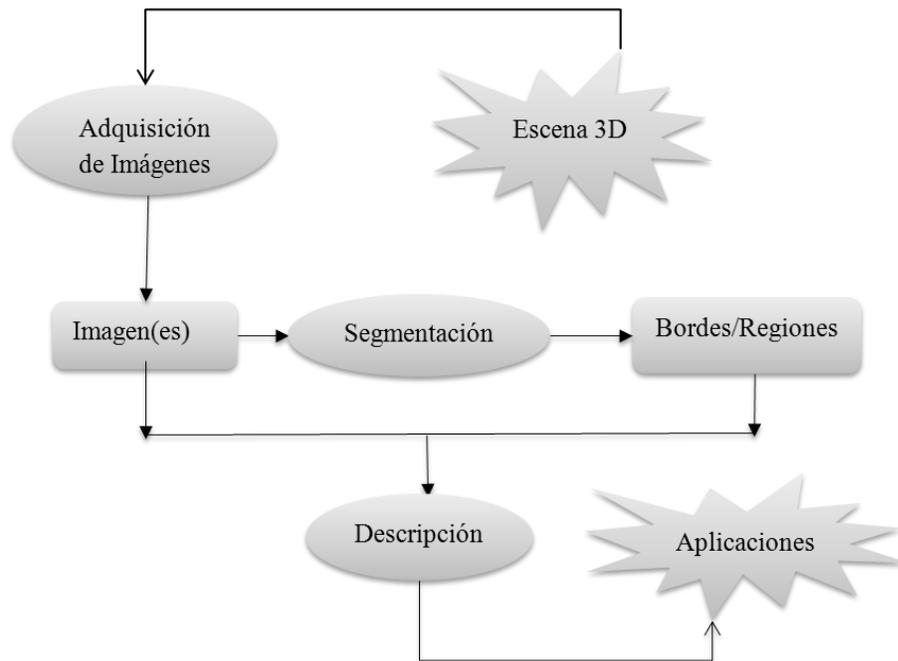


Figura 4. Diagrama de bloque mostrando el proceso completo de la visión artificial

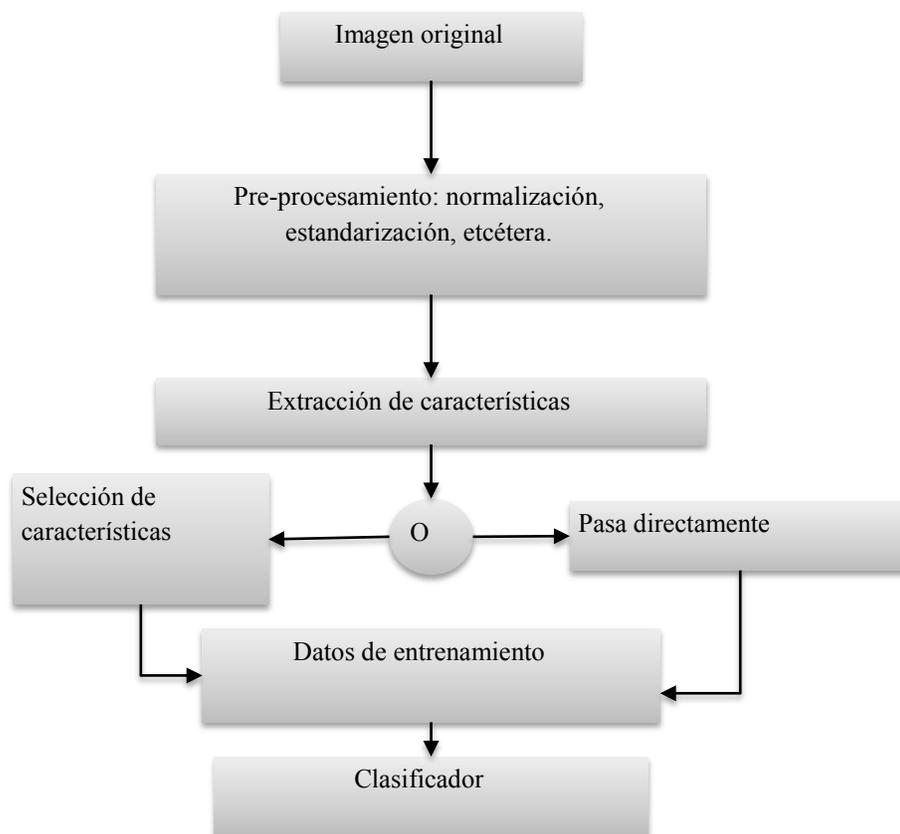


Figura 5. Metodología para un sistema de visión artificial con aprendizaje máquina.

### III.4.3. Herramientas para el desarrollo de aplicaciones para la visión artificial

Actualmente existen muchas herramientas y paquetes de software que permiten desarrollar aplicaciones de visión artificial. Muchas de estas aplicaciones se usan en la industria que es el área donde mayormente se utiliza la visión artificial.

A continuación se describe de manera muy breve sólo tres de las herramientas más usadas, herramientas que han probado ser de gran apoyo en el desarrollo de aplicaciones de visión por computadora.

IMAQ<sup>1</sup> de Vision National Instrument, es un paquete de software orientado a ingenieros y científicos para el desarrollo de aplicación de visión artificial. Ese paquete incluye módulos para desarrollo de aplicaciones de visión artificial, prueba de sistemas ya hecho y la posibilidad de aprender o conocer el funcionamiento los sistemas de visión artificial.

Matlab<sup>2</sup> es un paquete de software de propósito general pero que incluye módulos útiles en el desarrollo de aplicaciones para visión por computadora, tal es el caso de Image Acquisition Toolbox que permite la adquisición de imágenes a través de muchos dispositivos de captura. También se incluye el Image Processing Toolbox que incluye varias funciones para el procesamiento de imágenes.

OpenCV<sup>3</sup> aunque no es una herramienta en sí, sino más bien es una biblioteca de visión artificial. Desde su aparición es utilizada en muchas aplicaciones. Desde sistemas de seguridad con detección de movimiento, hasta aplicaciones de control de procesos donde se requiere el reconocimiento de objetos. OpenCV es multiplataforma lo cual ha apoyado a su gran difusión, aparte ser una librería libre. Contiene más de 500 funciones que abarcan una gran gama de áreas en el proceso de visión, como reconocimiento de

---

<sup>1</sup> [www.ni.com/vision/software.html](http://www.ni.com/vision/software.html)

<sup>2</sup> [www.mathworks.com](http://www.mathworks.com)

<sup>3</sup> <http://opencv.willowrage.com/wiki/>

objetos-ejemplo, reconocimiento facial-, calibración de cámaras, visión estéreo y visión robótica.

Independientemente de la herramienta que se use, las imágenes se presentan al sistema bajo diferentes formatos, entre los más comunes están: TIFF, BMP, GIF, JPG, PNG, DICOM, RAW.

En este capítulo se ha presentado la base teórica de los filtros de correlación y los sistemas de visión artificial. En el siguiente capítulo se describirán los principales filtros de correlación los cuales se han clasificados en básicos y avanzados.

## **Capítulo IV**

### **Filtros de correlación**

En la literatura especializada podremos encontrar una gran cantidad de filtros de correlación, los cuales optimizan diferentes criterios tales como la iluminación, el ruido, desplazamiento y rotación, entre otros. En esta sección mencionamos algunos de los filtros de correlación más conocidos ya sea por su robustez o por ser los pioneros en el área, también se menciona un par de métricas de desempeño de filtros de correlación.

#### **IV.1. Filtros básicos**

Suponga que queremos localizar todas las veces que aparece un avión en una fotografía del cielo. Una forma de lograr esto es mediante la correlación entre la imagen de referencia, en nuestro caso un avión, y una escena de entrada, la fotografía del cielo. La imagen de referencia se coloca en la esquina superior izquierda de la escena de entrada y se multiplican pixel por pixel. Los valores resultantes de esa multiplicación se suman para producir un valor de correlación. Este proceso se repite varias veces desplazando la imagen de referencia hacia la derecha y hacia abajo, esta operación producirá un arreglo bidimensional que se conoce como salida de correlación. Esta salida de correlación contendrá valores grandes que correspondan a aviones en la escena de entrada y cero para cualquier otra cosa. Un valor de correlación alto indica la presencia y localización de un avión en la escena.

Uno de los pioneros en el desarrollo de la correlación para el reconocimiento de patrones fue (VanderLugt, 1964) quien introdujo los filtros de acoplamiento. Este trabajo permitió usar correlacionadores ópticos para detectar y localizar imágenes de referencia en una escena.

#### **IV.1.1. Filtros de acoplamiento**

Los filtros de acoplamiento son óptimos en el sentido de que optimizan la razón señal a ruido (SNR) en la detección de una señal de referencia conocida y afectada por ruido aditivo blanco. Un defecto de este tipo de filtros es que son demasiados sensitivos a pequeños cambios en la señal o imagen de referencia, poca eficiencia en luz y relativamente baja capacidad de discriminación.

Considere a  $s(t)$  como la señal transmitida y  $r(t)$  como la señal recibida. La señal recibida puede obtenerse de entre dos hipótesis  $H_0$  y  $H_1$ . La hipótesis  $H_0$  establece que la señal recibida solo contiene ruido, el cual es representado por  $n(t)$ . La hipótesis  $H_1$  establece que la señal recibida es la suma de señal transmitida con ruido aditivo.

$$H_0: r(t) = n(t) \tag{13}$$

$$H_1: r(t) = s(t) + n(t) \tag{14}$$

El ruido es modelado como un proceso aleatorio con media cero y densidad espectral  $P_n(f)$ , nuestra tarea es escoger entre las dos hipótesis con la señal recibida  $r(t)$  y con lo que se conoce de  $s(t)$  y  $P_n(f)$ .

El enfoque básico usado para decidir cuál hipótesis seleccionar es usando un sistema LSI con una respuesta de impulso  $h(t)$ . La señal de salida  $y(t)$  producida por este sistema es examinada para encontrar su valor máximo  $y_{\max}$  el cual es comparado con un umbral  $T$  previamente seleccionado. Si  $y_{\max} \geq T$ , indica que la señal recibida contiene la señal transmitida y se escoge  $H_1$ . De otra forma, si  $y_{\max} < T$  indica que la señal recibida contiene solo ruido se escoge  $H_0$ . Esta operación se puede observar en la figura 6.

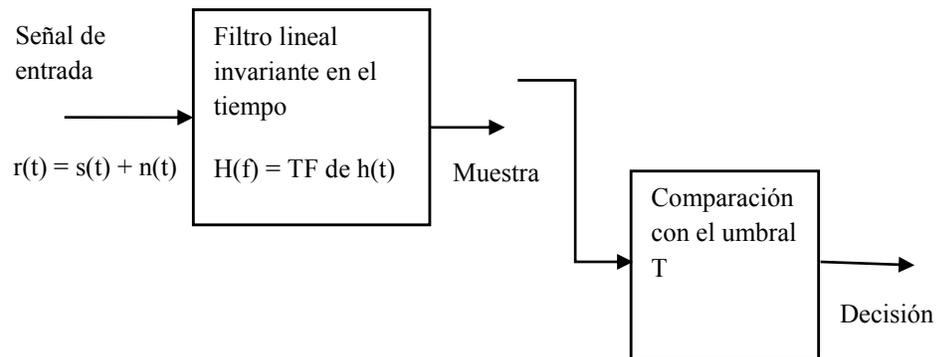


Figura 6. Filtro lineal para el problema de detección de señales binarias.

Si el umbral  $T$  preseleccionado es bajo, la probabilidad de un error en la identificación del objeto será menor, pero la probabilidad de una falsa alarma será mucho mayor. Si  $T$  es grande, ocurrirá lo contrario.

Un paso importante en este enfoque es el diseño del filtro  $H(f)$ . Un buen filtro puede producir una  $y_{\max}$  grande para la hipótesis  $H_1$  y hacer que el promedio de la

varianza del ruido inducido sea tan pequeño como sea posible. Lo cual requiere que  $H(f)$  maximice la señal a ruido (SNR), la cual se define de la siguiente forma:

$$SNR = \frac{|E\{y_{max}|H_1\}|^2}{var\{y_{max}\}}, \quad (15)$$

donde  $E\{y_{max}|H_1\}$  representa el pico de correlación esperado y  $var\{y_{max}\}$  representa la varianza. Cuando se supone que la media del ruido es cero,  $E\{y_{max}|H_1\}$  es el valor máximo de la salida del filtro cuando  $s(t)$  es la señal de entrada.

El filtro  $H(f)$  que maximiza la SNR está representado por la siguiente ecuación:

$$H(f) = \alpha \frac{S^*(f)}{P_n(f)}, \quad (16)$$

donde  $\alpha$  es una constante compleja arbitraria,  $S^*(f)$  es la Transformada de Fourier de la imagen de referencia conocida y  $P_n(f)$  es la densidad espectral del ruido aditivo. Para el caso de ruido blanco aditivo, el filtro que maximiza a SNR es,

$$H(f) = S^*(f) \quad (17)$$

#### IV.1.2. Filtros de sólo fase

Este tipo de filtros mejoran la eficiencia de procesamiento de señales luminosas.

Este filtro fue propuesto por (Horner et al. 1984) y se define de la siguiente forma:

$$H_{POF}(u, v) = \frac{S^*(u, v)}{|S(u, v)|} = e^{-j\theta(u, v)} \quad (18)$$

donde  $\theta(u, v)$  representa la fase de  $S(u, v)$ . Como se puede observar, POF tiene una magnitud de 1 en todas las frecuencias espaciales, permitiendo así el paso de toda la luz. Este filtro considera que la fase de una transformada de Fourier 2-D de una imagen contiene más información de la misma imagen que la magnitud de la transformada de Fourier de la imagen (véase figura 7). Este tipo de filtros ofrece un desempeño aceptable siempre y cuando se ignore la información de la magnitud.

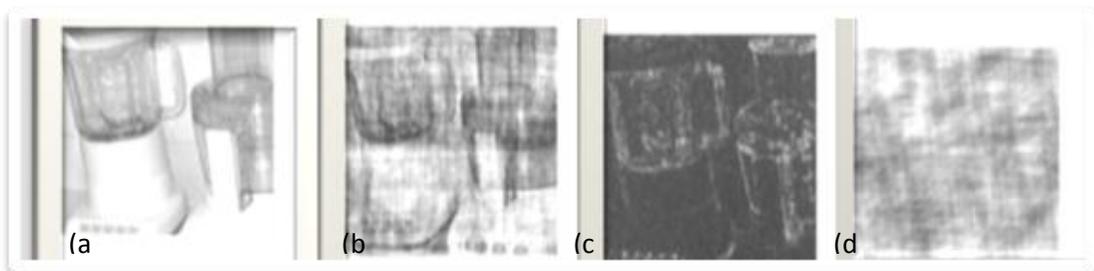


Figura 7.- Descomposición de la imagen en fase y magnitud: a) Imagen original, b) Fase de la imagen original, c) Fase de la imagen original pero con magnitud pseudo-aleatoria, d) la fase domina a la magnitud en sus efectos relativos para representar la información de la imagen

### IV.1.3. Filtros de sólo fase binarios

Los filtros de solo fase binarios (BPOF) se pueden obtener de varias formas, tales como la binarización de la parte real del filtro de acoplamiento, binarizando la parte imaginaria del filtro de acoplamiento o binarizando la suma de la parte imaginaria del filtro de acoplamiento. Todas esas diferentes binarizaciones pueden ser unificadas en un simple método con la ayuda del ángulo de línea de umbral (TLA, threshold line angle).

Usando TLA y  $\theta$  (que puede variar entre  $-\frac{\pi}{2}$  y  $\frac{\pi}{2}$ ), podemos definir un BPOF de la siguiente forma:

$$H_{BPOF}(f) = \text{sgn}[\text{Re}\{e^{-j\theta}S(f)\}] = \begin{cases} +1 & \text{si } -\frac{\pi}{2} \leq (\theta(f) - \theta) < \frac{\pi}{2} \\ -1 & \text{otro caso} \end{cases} \quad (19)$$

donde la función signo se define por  $\text{sgn}(x) = \begin{cases} +1 & \text{si } x \geq 0 \\ -1 & \text{si } x < 0 \end{cases}$

Para  $\theta = 0$ , BPOF se obtiene binarizando la parte real de la transformada de Fourier de la imagen. Si  $\theta = \frac{\pi}{2}$ , la binarización se refiere a la parte imaginaria de la transformada de Fourier de la imagen.

La ventaja de usar los BPOF's es que los filtros requieren muy poca capacidad de almacenamiento (un bit por pixel de frecuencia) y además se requieren pocos cálculos (no se realizan multiplicaciones en el dominio de la frecuencia, sólo se necesitan sumas y restas). Una desventaja de este filtro, al igual que los filtros de sólo fase es que no suprimen el ruido, lo cual produce una salida SNR muy baja.

## IV.2. Filtros avanzados

Como se comentó anteriormente, los filtros de acoplamiento son eficientes para reconocer patrones en presencia de ruido blanco aditivo ya que produce una salida de SNR alta. En el procesamiento de imágenes, la eficiencia de un filtro de correlación radica en su capacidad para producir picos de correlación invariantes a desplazamientos. Desafortunadamente los filtros de acoplamiento no cumplen con esta capacidad lo cual

los hace inadecuados para el reconocimiento de patrones porque la respuesta que produce se degrada cuando existe una variación en la imagen de entrada con respecto a la imagen de referencia. Estas variaciones en la imagen de entradas pueden ser cambios en escala, rotaciones o efectos de la iluminación.

Una solución a este problema pudiera ser la aplicación de una gran cantidad de filtros de acoplamiento, uno para cada variación en particular. Pero esto representa un gran costo computacional, pues requiere de mucho espacio de almacenamiento y una gran capacidad de procesamiento, lo cual lo hace un método poco práctico.

Para solucionar este problema es necesario diseñar filtros de correlación robustos que puedan sobrellevar las limitaciones de los filtros de acoplamiento. Una parte importante de los filtros de correlación tiene que ver con el diseño del filtro  $h(m, n)$  de tal forma que cumpla con los siguientes tres objetivos:

- Reconocer versiones distorsionadas de la imagen de referencia.
- Proveer robustez en presencia de ruido.
- Proveer un reconocimiento correcto manteniendo un intervalo de error bajo.

#### **IV.2.1. Invariancia a distorsiones en el plano**

Una imagen puede tener distorsiones debido a escalamiento, rotaciones en el plano y desplazamientos, estas distorsiones pueden ser modeladas matemáticamente en forma de transformaciones de coordenadas. Estas distorsiones pueden ser provocadas

por varios factores, algunos de ellos son: escalamiento de la imagen, desplazamiento de la imagen dentro del mismo plano, rotación en el plano, cambios provocados por los dispositivos de captura (dispositivo mal centrado, lente inadecuado, movimiento constante) y efectos de iluminación.

Consideremos a  $f(x, y)$  como una imagen en coordenadas cartesianas. Su transformación polar  $f(p, \theta)$  se calcula mediante la siguiente fórmula:

$$p = \ln(r) = \ln\{\sqrt{(x - x_0)^2 + (y - y_0)^2}\} \quad (20)$$

$$y \quad \theta = \text{atan}((y - y_0)/(x - x_0)) \quad (21)$$

Como se puede observar, la imagen es mapeada a una nueva sistema de coordenadas midiendo el logaritmo natural de la distancia radial y el ángulo con respecto al origen  $\{x_0, y_0\}$ .

#### **IV.2.2. Funciones armónicas circulares**

Las funciones armónicas circulares son muy útiles en el diseño de filtros de correlación invariantes a rotación en el plano. Una función armónica circular se puede definir, transformando de coordenadas cartesianas  $\{x, y\}$  a coordenadas polares  $\{r, \theta\}$ , de la siguiente forma:

$$r = \sqrt{(x - x_0)^2 + (y - y_0)^2} \quad (22)$$

$$y \quad \theta = \text{atan}((y - y_0)/(x - x_0)) \quad (23)$$

donde  $\{x_0, y_0\}$  es el centro de rotación de la imagen en el sistema de coordenadas cartesianas.

En algunas aplicaciones conviene expandir la función armónica circular a una imagen en el dominio de la frecuencia. Consideremos  $F(r, \theta)$  como la transformación polar de  $F(u, v)$  es decir, la transformada de Fourier de  $f(x, y)$ . La función armónica circular correspondiente se define de la siguiente manera:

$$F_k(r) = \frac{1}{2\pi} \int_0^{2\pi} F(r, \theta) e^{-jk\theta} d\theta \quad (24)$$

De modo que la transformada de Fourier de la imagen puede ser expandida en términos de funciones armónicas circulares como:

$$F(r, \theta) = \sum_{k=-\infty}^{\infty} F_k(r) e^{jk\theta} \quad (25)$$

Esta función armónica circular puede ser usada para obtener un filtro invariante a rotación, consideremos  $C(t_\theta)$  como la salida del filtro cuando la imagen de entrada es rotada  $t_\theta$  grados. Cuando la imagen tiene una rotación de  $t_\theta = 0^\circ$  la salida del filtro se expresa de la siguiente manera:

$$\begin{aligned} C(0) &= \iint F(u, v) H^*(u, v) du dv \\ &= \iint F(r, \theta) H^*(r, \theta) r dr d\theta \end{aligned} \quad (26)$$

donde  $H(u, v)$  es la transformada de Fourier del filtro. Un filtro para una imagen con una rotación de  $t_\theta$  grados es la siguiente:

$$C(t_\theta) = \sum_{k=-\infty}^{\infty} C_k e^{jkT_\theta}, \quad (27)$$

$$\text{donde } C_k = 2\pi \int_{r=0}^{\infty} F_k(r) H_k^*(r) r dr \quad (28)$$

Como se acaba de ver, se puede obtener un filtro invariante a rotación mediante la función armónica circular de la imagen de referencia. Sin embargo, estos filtros son muy pobres para discriminar información inservible.

### IV.2.3. Filtros compuestos de correlación

Los filtros compuestos de correlación se desarrollaron para manejar los diferentes tipos de distorsiones que no pueden ser modeladas matemáticamente por transformaciones de coordenadas cartesianas o por funciones armónicas circulares. Estos filtros se obtienen de varias imágenes de entrenamiento, las cuales representan diferentes vistas de la imagen del objeto o patrón a identificar, además que representan las distorsiones esperadas para el objeto.

Una ventaja de estos filtros es la invariancia a distorsiones geométricas, cambios espectrales y fenómenos provocados por el sensor de visión.

### IV.2.4. Funciones sintéticas discriminantes (SDF)

Considere a  $x_1(x, y)$ ,  $x_2(x, y)$ , ...,  $x_N(x, y)$  como las  $N$  imágenes de entrenamiento que representan las posibles distorsiones para la imagen de referencia  $x(x, y)$ . Ahora considere  $x(u, v)$  su transformada de Fourier bidimensional:

$$X(u, v) = \iint x(x, y) \exp[-i2\pi(ux + vy)] dx dy \quad (29)$$

donde  $u$  y  $v$  representan las frecuencias espaciales.

El objetivo de la función sintética discriminante consiste en diseñar una imagen compuesta  $h(x, y)$  de tal forma que representen y describan las distorsiones esperadas para la imagen de referencia. Un filtro SDF es una combinación lineal de filtros de acoplamiento espacial (MSF) para diferentes patrones.

Sean  $\{x_i(x, y), i=1, 2, \dots, N\}$  un conjunto de imágenes de entrenamiento cada una con  $d$  píxeles y  $h(x, y)$  con representaciones vectoriales (que se obtienen escaneando un arreglo 2-D de izquierda a derecha y de arriba hacia abajo, la secuencia de números resultante se coloca dentro de un vector columna) como se muestra en la figura 8.

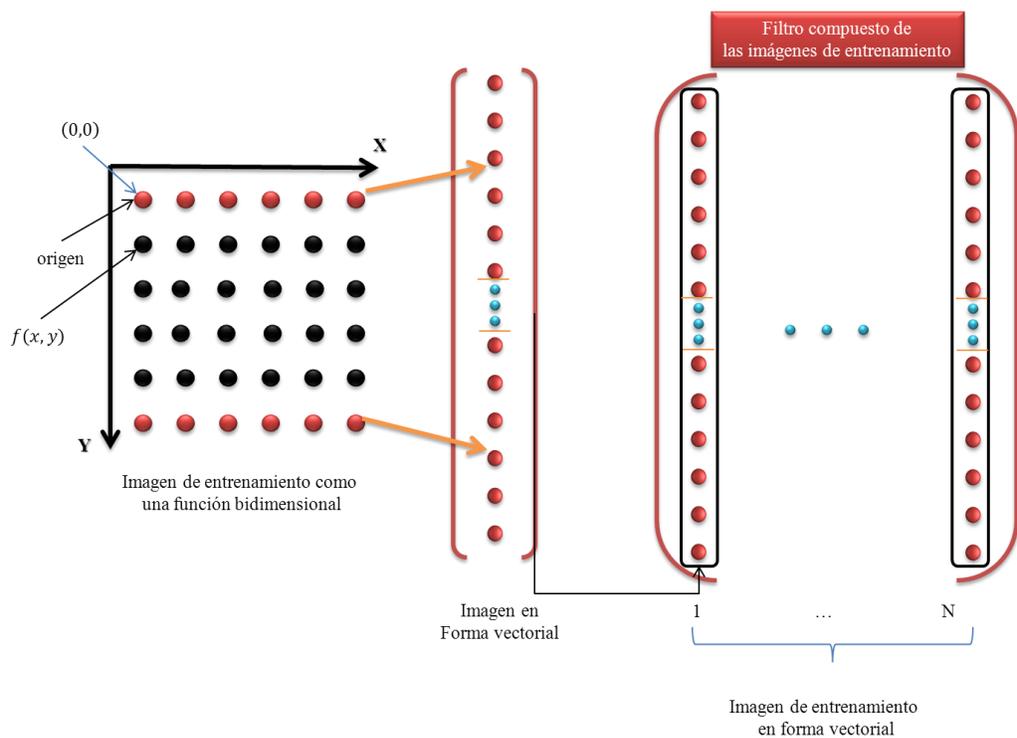


Figura 8. Conversión de una imagen bidimensional a su representación vectorial y su inclusión en la matriz que generará el filtro compuesto.

El valor en el origen del plano de correlación producida por el filtro  $h(x, y)$  en respuesta a una imagen de entrenamiento  $x_i(x, y)$  está dado por la siguiente ecuación:

$$X_i^T h = u \quad (30)$$

donde  $\mathbf{h}$  es el vector del filtro que contiene  $d$  pixeles,  $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]$  es una matriz de tamaño  $d \times N$  con los  $N$  vectores imágenes de entrenamiento como su columna, y  $\mathbf{u}=[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_N]^T$  es un vector de tamaño  $N \times 1$  que contiene los valores del pico esperado para cada imagen de entrenamiento. Sin embargo, el número de imágenes de entrenamiento  $N$  por lo general es mucho menor que el número de frecuencias en el filtro. Para encontrar una única solución, suponemos que  $\mathbf{h}$  es una combinación lineal de las imágenes de entrenamiento, es decir,  $\mathbf{h}=\mathbf{X}\mathbf{a}$  donde  $\mathbf{a}$  es el vector de pesos para las combinaciones lineales de las columnas de la matriz de datos  $\mathbf{X}$ . Para determinar  $\mathbf{a}$ , se substituye por  $\mathbf{h}$  en la ecuación 30.

$X^T(\mathbf{X}\mathbf{a})=\mathbf{u} \Rightarrow \mathbf{a}=(X^T X)^{-1} \mathbf{u}$  y substituyendo esta solución en  $\mathbf{h}=\mathbf{X}\mathbf{a}$  obtenemos el filtro SDF:

$$\mathbf{h} = X(X^T X)^{-1} \mathbf{u} \quad (31)$$

donde  $\mathbf{h}$  es el vector  $d \times 1$  del filtro.

El proceso para determinar la clase de un patrón de prueba requiere de la localización del pico de correlación en la superficie de correlación. La posición del patrón en la escena de entrada se determina localizando el pico. Una vez que

localizamos el pico, podemos decir que el filtro “reconoce” el patrón cuando el valor del pico excede de un umbral.

Sin embargo, un filtro SDF convencional es muy dado a producir un pico de correlación rodeado por un gran lóbulo lateral que nos lleva a errores si este es más grande que el pico principal.

#### **IV.2.5. Filtro del Mínimo Promedio de la Energía de Correlación (MACE)**

Una forma de suprimir los lóbulos laterales producidos por un filtro SDF convencional es minimizando la energía en el plano de correlación. La energía de correlación promedio (ACE – average correlation energy) para  $N$  imágenes de entrenamiento se define de la siguiente forma:

$$ACE = \frac{1}{N} \sum_{i=1}^N \sum_m^{d_1} \sum_n^{d_2} |g_i(m, n)|^2 \quad (32)$$

donde  $d$  es un vector columna de una imagen de entrenamiento. Usando el teorema de Parseval, la ACE se puede expresar en el dominio de la frecuencia como

$$ACE = \frac{1}{d.N} \sum_{i=1}^N \sum_k^{d_1} \sum_l^{d_2} |G_i(k, l)|^2 \quad (33)$$

donde  $G_i(k, l)$  es la transformada de Fourier de  $g_i(m, n)$ . De esta forma  $G_i(k, l) = H(k, l)X_i^*(k, l)$ , y la ACE expresada en el dominio de las frecuencias se convierte en:

$$ACE = \frac{1}{d.N} \sum_{i=1}^N \sum_k^{d_1} \sum_l^{d_2} |H(k, l)|^2 |X_i(k, l)|^2 \quad (34)$$

Por simplicidad, consideremos  $\mathbf{H}(\mathbf{k}, \mathbf{l})$  como un vector  $\mathbf{h}$ , y definamos una matriz diagonal  $X_i$  cuyos elementos a lo largo de la diagonal principal son  $X_i(k, l)$ . De esta forma  $\mathbf{h}$  y  $X_i$  representan el filtro y las imágenes de entrenamiento en el dominio de la frecuencia respectivamente. Por lo tanto, expresamos a ACE de la siguiente forma:

$$ACE = \frac{1}{d.N} \sum_{i=1}^N (H^+ X_i)(X_i^* h) = h^+ \left[ \frac{1}{d.N} \sum_{i=1}^N X_i X_i^* \right] h = h^+ D h \quad (35)$$

donde  $D = \frac{1}{d.N} \sum_{i=1}^N X_i X_i^*$  es una matriz diagonal de tamaño  $d \times d$ . Las restricciones sobre el pico de correlación deben ser expresadas en el dominio de la frecuencia. El producto interno en el dominio espacial es directamente proporcional a los productos internos en el dominio de la frecuencia, por lo tanto la ecuación se expresa de la siguiente manera:

$$X^+ h = d \cdot u.$$

donde  $X$  es una matriz cuyas columnas  $x_i$  son la representación vectorial de la transformada de Fourier de las imágenes de entrenamiento.

Una solución óptima se logra usando el método de los multiplicadores de Lagrange, el cual produce el filtro dado por la siguiente ecuación:

$$\mathbf{h} = D^{-1} X (X^+ D^{-1} X)^{-1} \mathbf{u} \quad (36)$$

Esta es una solución en el dominio de las frecuencias y  $\mathbf{h}$  es por lo tanto el vector columna que contiene el arreglo del filtro  $H(k, l)$  en el dominio de las frecuencias.

Los filtros MACE son efectivos para encontrar imágenes de entrenamiento en una escena con fondo confuso. Eliminan el lóbulo lateral para asegurar un pico de

correlación agudo y fácil de distinguir, reduciendo la probabilidad de errores. Sin embargo, tienen dos inconvenientes; no son inmunes a ruidos y son demasiado sensibles a cambios de variaciones intra-clases.

#### IV.2.6 Filtro SDF de mínima varianza (MVSDF)

Consideremos la situación donde la imagen de entrada es uno de las imágenes de entrenamiento  $s_i$  afectados por ruido blanco aditivo  $n$ . La salida del filtro en respuesta a un vector de entrenamiento  $x_i$  está representado por:

$$(x_i + v)^T h = x_i^T h + v^T h = u_i \delta \quad (37)$$

MVSDF minimiza la varianza de  $\delta$  para minimizar las fluctuaciones en la salida del filtro. Usando el método de los multiplicadores de Lagrange obtenemos el siguiente filtro MVSDF:

$$h = C^{-1} X (X^T C^{-1} X)^{-1} u \quad (38)$$

donde  $C$  es la matriz de covarianza del ruido.

Podemos representar MVSDF en el dominio de las frecuencias considerando a  $X$  como una matriz con columnas que representan a las imágenes de entrenamiento, a cada imagen se le aplicó la transformada de Fourier. La ecuación queda de la siguiente manera:

$$h = P^{-1} X (X^+ P^{-1} X)^{-1} u \quad (39)$$

### IV.2.7. Filtros tolerantes a distorsiones

Los filtros SDF son tolerantes a distorsiones siempre y cuando existan restricciones que controlen los picos de correlación para las imágenes de entrenamiento. Aquí se presenta otro enfoque, el cual consiste en eliminar las restricciones. Los motivos para usar este enfoque son los siguientes: las imágenes de entrenamiento siempre producen valores diferentes del especificado para las imágenes de entrenamiento, no existe una relación formal entre las restricciones existentes impuestas sobre la salida del filtro y su habilidad para tolerar distorsiones.

La idea clave de este enfoque consiste en tratar el plano de correlación como un nuevo patrón generado por el filtro en respuesta a la imagen de entrada. Una métrica para la distorsión se define como la variación promedia en imágenes después del filtrado. Si  $g_i(m, n)$  es la superficie de correlación producida en respuesta a la  $i$ -ésima imagen de entrenamiento, la variación promedia entre la superficie de correlación de la imagen de entrenamiento en error cuadrado medio (MSE) es:

$$\text{métrica de distorsión} = \frac{1}{N(N-1)} \sum_{i=1}^N \sum_{j=1}^N \sum_m \sum_n [g_i(m, n) - g_j(m, n)]^2 \quad (40)$$

Esta métrica se conoce también como la medida de similitud promedio (ASM).

Consideremos  $x_i$  como un vector que represente la transformada de Fourier de la  $i$ -ésima imagen de entrenamiento y  $m = \frac{1}{N} \sum_{i=1}^N x_i$  es el promedio de las imágenes de entrenamiento. Definimos la matriz diagonal  $\mathbf{M}$  y  $\mathbf{X}_i$  con los elementos de  $m$  y  $x_i$  a lo

largo de la diagonal principal. La transformada de Fourier del plano de correlación que se produce en respuesta a la  $i$ -ésima imagen de entrenamiento se puede obtener con:

$g = X^*h$ , donde  $h$  es el vector del filtro en el dominio de las frecuencias. De esta forma multiplicando el vector del filtro  $h$  por la matriz diagonal  $X^*$  obtenemos  $g$  que alcanza los mismos resultados que

$$G_i(k, l) = X_i^*(k, l)H(k, l) \quad (41)$$

Los requisitos para que un filtro sea tolerante a distorsión, es que debe producir un valor de pico muy grande que facilite la detección del patrón y así localizar su posición en la escena de entrada.

Mientras que el desempeño de un filtro SDF está controlado por las restricciones que se le imponen a la salida del filtro, el filtro MACH las evita pues su objetivo es sólo obtener un gran pico de correlación.

#### **IV.2.8. Filtro compuesto no lineal k-law**

Un filtro compuesto no lineal que logra una invariancia a rotación, escala, y/o ruido son los filtros compuestos k-law. Este tipo de filtros realizan varias correlaciones no lineales entre una escena de entrada y diferentes imágenes de referencia. Un filtro no lineal se puede expresar de la siguiente forma:

$$H(u, v)_{KL} = |F(u, v)|^k e^{-i\phi(u, v)}, \quad 0 < k < 1 \quad (42)$$

donde  $F(u, v)$  es la transformada de Fourier del objeto que deseamos reconocer,  $\phi$  es la fase de la imagen objetivo  $|\cdot|$  es el modulo de  $F(u, v)$  y  $k$  es el factor de no linealidad.

Si se cambia los valores de  $k$  a 1 obtenemos un filtro de acoplamiento clásico, si se cambia a 0 obtenemos un filtro de solo fase y si se cambia a -1 obtenemos un filtro inverso. Cuando el operador de no linealidad modifica la transformada de Fourier en ambas imágenes, es decir en la escena de entrada y en la imagen de referencia, se considera que el correlacionador no es lineal.

#### IV.2.9. Filtros de correlación clasificadores de distancia

Algunos de los filtros descritos anteriormente son sistemas lineales cuya respuesta a patrones de interés son controlados de forma cuidadosa por varias técnicas de optimización. Otro enfoque se basa en la idea de aplicar transformaciones a los datos de entrada.

Considere la correlación de dos secuencias discretas  $x(n)$  y  $h(n)$ , cada uno con una longitud  $L$ . El resultado será una secuencia de longitud  $2L-1$ , el cual se obtiene con la siguiente ecuación:

$$g(n) = x(n) \otimes h(n) = \sum_{k=1}^L x(n+k)h(k), -L+1 \leq n \leq n-1 \quad (43)$$

Esta ecuación se puede escribir en notación matricial y vectorial de la siguiente manera:

$$\mathbf{g} = \mathbf{H}\mathbf{x}, \quad (44)$$

donde  $H$  es una matriz cuyas filas son versiones desplazadas de la secuencia  $h(n)$ . Así, la estructura de  $H$  es de tal forma que su multiplicación por el vector de datos de entrada  $x$  implementa el desplazamiento, multiplicación y agrega operaciones de correlación para producir el vector de salida  $g$ . Los vectores de entrada y salida se obtienen ordenando lexicográficamente las filas o columnas, mientras que la matriz  $H$  contiene bloques ordenados siguiendo la estructura Toeplitz y cada bloque por sí mismo es una matriz Toeplitz.

La distancia de un vector de prueba  $x$  a una referencia  $m_k$  bajo una transformación lineal  $H$  está dado por la siguiente ecuación:

$$d_k = [H^*x - H^*m_k]^2 = (x - m_k)^+ HH^*(x - m_k) \quad (45)$$

Aplicando la distancia de Mahalanobis se normaliza las características espaciales. Entonces dado una entrada desconocida, se calcula su distancia a todas las clases centradas en el espacio transformado. La entrada es asignada a la clase en la que la distancia entre el objetivo y la entrada es la más menor. La distancia de Mahalanobis y otras transformaciones son difíciles de aplicar a imágenes, ya que requieren de extracción de características o reducción de dimensiones, segmentación y almacenamiento.

El filtro de correlación clasificador de distancias transforma las imágenes de entrenamiento en diferentes clases, las cuales se encuentran separadas unas de otras, y son más compactos en el nuevo espacio de transformación que en el espacio original.

Para una entrada de prueba, esta se clasifica de acuerdo a la distancia existente entre la entrada y las diferentes clases.

#### IV.2.10. Filtros polinomiales de correlación

La mayoría de los filtros aquí descritos son lineales. Sin embargo, el filtro de correlación polinomial (PCF) permite la no linealidad de los filtros para lograr un mejor desempeño en el reconocimiento de patrones. En el PCF, cada energía no lineal introduce una nueva rama en la arquitectura de correlación.

La principal diferencia entre un enfoque tradicional y el enfoque polinomial es que la salida de un filtro PCF es una función no lineal de la entrada. Para una entrada como una imagen, la salida se obtiene mediante la ecuación:

$$g_x = A_1x^1 + A_2x^2 + \dots + A_Nx^N \quad (46)$$

Esta ecuación se conoce como PCF,  $x$  representa la imagen de entrada en su forma vectorial,  $g_x$  es un vector que representa el plano de correlación resultante como una función polinomial de  $x$ . Para asegurarse de que el filtro sea invariante a desplazamiento, todos los coeficientes matriciales tienen que ser Toeplitz. Como se puede observar, cada término en el polinomio se puede calcular como una operación de filtrado invariante a desplazamiento, es decir,

$$A_i x^i \equiv h_i(m, n) \otimes x^i(m, n) \quad (47)$$

La salida del filtro de correlación polinomial se expresa matemáticamente como:

$$g_x(m, n) = \sum_{i=1}^N h_i(m, n) \otimes x^i(m, n) \quad (48)$$

Las áreas de aplicación de los filtros de correlación son diversas. A través de los años, muchos investigadores han contribuido al desarrollo de esta área, de los cuales solo unos pocos se mencionan en este trabajo.

### **IV.3. Medidas de desempeño de los filtros de correlación**

Existen muchas medidas de desempeño para un filtro de correlación, algunas de esas medidas basadas en la capacidad del filtro para reconocer a un objeto afectado por ruido blanco aditivo, otros en la capacidad de filtro para rechazar objetos false, entre otros. A continuación se mencionan algunas de esas métricas de desempeño.

Para caracterizar la agudeza de un pico de correlación se han propuesto varias métricas, una de esas métricas es la razón del pico de correlación a lóbulos laterales (PSR). El PSR es la razón del pico de correlación para la desviación estándar de los valores de correlación en una región que está centrada sobre el pico, pero excluyendo una región pequeña alrededor del pico.

$$PSR = \frac{(pico-media)}{desviación\ estandar} \quad (49)$$

PSR se usa para conocer que tan bien, una región en la escena a identificar, se acopla con la imagen objetivo, es decir, cuando la salida de correlación contiene múltiples picos, los cuales se ordenan de acuerdo a sus valores PSR.

Otra medida para la agudeza del pico de correlación es la energía pico-correlación (PCE), el cual se define de la siguiente manera:

$$PCE = \frac{|c(o)|^2}{\int |c(x)|^2 dx} = \frac{|\int s(f)H(f)df|^2}{|\int s(f)H(f)|^2 df} \quad (50)$$

El propósito de PCE es caracterizar la agudeza del pico resultante de un filtro de correlación. Como se puede observar que un máximo PCE se obtiene cuando la salida de correlación es una función delta y que es cero o cercana a cero cuando la salida es una constante.

Otra métrica para medir el desempeño de un filtro de correlación es la que se denomina como la capacidad de discriminación (DC), también conocido como coeficiente de discriminación. La capacidad de discriminación es la habilidad que posee un filtro de correlación para distinguir a un objeto de otros diferentes.

$$DC = 1 - (C^R(0,0)^2 / C^O(0,0)^2), \quad (51)$$

donde  $C^R$  es el valor máximo del plano de correlación sobre la región a rechazar y  $C^O$  es valor máximo del plano de correlación sobre la región donde se encuentra el objeto patrón. Si DC es un valor muy cercano a uno, indica que el filtro tiene un buen desempeño, en caso contrario si es un valor que tiende a cero o valores negativos indica que el filtro tiene una mala capacidad de discriminación.

## Capítulo V

### Evaluación del desempeño de algunos filtros de correlación

Los filtros de correlación están teniendo actualmente una gran aceptación por parte de científicos, ingenieros y tecnólogos debido a su gran fundamento matemático. En la literatura especializada podemos encontrar mucha información sobre los filtros de correlación. Si indagamos en la red, encontraremos una gran cantidad de aplicaciones de reconocimiento de patrones por correlación, desde una simple aplicación computacional para fines educativos hasta un sofisticado sistema de visión artificial para un robot.

El uso de los filtros de correlación ha sido posible gracias al desarrollo que ha tenido el poder computacional, pues es muy sencillo conseguir una computadora con una capacidad de procesamiento y de almacenamiento a muy bajo costo. Cabe mencionar también el gran desarrollo que ha tenido el software de paquete que viene con grandes capacidades de procesamiento matemático que facilita el proceso de desarrollo de aplicaciones para sistemas de visión artificial.

Para que un sistema de reconocimiento automático de objetos por correlación funcione correctamente en un ambiente real, debe de contener un filtro invariante a distorsiones, tales como rotación, desplazamiento, cambios de escala e iluminación. Cada filtro propuesto en la literatura maximiza uno o más criterios para ofrecer invariancia a las distorsiones (véase apéndice A) y maximizar su desempeño.

El objetivo de esta sección es evaluar el desempeño de tres filtros de correlación con el fin de seleccionar el más adecuado para implementarlo en un sistema de visión artificial. El primero es el filtro de acoplamiento para efectos de demostración y comparación ante dos filtros compuestos-uno lineal y otro no-lineal-, el segundo es un filtro compuesto SDF de sólo fase el cual logra una máxima eficiencia frente a la iluminación y por último el filtro compuesto no lineal SDF de sólo fase k-law, el cual ofrece invariancia a la rotación, escalamiento e iluminación.

Para obtener el desempeño de cada filtro se usa la métrica conocida como la capacidad de discriminación (DC), la cual nos permite conocer que tan bueno es un filtro para reconocer objetos verdaderos y rechazar aquellos que sean falsos.

## **V.1. Filtro de acoplamiento**

Los filtros de acoplamiento proveen una optimización a la razón señal-a-ruido en el reconocimiento de una señal de referencia que contiene ruido blanco aditivo. Sin embargo este tipo de filtro no es tolerante a distorsiones tales como la rotación, el desplazamiento y los efectos de la iluminación.

La figura 9 esquematiza las operaciones necesarias para obtener el plano de correlación usando el filtro de acoplamiento. Como se puede observar, es un proceso simple que solo involucra la aplicación de la transformada de Fourier tanto a la imagen de prueba como a la imagen de referencia y la operación de correlación entre ambas para obtener el plano de correlación.

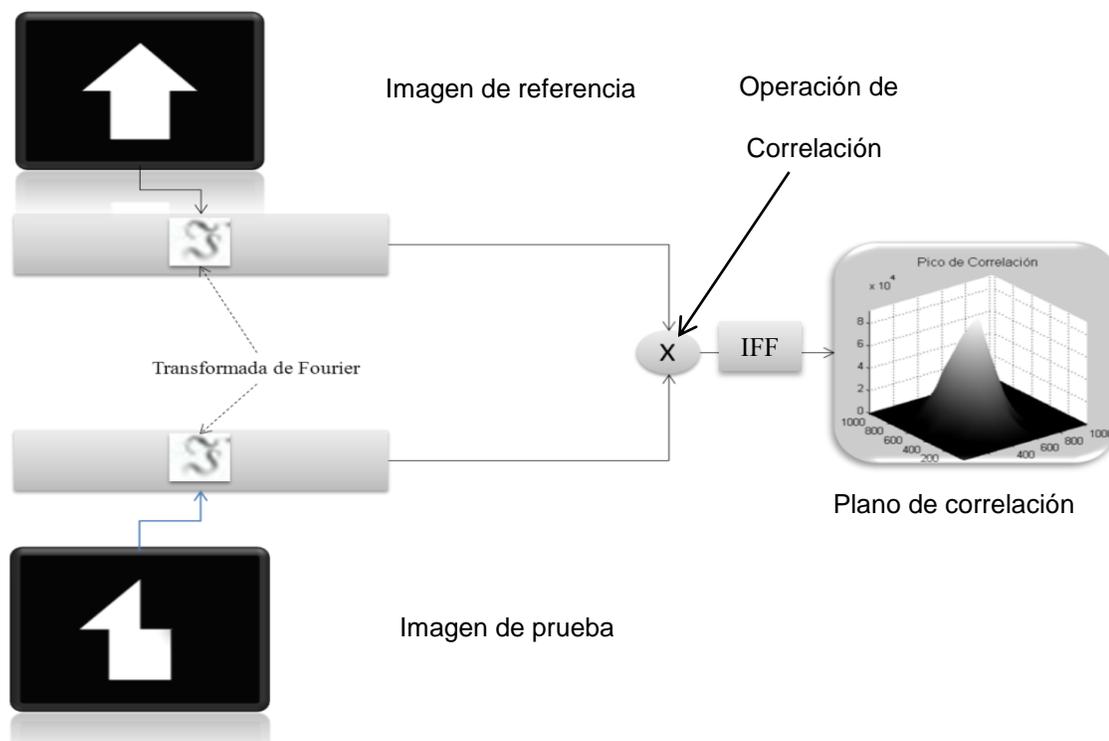


Figura 9. Operaciones necesarias para obtener el plano de correlación con el filtro de acoplamiento.

Considere la figura 10 como nuestra imagen de referencia o patrón, y configuraciones parciales de la misma como escenas de entrada. Usando el filtro de acoplamiento para obtener el plano de correlación entre la imagen de referencia con una imagen de prueba (en este caso la misma imagen), se obtiene el resultado mostrado en la figura 11.

Como se puede observar en la figura 11, se realizó una operación de auto-correlación, es decir, la imagen de referencia correlacionada consigo misma. Esta operación produce un pico de correlación con una DC = 0.720225, lo cual muestra la pobre capacidad de discriminación de este filtro.

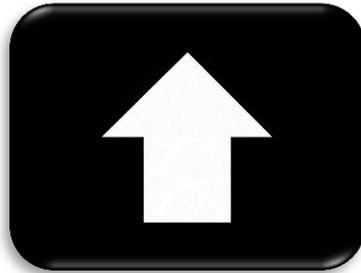


Figura 10.-Imagen de referencia

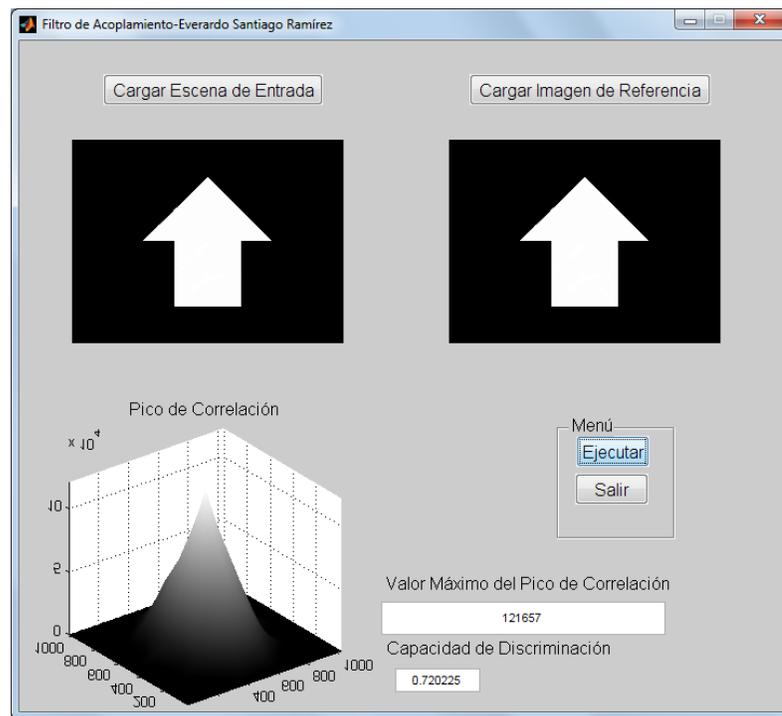


Figura 11.- Resultado producido por el filtro de acoplamiento

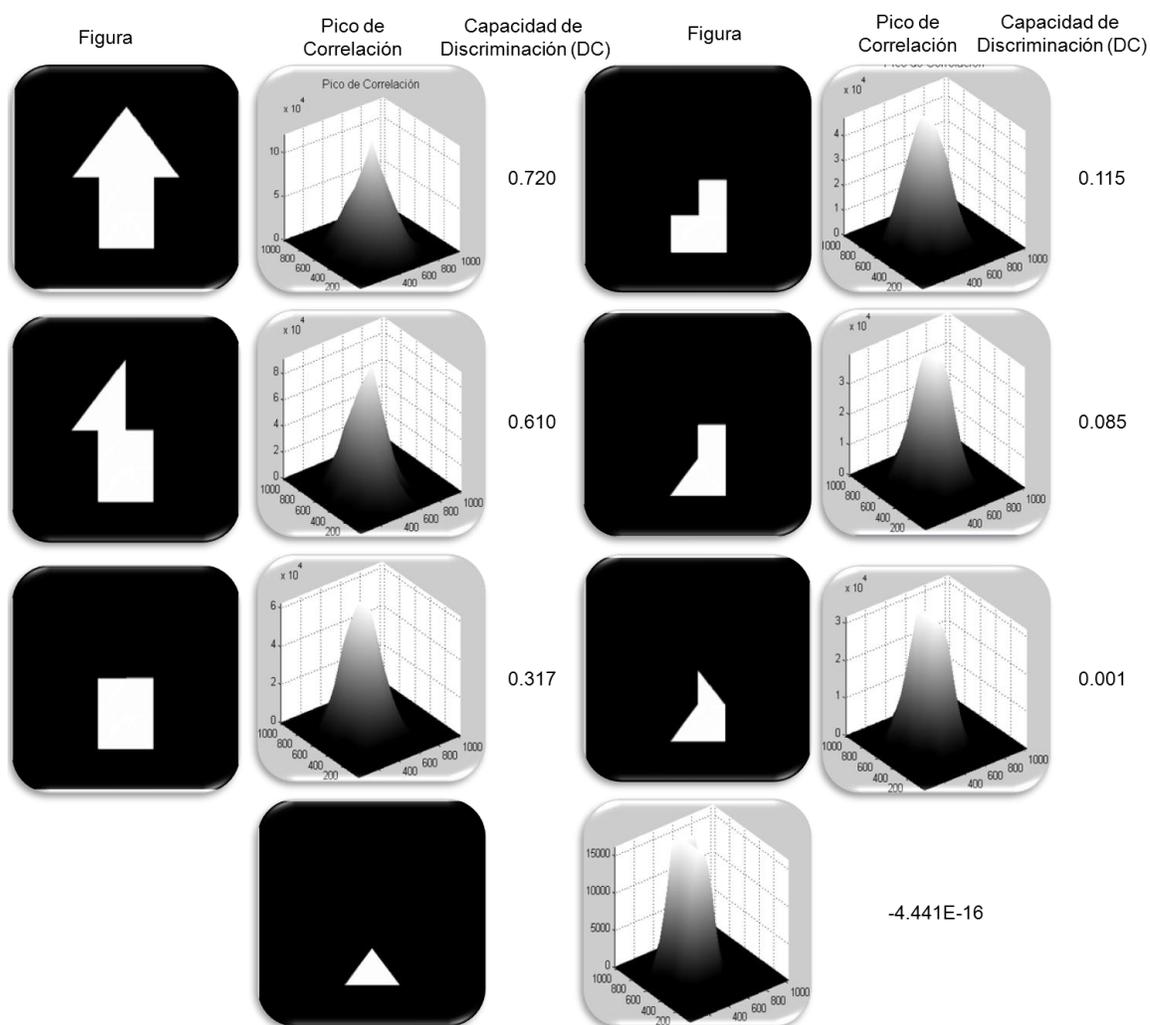


Figura 12.- Gráfica de resultados del filtro de acoplamiento

Para observar el comportamiento de este filtro, se probó con configuraciones parciales de la figura 10 obteniendo los resultados representados por la figura 12. Como se puede observar, cada pieza del tangrama contribuye al valor del pico de correlación, entre más piezas sean mayor será el valor del pico de correlación. También se puede observar que conforme la escena de entrada tenga menos piezas, la capacidad del filtro

para discriminar objetos falsos se reduce considerablemente, otra cosa que se puede observar es que mientras va creciendo el valor del pico de correlación, la punta del pico se va haciendo más aguda, pero siempre están presentes los lóbulos laterales.

## **V.2. Filtro compuesto SDF de sólo fase**

La deficiencia de luminosidad de los filtros de acoplamiento se mejora usando los filtros de solo fase (POF, Phase Only Filter) porque en el plano de la frecuencia pasa toda la luz incidente. Sin embargo, en el procesamiento de imágenes los filtros deben ser capaces de producir picos invariantes a desplazamientos, escala, rotación o pequeñas deformaciones de contornos causados por la iluminación.

El enfoque de los filtros compuestos de correlación propone resolver estos problemas creando una imagen compuesta. Dicha imagen se compone de imágenes que representan las posibles distorsiones esperadas para la imagen de referencia. Este conjunto de imágenes es conocido como conjunto de imágenes entrenamiento. Las imágenes consideradas en el conjunto de entrenamiento para el diseño del filtro son variaciones en escala, rotación y/o iluminación de la imagen de referencia, refiérase a la figura 13 como un ejemplo de un conjunto de imágenes de entrenamiento. La imagen compuesta es una combinación lineal de las imágenes de entrenamiento tal como se puede apreciar en la figura 14.

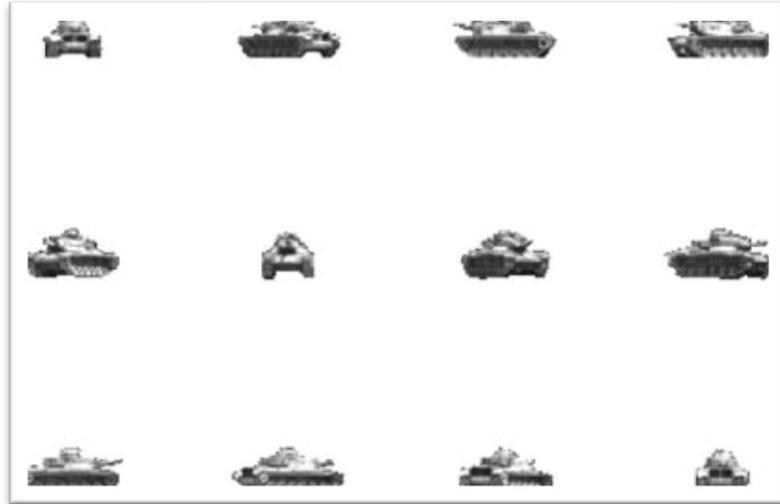


Figura 13. Ejemplo de conjunto de imágenes de entrenamiento.



Figura 14. Imagen compuesta formada por el conjunto de entrenamiento representado por la figura 13

Para formar la imagen compuesta de la Figura 14, se utilizó cada uno de los objetos en la figura 13 de forma independiente y centrada en un arreglo bidimensional de 256 x 256 píxeles. Para el caso del tangrama, el cual consiste de siete piezas que colocadas en una posición específica forman una figura, cada pieza del Tangrama debe permanecer en la posición determinada con el objetivo de que cada imagen de

entrenamiento contribuya en la magnitud del pico de correlación en la salida. Lo anterior significa que si solo una imagen es colocada en posición correcta, la magnitud obtenida será aproximadamente  $1/7$  de la magnitud que se obtendría si se tuvieran las 7 piezas formando la figura deseada.

Usando una función sintética discriminante (SDF) de solo fase, se logra un sistema de reconocimiento de patrones invariante a distorsiones. Las operaciones necesarias para obtener el pico de correlación con este filtro es mostrado en la figura 15. El primer paso en este proceso consiste en seleccionar la figura que deseamos armar, es decir, nuestra imagen de referencia. Posteriormente se diseña el filtro SDF  $h(x, y)$  usando las piezas del tangrama configuradas de tal forma que la unión de todas forma la imagen de referencia. Al filtro se le calcula el filtro-fase SDF  $H(u, v)$ . A la imagen que queremos obtener su grado de similitud con nuestra imagen de referencia le aplicamos la transformada de Fourier. Una vez realizado este paso, realizamos la operación de correlación entre la imagen de prueba y el filtro SDF  $H(u, v)$ , produciendo de esta forma el plano de correlación.

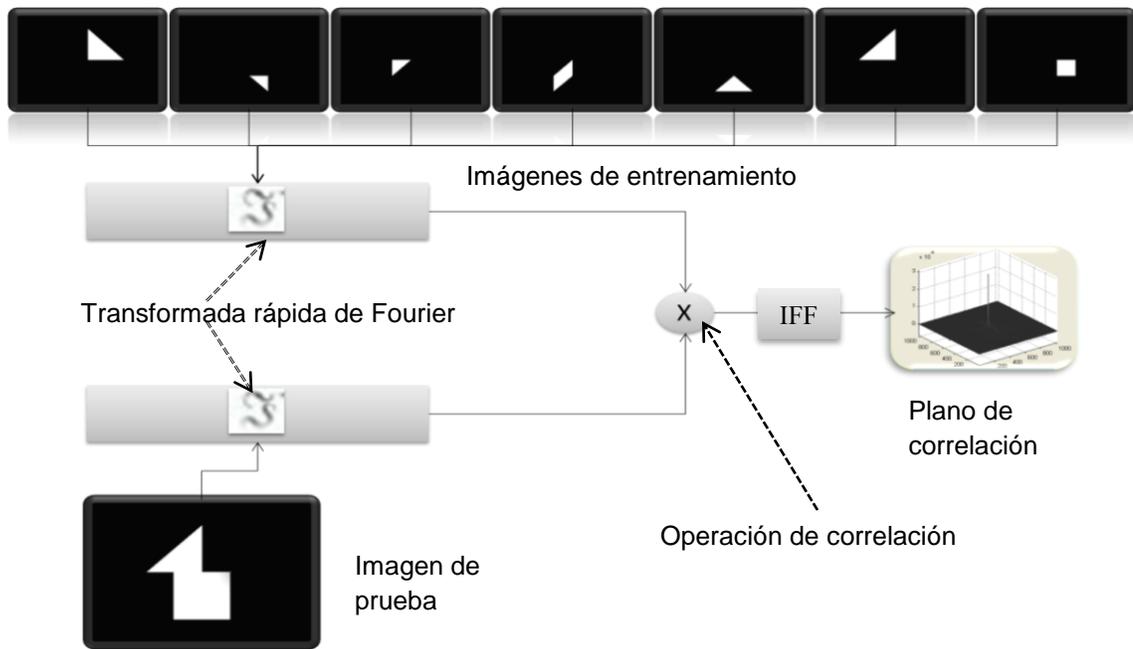


Figura 15. Operaciones necesarias para obtener el plano de correlación con el filtro compuesto SDF de sólo fase.

Para la evaluación de este filtro considere la figura 10 como imagen de referencia. Las imágenes de entrenamiento en este caso son cada uno de las piezas del tangrama, configuradas de tal forma que la unión de las mismas en una sola imagen forma la imagen de referencia.

Este filtro produce un pico de correlación más agudo y ya no se presentan los lóbulos laterales, además de que la  $DC = 0.984563$  está más cerca de 1, lo cual indica su gran capacidad de discriminación. Tanto a la imagen compuesta por el conjunto de entrenamiento como a la escena de entrada, se les aplica la transformada de Fourier,

posteriormente la operación de correlación produce el plano de salida, ver resultado en la figura 16.

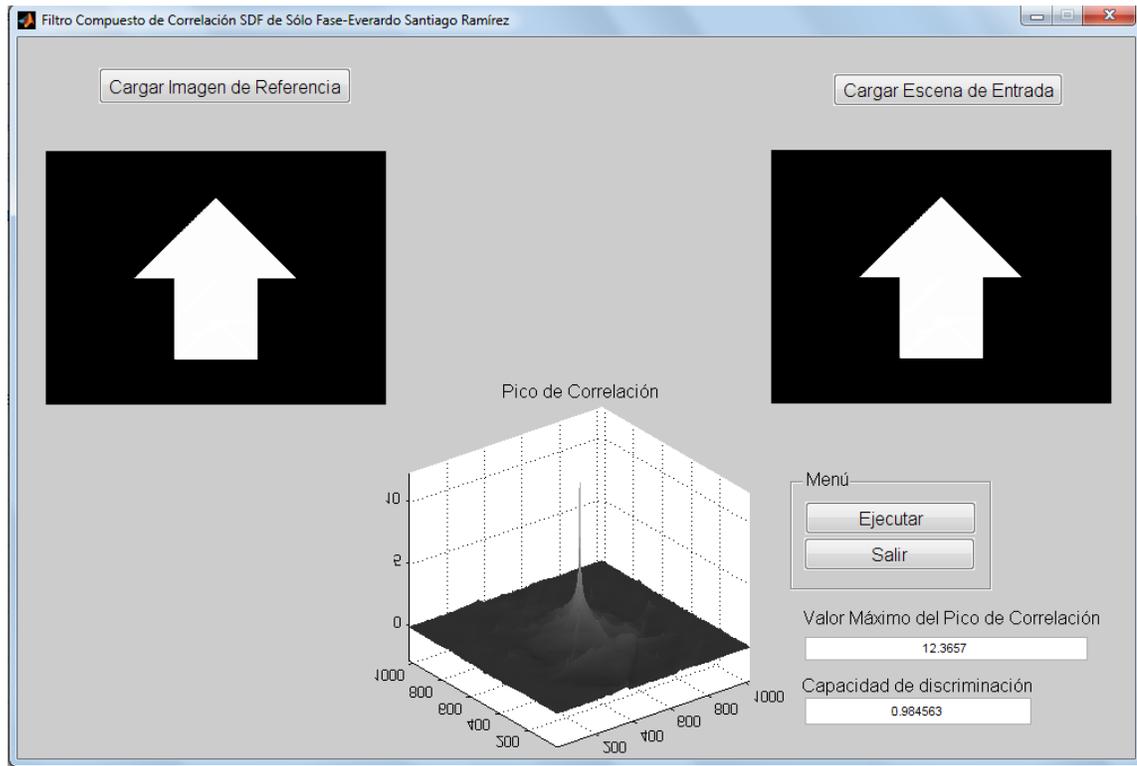


Figura 16.- Reconocimiento de objetos con el filtro SDF de sólo fase

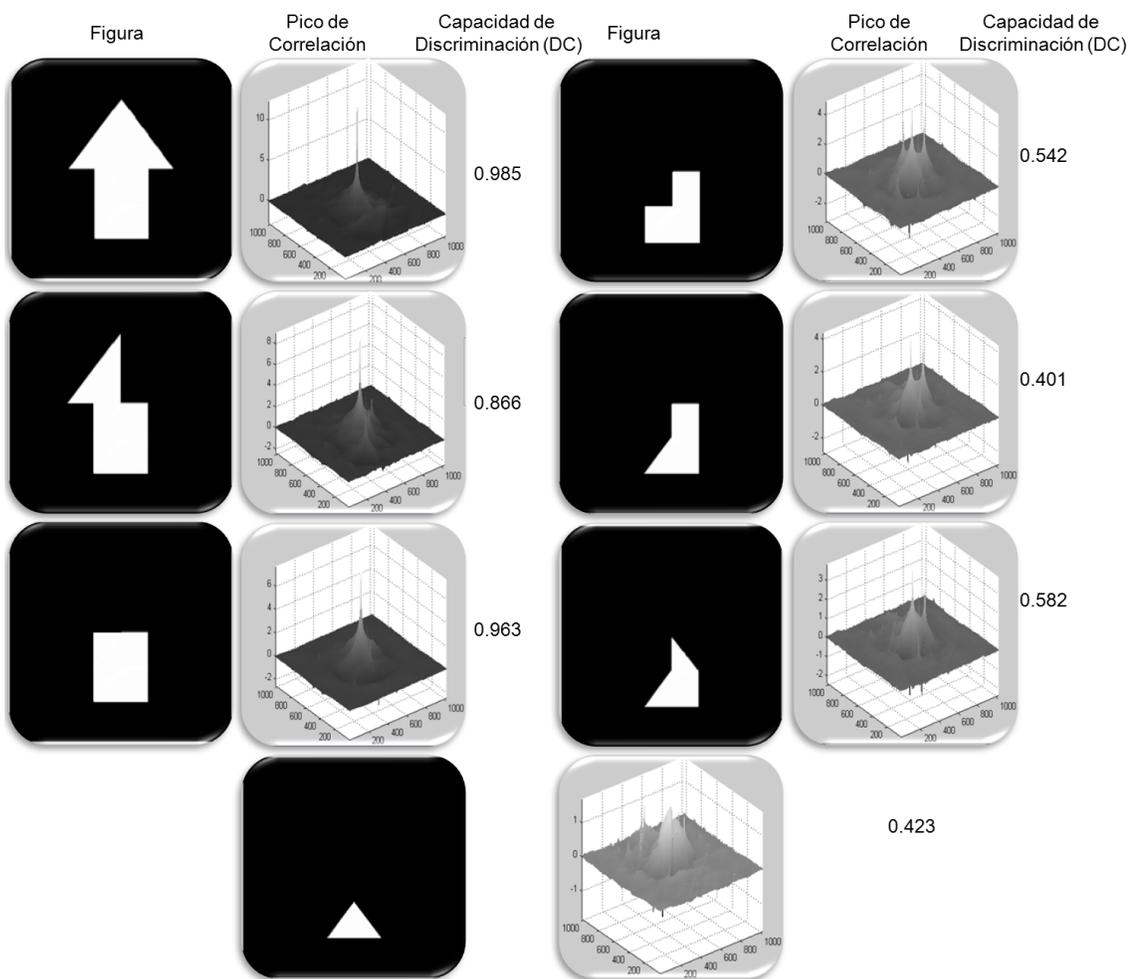


Figura 17.- Resultados de aplicar el filtro SDF de sólo fase a configuraciones parciales del Tangrama

En la figura 17 se muestra el comportamiento del filtro compuesto lineal de sólo fase ante configuraciones parciales de la figura. Como se puede apreciar, el filtro compuesto SDF de solo fase produce picos de correlación más agudos a comparación del filtro de acoplamiento, de tal forma que entre más piezas del tangrama contenga la figura, más agudo es el pico de correlación y mayor es el valor de DC. Lo ideal, o lo que

se espera de un filtro es que pueda reconocer una versión distorsionada de la imagen de referencia y rechazar cualquier otro, este es el enfoque de los filtros compuestos.

### V.3. Filtro compuesto no lineal k-law

Los filtros compuestos no lineales, como se verá a continuación, muestran notoriamente ventajas comparados con el filtro clásicos de acoplamiento, filtros de sólo fase, y otros filtros lineales ya que poseen una gran capacidad para discriminar objetos, producir un pico de correlación fácil de localizar y un plano de salida de correlación con menos ruido (Guerrero et al. 2009).

Como se mencionó en la sección IV.2.4, los filtros SDF son un tipo de filtros para reconocimiento de patrones invariantes a la iluminación, el cual se puede de expresar de la siguiente manera:

$$H = S(S^T S)^{-1}c, \quad (51)$$

donde  $S = [S_1, S_2, \dots, S_N]$ , es la matriz que contiene a las transformadas de Fourier de las  $N$  imágenes de entrenamiento  $S_i$  pero ordenadas lexicográficamente.

Si aplicamos la operación de no linealidad a  $S$  se obtiene como resultado el filtro SDF no lineal en el plano de Fourier:

$$H^k = S^k((S^k)^T S^k)^{-1}c \quad (52)$$

Si  $I_o(x, y)$  es la escena de entrada, entonces la salida de correlación está dada por la siguiente ecuación:

$$C(x, y) = \mathcal{F}^{-1}\{\mathcal{F}[H^k(x, y)]\mathcal{F}^*[I_o^k(x, y)]\} \quad (53)$$

El factor de no linealidad  $k$  con valores  $0 < k < 1$ , modificará la transformada de Fourier, tanto de nuestra imagen de referencia como de la escena de entrada. En la figura 18 se muestran las operaciones necesarias para obtener el plano de correlación un sistema no lineal de reconocimiento automático de objetos.

Como se puede apreciar, el proceso es similar al del filtro SDF de solo fase, la diferencia radica en el factor de no linealidad. Primero seleccionamos la figura con la que deseamos trabajar, misma que será nuestra imagen de referencia. Aplicamos la transformada de Fourier y la operación de no-linealidad a cada imagen de entrenamiento y creamos el filtro SDF con dichas imágenes. A la imagen de prueba le aplicamos también la transformada de Fourier y la operación de no-linealidad. Posteriormente realizamos la operación de correlación entre ambas imágenes.

Nuestro siguiente experimento consiste evaluar el filtro compuesto no lineal  $k$ -law, para lo cual se consideran como imagen de prueba y el conjunto de entrenamiento, las imágenes representadas en la figura 18. Utilizando un factor de no linealidad  $k=0.3$  el resultado se presenta en la figura 19.

Como se puede observar, este filtro resulta más eficaz que el filtro de acoplamiento y el filtro compuesto SDF de solo fase. Para comprobar la eficacia de este filtro se probó, con la misma figura pero con diferentes valores para el factor de no linealidad, el resultado se muestra en la gráfica de la figura 20:

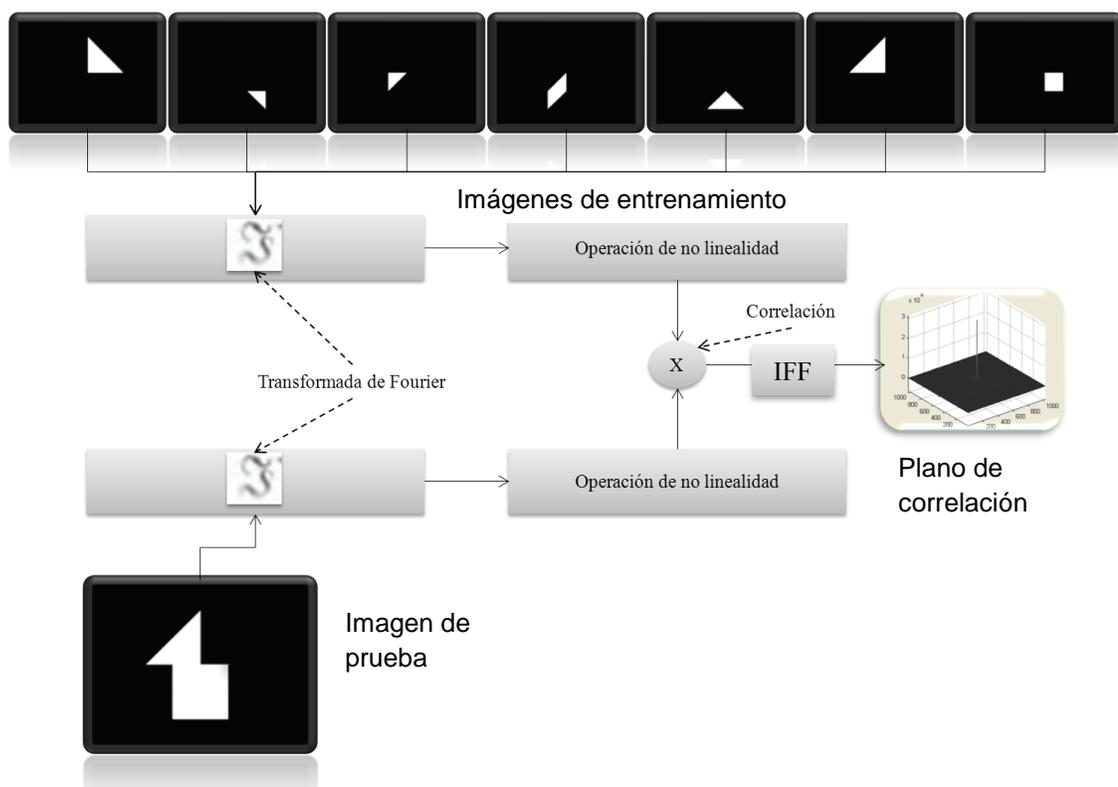


Figura 18. Operaciones necesarias para obtener el plano de correlación con el filtro compuesto no lineal k-law.

El filtro compuesto no lineal SDF de sólo fase k-law muestra un mejor desempeño que el filtro de sólo fase, produciendo un pico de correlación agudo y con mejor capacidad de discriminación con un factor de no linealidad con valores entre 0.1 y 0.4. En la figura 21 se puede apreciar que el desempeño de este filtro empieza a disminuir con un factor de no linealidad entre los valores 0.5 y 0.9, empiezan a aparecer los lóbulos laterales y a disminuir la capacidad de discriminación, es decir, tiende a comportarse como un filtro de acoplamiento.

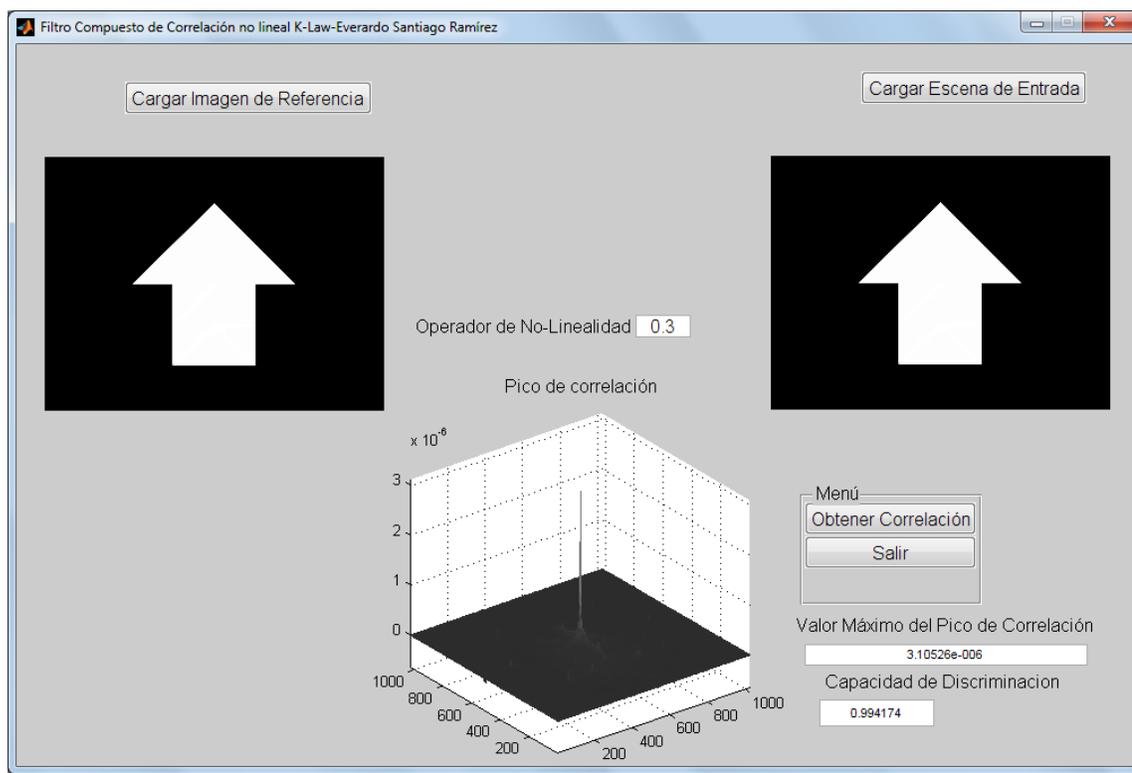


Figura 19.- Reconocimiento automático de objetos con el filtro compuesto no lineal SDF del tipo k-law

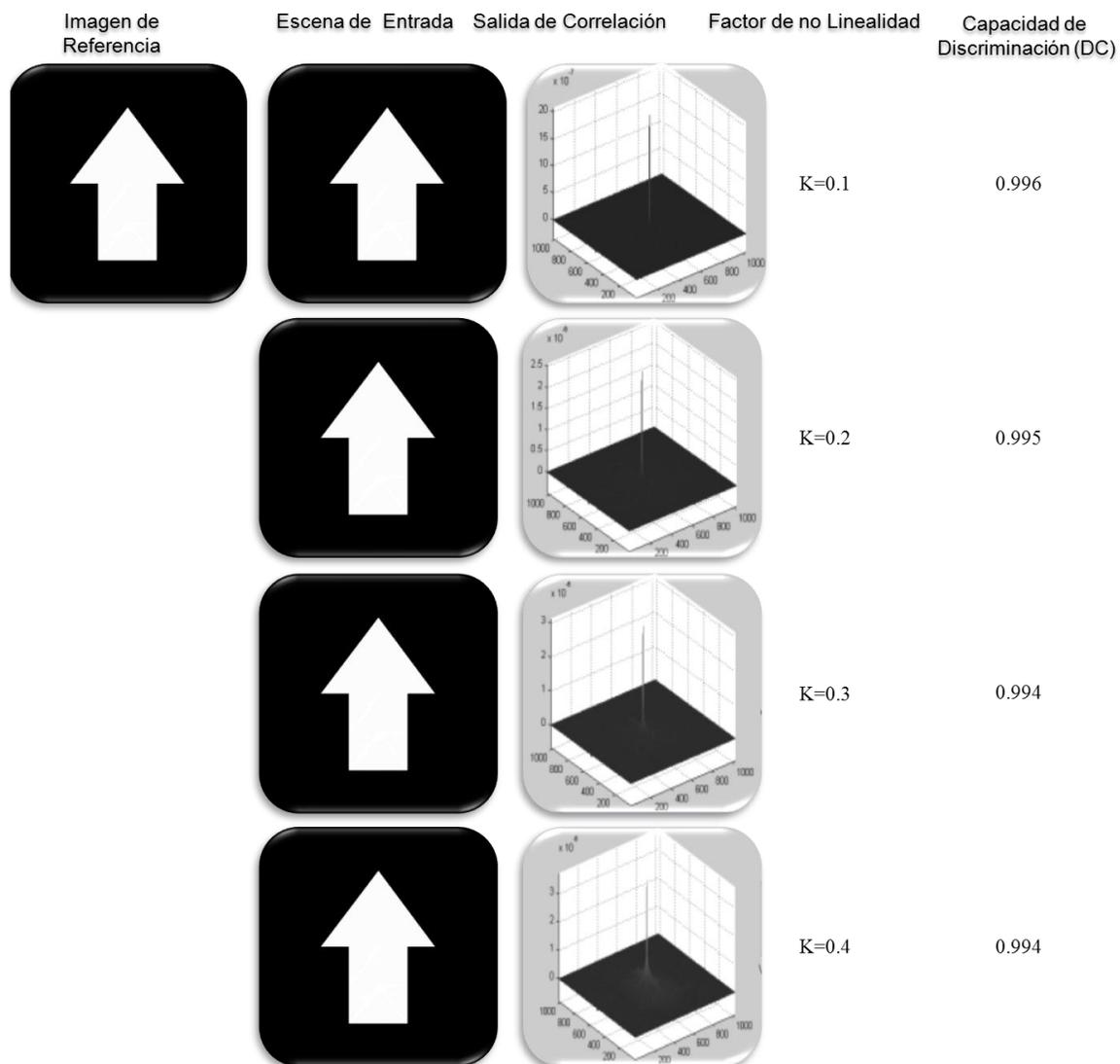


Figura 20.- Resultado de aplicar el filtro compuesto no lineal k-law con  $k=0.1, \dots, 0.4$ .

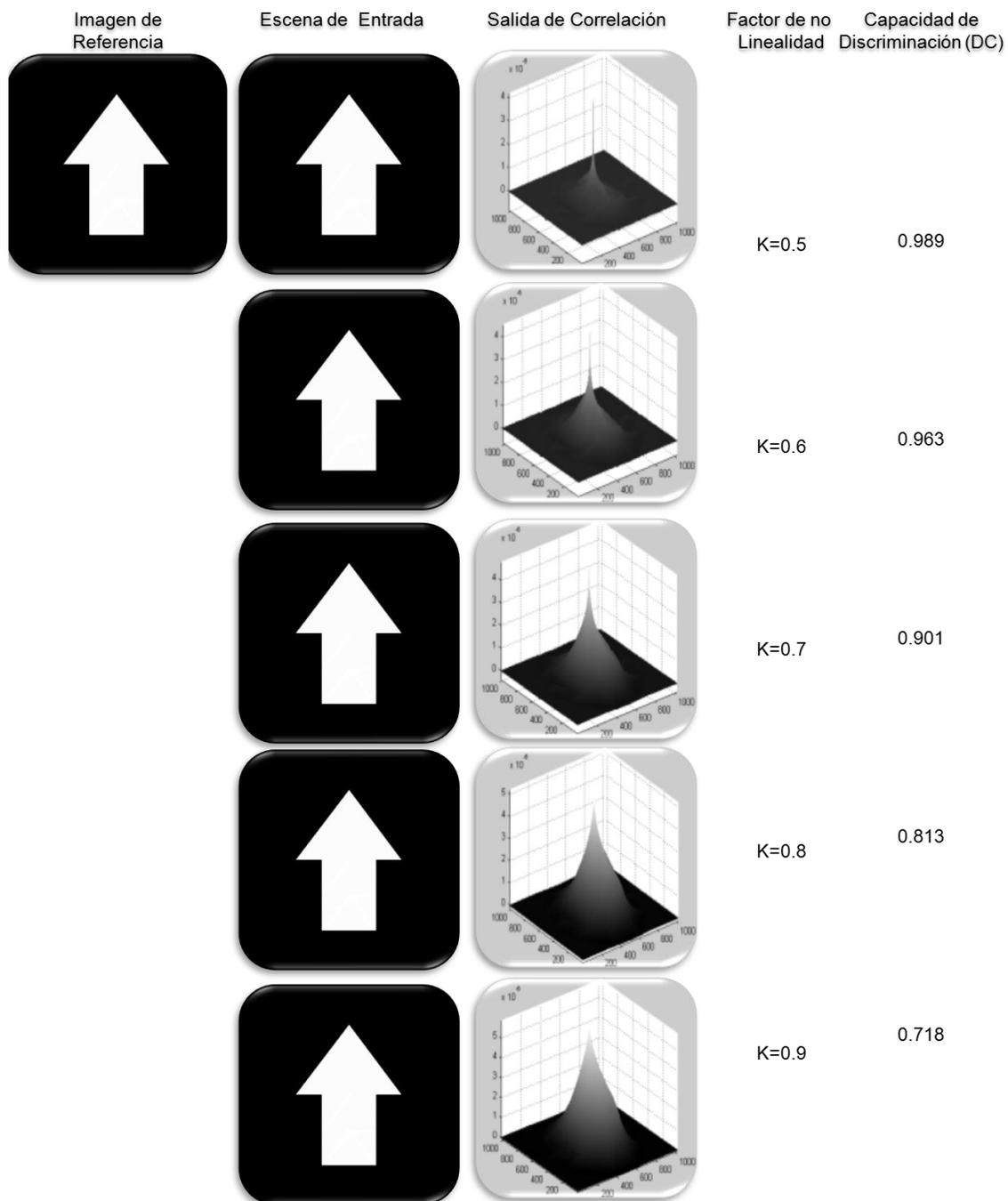


Figura 21.- Resultados de aplicar el filtro compuesto no lineal k-law con  $k=0.5, \dots, 0.9$ .

Los resultados producidos por cada filtro que se analizó en este capítulo son diferentes entre ellos. Los filtros de acoplamiento muestran una pobre capacidad de discriminación ya que producen un plano de correlación con lóbulos laterales que dificultan la localización del valor máximo de correlación. Los filtros compuestos no lineales SDF de solo fase producen mejores resultados que el filtro de acoplamiento, ya que producen un pico más agudo y en el origen, además de su invariancia a la iluminación. El filtro compuesto no lineal SDF de sólo fase k-law ofrece resultados parecidos a los del filtros compuesto SDF de solo fase, pero la ventaja que ofrece este filtro es la producción de picos más agudos ante configuraciones parciales de la figura y un capacidad de discriminación cercano a 1 con un factor de no linealidad con valores entre 0.1 y 0.4.

Para el caso de estudio que aquí se presenta, se decide implementar el filtro compuesto no lineal SDF de sólo fase k-law, debido a su robustez al permitir la creación de un sistema correlacionador invariante a distorsiones tales como las causadas por los cambios constantes de iluminación y/o defectos en el sensor de visión. El siguiente capítulo se presenta los resultados que produce este filtro en un sistema de visión artificial que supervisa el proceso de armado de un tangrama.

## **Capítulo VI**

### **Aplicación del filtro compuesto no lineal k-law en la supervisión del armado de un tangrama**

En el capítulo V se evaluaron tres filtros de correlación, los cuales fueron el filtro de acoplamiento (MF), filtro compuesto SDF de sólo fase y el filtro compuesto no lineal k-law. El objetivo de este capítulo es mostrar el desempeño del filtro compuesto no lineal k-law, aplicado a la problemática de supervisión del armado de un tangrama tangible.

Con el fin de lograr este objetivo, primero se presenta una pequeña descripción del uso de tangrama como herramientas de estimulación cognitiva y posteriormente un análisis del desempeño de este filtro no lineal con valores para el factor de no linealidad con valores entre 0.1 y 0.4, intervalo que mostró mejor capacidad de discriminación según los resultado de la evaluación (véase sección V.3).

#### **VI.1. Uso de tangrama como herramienta de estimulación cognitiva**

Estudios han revelado que los adultos que han acudido a sesiones de estimulación cognitiva y motora, presentan una situación afectiva más positiva. Además, presentan menos problemas conductuales y gozan de una calidad de vida superior, con

menos episodios de agitación, menor irritabilidad y mejor relación con otras personas (Rodríguez, 2009).

Hoy en día el tangrama no solo se usa como un entretenimiento sino en diferentes áreas del saber humano, tales como la psicología, el diseño, la filosofía y especialmente en la pedagogía. En el área de las enseñanzas de las matemáticas el tangrama se utiliza para introducir conceptos de geometría plana y para promover el desarrollo de las capacidades psicomotrices e intelectuales de los niños.

## **VI.2 Escenario de uso de la visión artificial durante el desarrollo de las actividades de estimulación cognitiva**

Imagínese el siguiente escenario: una casa para adultos mayores, la mayoría de ellos con problemas de deterioro cognitivo. Este problema provoca que el adulto pierda la noción del tiempo, su identidad, el placer de disfrutar su entorno y hasta el propósito de su vida.

En este lugar, las personas que se encargan de cuidarlos aplican una terapia basada en la resolución de rompecabezas para la estimulación cognitiva, la cual ayuda a los adultos mayores con problemas de deterioro cognitivo a mantener sus capacidades psicomotrices, posibilitándoles de esta forma una mayor autonomía y reflexión de sus actos cotidianos.

La terapia es buena y da resultado, pero hay un problema: Las personas que se encargan de cuidar a los adultos mayores deben estar al pendiente de 20 o más pacientes

al mismo tiempo y en ocasiones les es imposible estar atentos al progreso del armado del rompecabezas, mucho menos se dan cuenta del momento en que un adulto armó la figura con éxito. Las personas que cuidan a los ancianos no tienen la oportunidad de felicitar a la persona y a animarlo a seguir jugando, lo cual provoca que el adulto se sienta solo y sin atención.

Ahora imagínese en este mismo escenario pero ahora con un sistema de visión artificial con capacidad de supervisar de una manera automática el progreso de armado de un Tangrama e indicar el avance a los participantes sobre la superficie de armado. Además, dicho sistema es capaz de avisar mediante el uso de un dispositivo móvil al encargado de las personas sobre el éxito del jugador y así acudir a felicitarlo, dándole a saber que se está atento a él. El adulto se sentirá querido y estará animado a seguir jugando. Y todo de una forma interactiva con el apoyo de la visión artificial en un sistema consciente del contexto del juego.

### **VI.3. Selección de valor para el factor de no linealidad $k$**

El filtro seleccionado para resolver la problemática planteada en este trabajo es el filtro compuesto no lineal  $k$ -law, sin embargo debemos de analizar el comportamiento de este filtro con los diferentes valores de  $k$  que oscilan entre 0.1 y 0.4 con el fin de seleccionar el valor adecuado para usarse en el sistema.

A continuación se presenta el resultado de usar este filtro con el intervalo seleccionado y configuraciones parciales del tangrama. Para tal propósito considérese la

imagen de referencia y el conjunto de imágenes de entrenamiento representados en los diagramas de las figuras 15 y 18.

La figura 22 muestra el desempeño del filtro con un factor de no linealidad  $k=0.1$ , el pico de correlación cuando la figura está completa es aguda y con un coeficiente de discriminación muy cercana a 1, pero conforme se le restan piezas a la figura, este pico disminuye considerablemente haciéndola difícil de distinguir.

Con  $k=0.2$ , el filtro se comporta de manera similar que con  $k=0.1$ , se tiene un buen desempeño cuando la figura está completa pero va disminuyendo considerablemente, hasta el grado que es difícil de distinguir conforme se le restan piezas a la figura completa (ver figura 23).

En la figura 24 y 25 se presenta el resultado del filtro con los valores  $k=0.3$  y  $k=0.4$ , en donde el filtro ofrece su mejor desempeño para discriminar objetos, el factor de no linealidad con un valor de  $k=0.2$  (véase figura 23) ofrece un desempeño similar, pero con  $k=0.1$  (figura 22), aunque ofrece una mejor capacidad de discriminación cuando la figura está completa, los valores de DC con configuraciones parciales disminuye de forma considerable, ofreciendo una capacidad de discriminación menor que los otros valores.

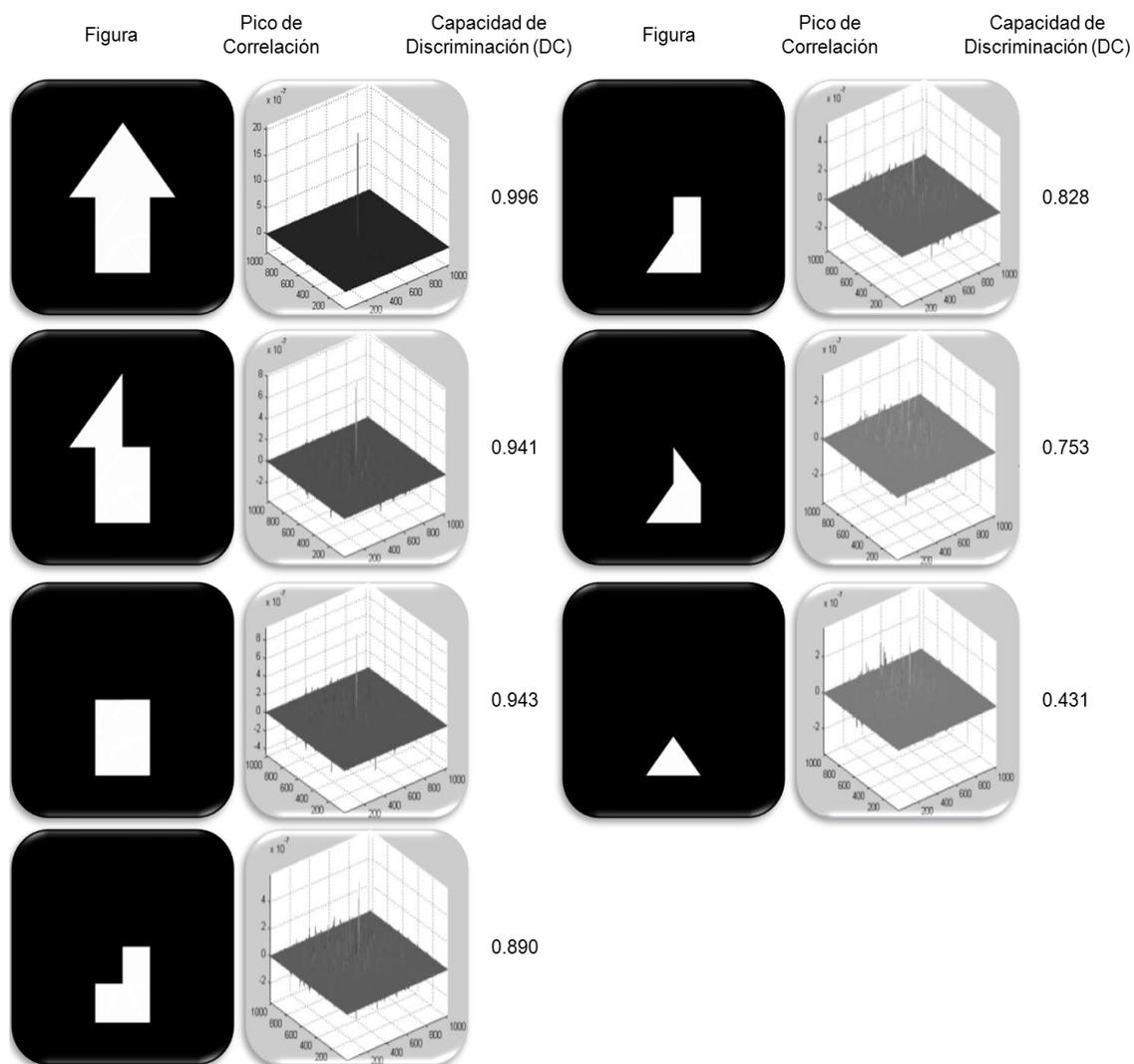


Figura 22.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con  $k=0.1$

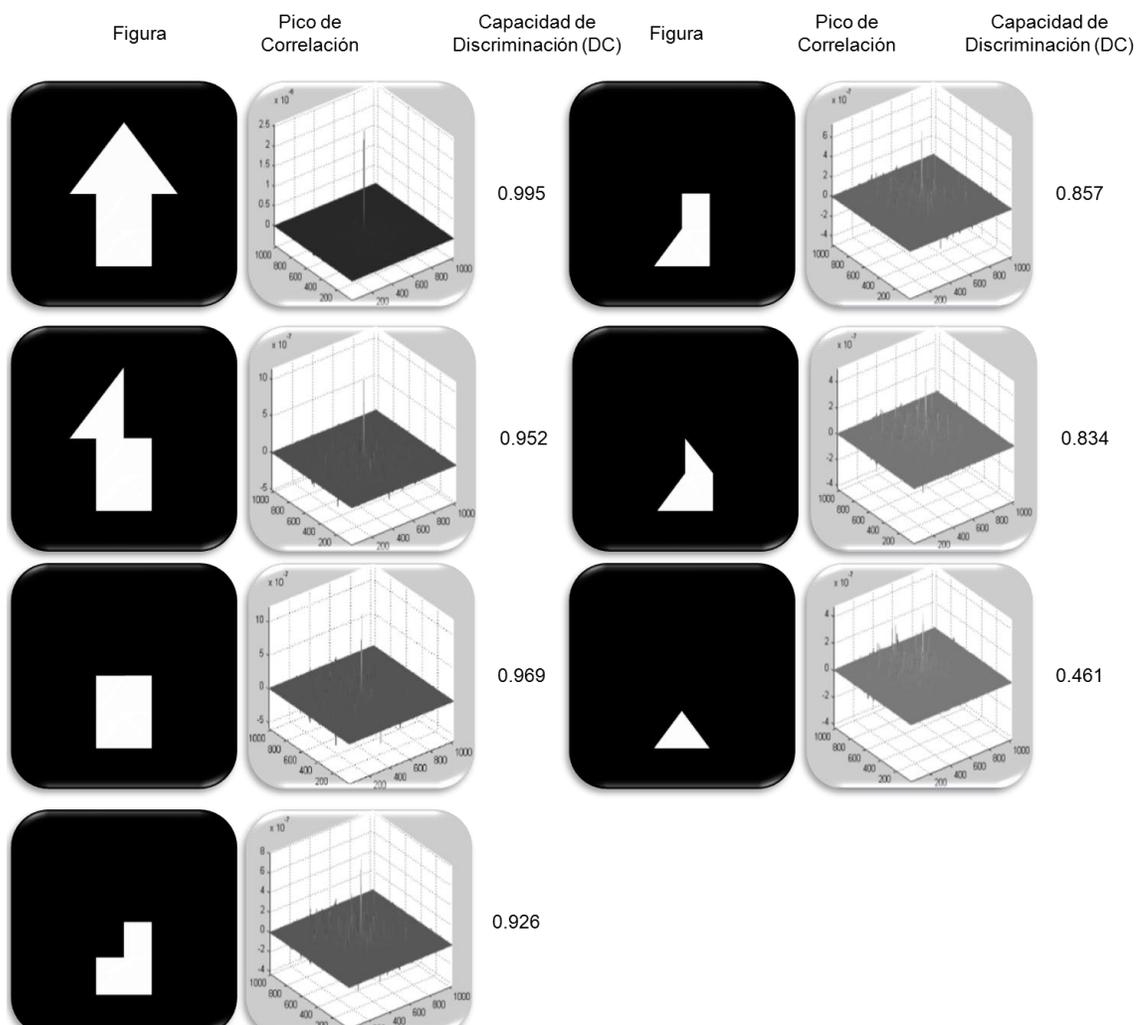


Figura 23.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con  $k=0.2$

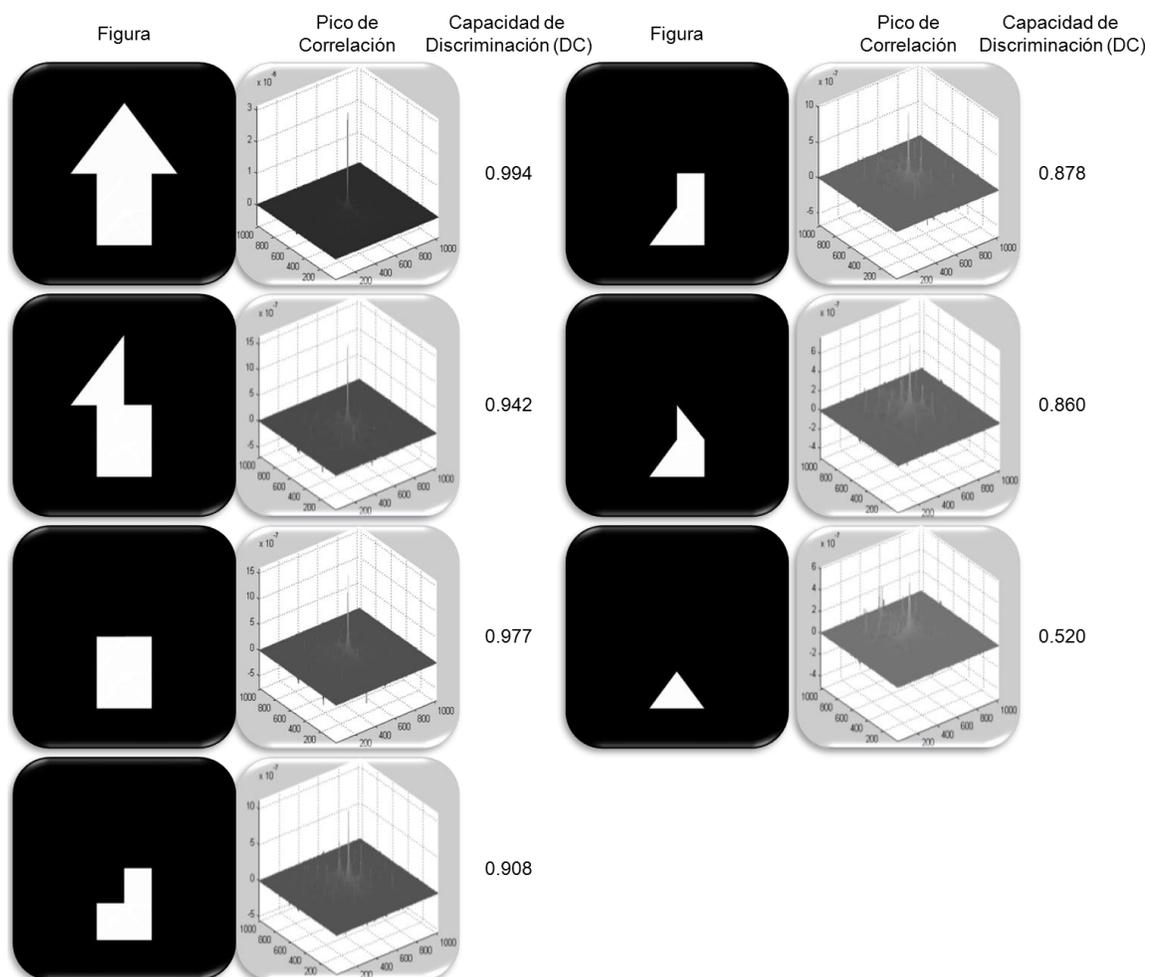


Figura 24.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con  $k=0.3$

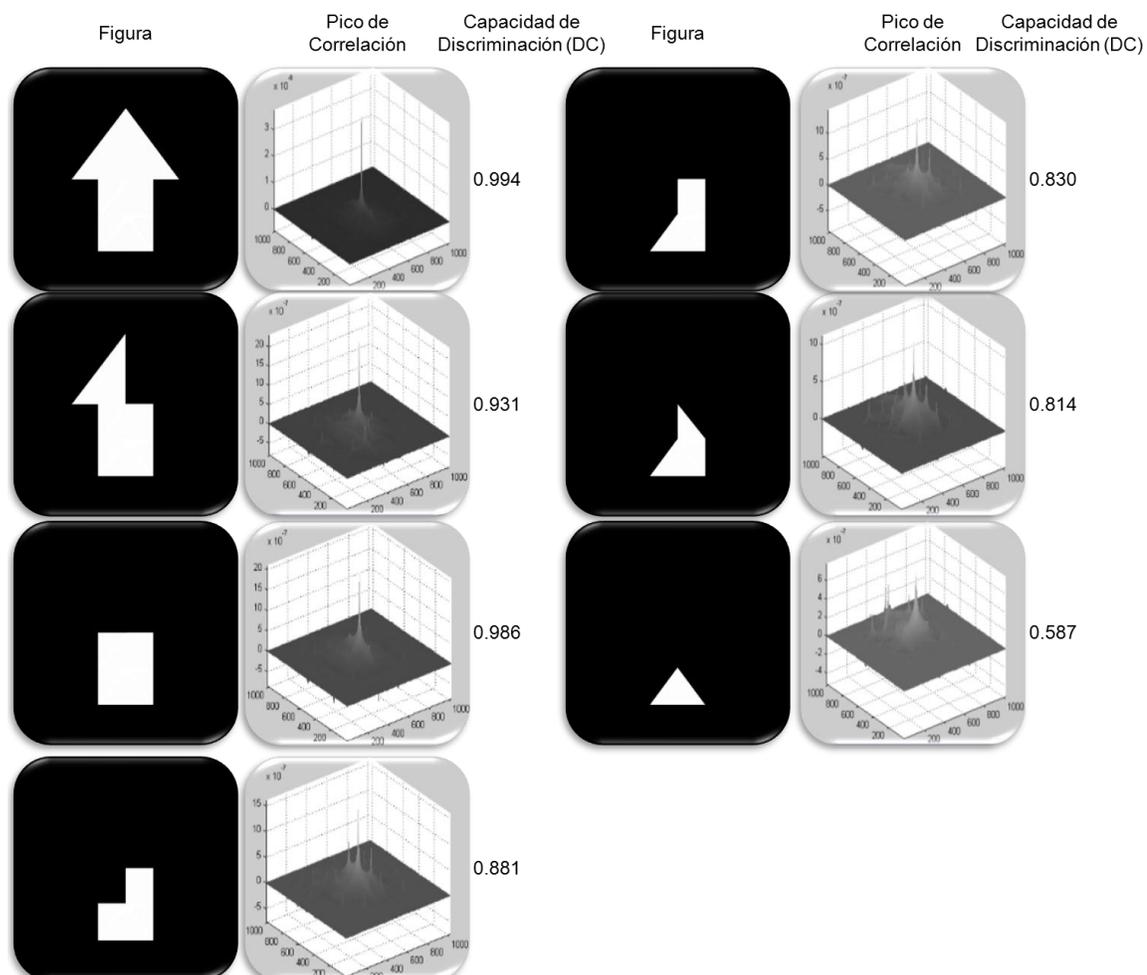


Figura 25.- Resultado de aplicar el filtro compuesto k-law a configuraciones parciales del tangrama con  $k=0.4$

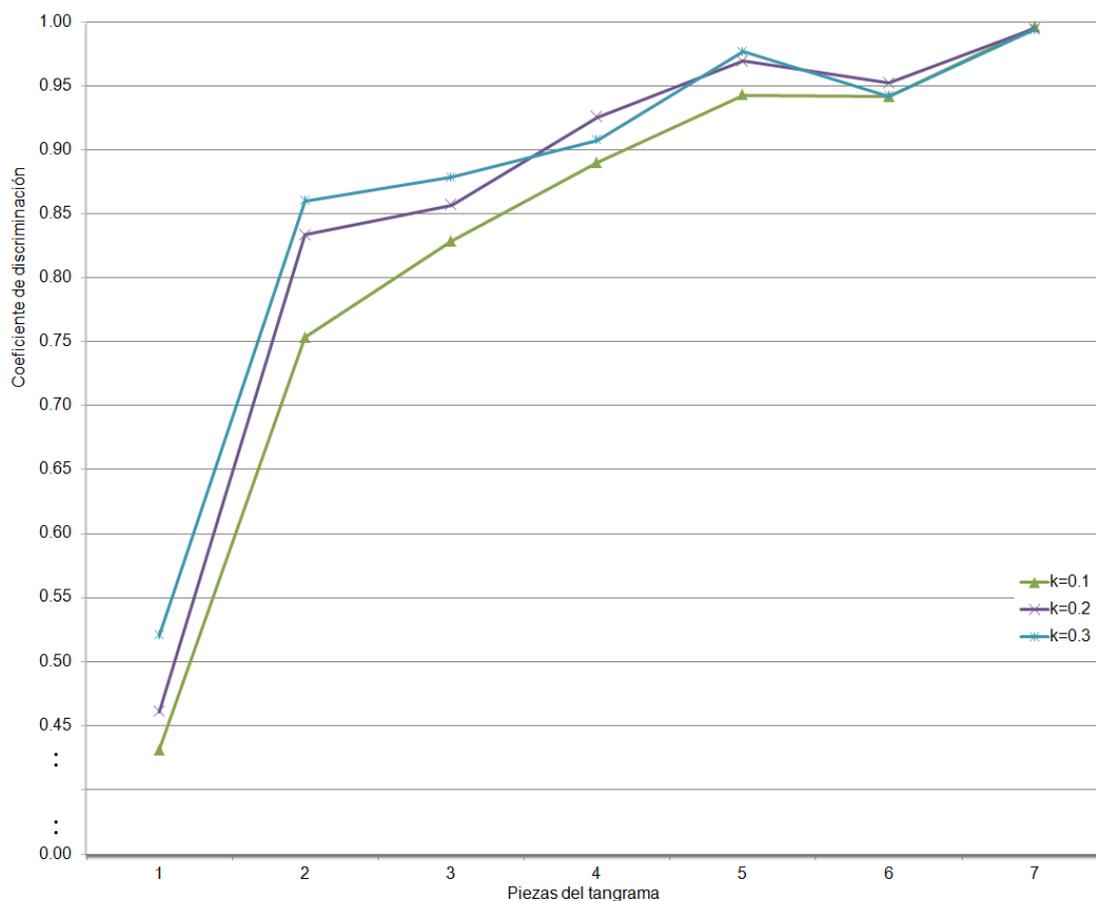


Figura 26.- Desempeño del filtro compuesto no-lineal k-law ante configuraciones parciales de una figura del tangrama

En la figura 26 se muestra el desempeño del filtro con valores para el factor de no linealidad con valores entre 0.1 y 0.4. Como se puede observar. Aunque el desempeño del filtro no varía mucho en ese intervalo, ya que cuando la figura está completa el coeficiente de discriminación se aproxima mucho a 1, pero con las figuras parciales el factor de no linealidad con los valores  $k = \{0.1, 0.2, 0.4\}$  la capacidad de discriminación se reduce considerablemente, siendo el factor de no linealidad  $k = 0.3$

quien nos ofrece resultados más uniformes. Esto se deduce del hecho de que  $k = 0.3$  ofrece una media aritmética de coeficientes de discriminación igual 0.869, mientras que  $k = \{0.1, 0.2, 0.4\}$  ofrecen una media aritmética de 0.826, 0.856 y 0.860, respectivamente.

En la gráfica presentada en la figura 27 se puede apreciar que el filtro compuesto no lineal con un factor de no linealidad de  $k = 0.3$  ofrece una capacidad de discriminación mayor que el ofrecido por el filtro compuesto SDF de solo fase y el filtro de acoplamiento frente a configuraciones parciales del tangrama.

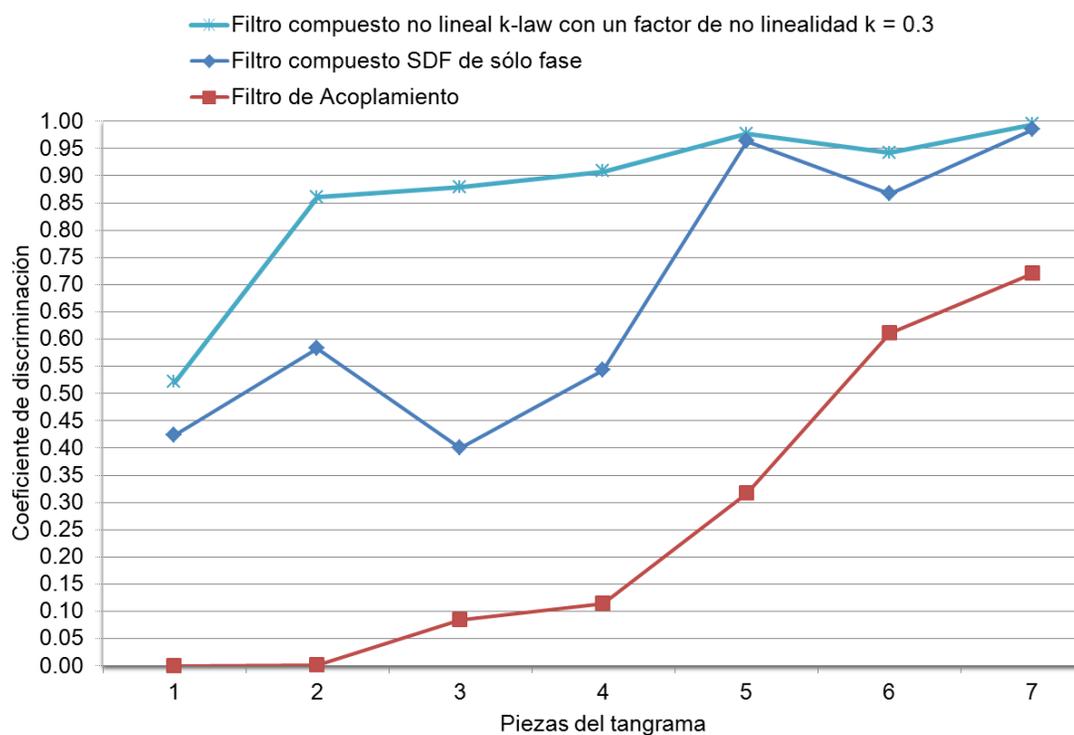


Figura 27.- Desempeño del filtro compuesto no lineal k-law con un factor de no linealidad  $k=0.3$  ante el filtro de sólo fase y el filtro de acoplamiento

El factor de no linealidad  $k=0.3$  ofrece una mayor capacidad de discriminación a los filtros, por tal motivo se selecciona este valor para usarlo en la implementación del sistema de visión artificial para la supervisión del armado de un tangrama tangible (véase el apéndice B), el cual se presenta en la siguiente sección.

#### **VI.4. Sistema de visión artificial para la supervisión del armado de un tangrama**

El sistema de visión artificial para la supervisión del proceso de armado de un tangrama tangible, denominado desde ahora en adelante como SISVIASAT, implementa el filtro compuesto no lineal SDF de sólo fase  $k$ -law para obtener el grado de similitud entre la imagen de referencia y la figura que arma el usuario. Para probar el funcionamiento de este sistema, se consideraron las imágenes de referencia en la figura 28.

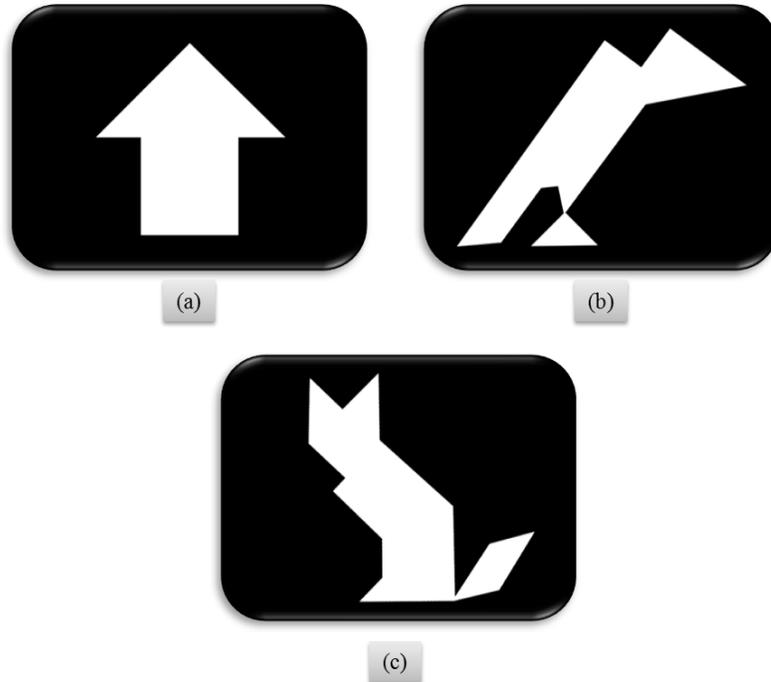


Figura 28.- Imágenes de referencia a probar con SISVASAT

El algoritmo para trabajar con SISVIASAT se resume de la siguiente manera (refiérase al apéndice C para el manual de usuario):

- El usuario selecciona la imagen referencia con la que desea trabajar.
- El sistema crea de forma automática un filtro SDF con las piezas del tangrama, siendo estos el conjunto de entrenamiento. Para cada imagen del conjunto de entrenamiento, se le aplica el factor de no linealidad.
- El usuario ejecuta el sistema.
- El sistema empieza a capturar la escena de entrada y a evaluarlo para calcular el valor de correlación cada cinco segundos.

- El sistema calcula el progreso del usuario de acuerdo al valor máximo de correlación para la figura seleccionada cuando la figura está completamente armada. El progreso se muestra en el indicador gráfico.

Considere configuración del tangrama en la figura 29 como la imagen de referencia, es decir, figura la que deseamos llegar con las siete piezas del rompecabezas. Una configuración completa realizada por el usuario es evaluada de forma correcta por el SISVIASAT, esto se puede ver en la siguiente figura en donde la barra del indicador está en el tope, al 100%.

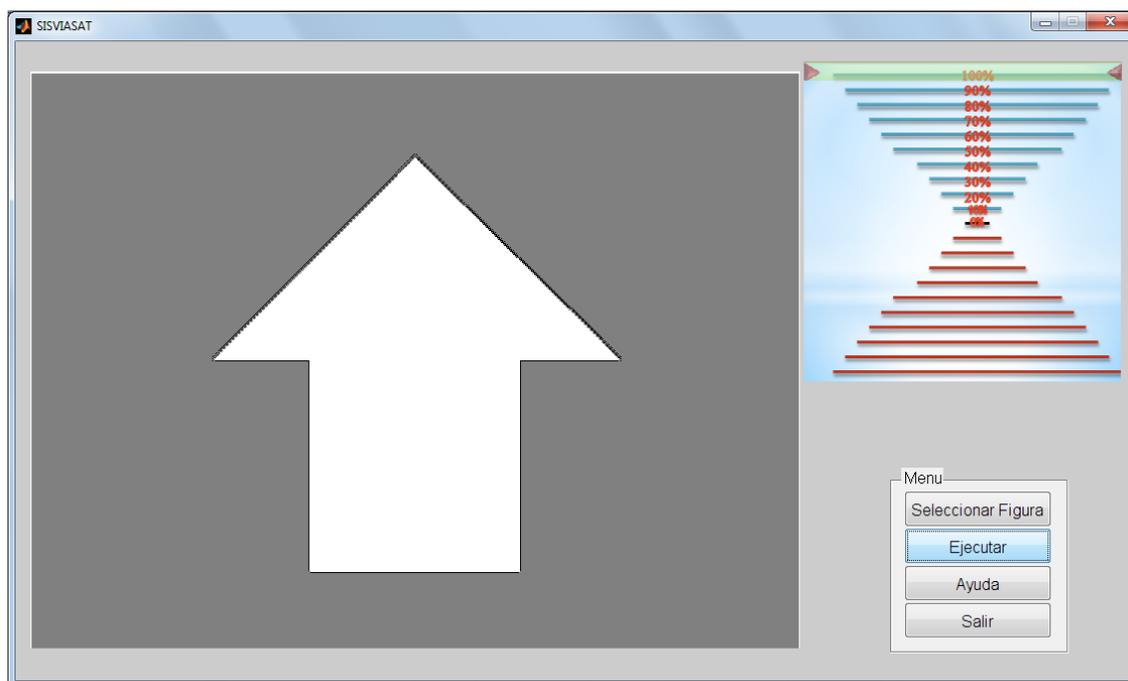


Figura 29.- Figura completa de tangrama en SISVIASAT

El filtro ofrece una buena capacidad de discriminación en un ambiente real donde las distorsiones provocadas por la iluminación están presentes. En este caso el filtro generó un plano de correlación con un pico agudo y fácil de localizar con un coeficiente de discriminación igual a 0.983, un valor muy cercano a 1 (véase figura 30).

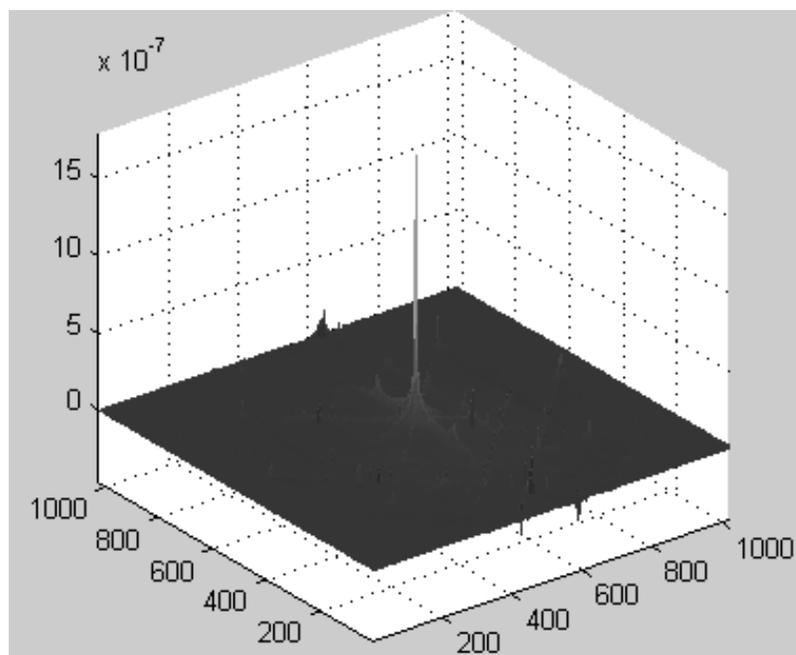


Figura 30.- Plano de correlación de una figura realizada en SISVIASAT con un DC = 0.983

En las figuras 31 y 32 se muestra el resultado producido por el sistema ante configuraciones parciales del tangrama, como se puede observar cada pieza contribuye con un porcentaje a la configuración final.

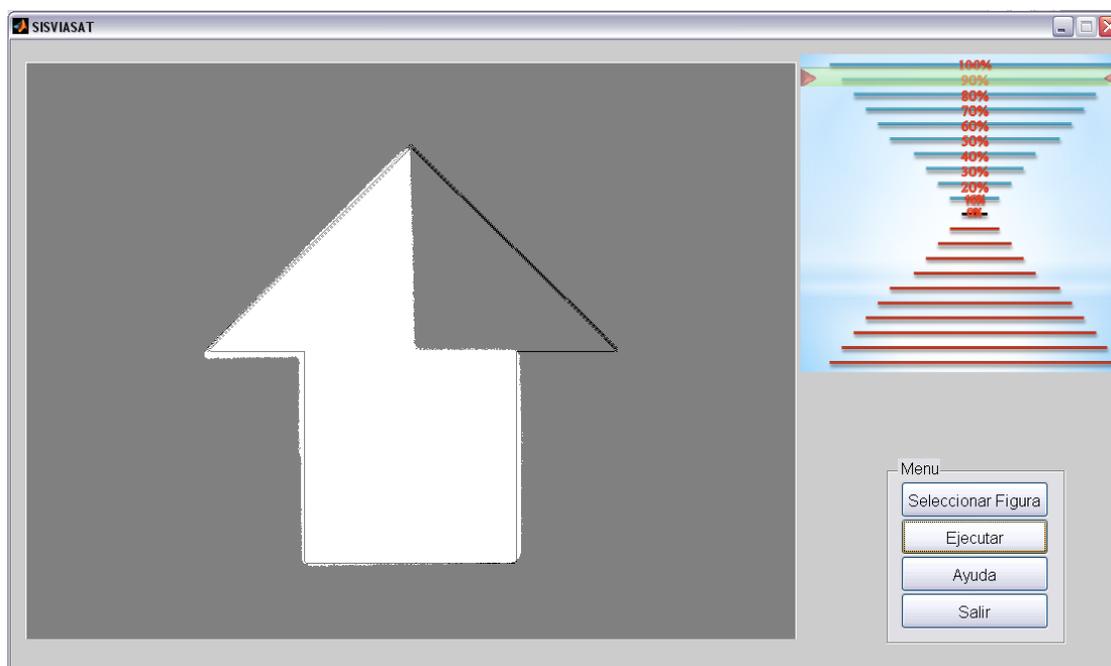


Figura 31.- Resultado de una configuración parcial de 6 piezas del tangrama con SISVIASAT.

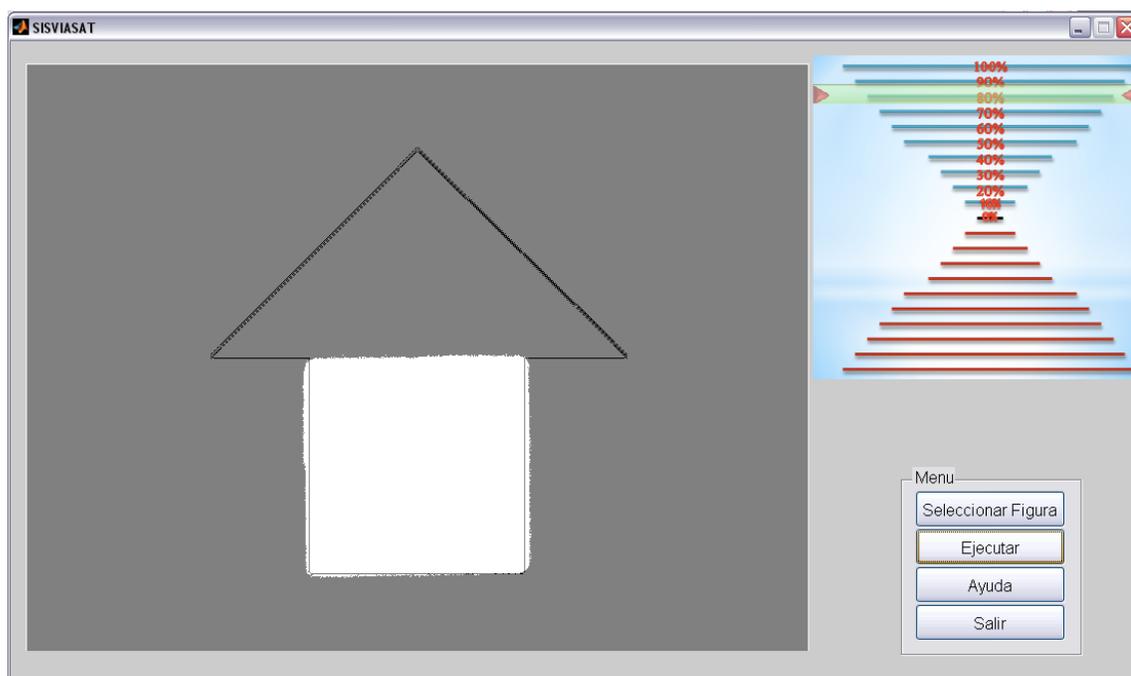


Figura 32.- Resultado de una configuración parcial 5 piezas del tangrama con SISVIASAT.

En la figura 33 se muestran los resultados con las figuras 28(b) y 28(c), se puede observar que el sistema se desempeña correctamente con dichas figuras al reconocer que están completas al 100%.

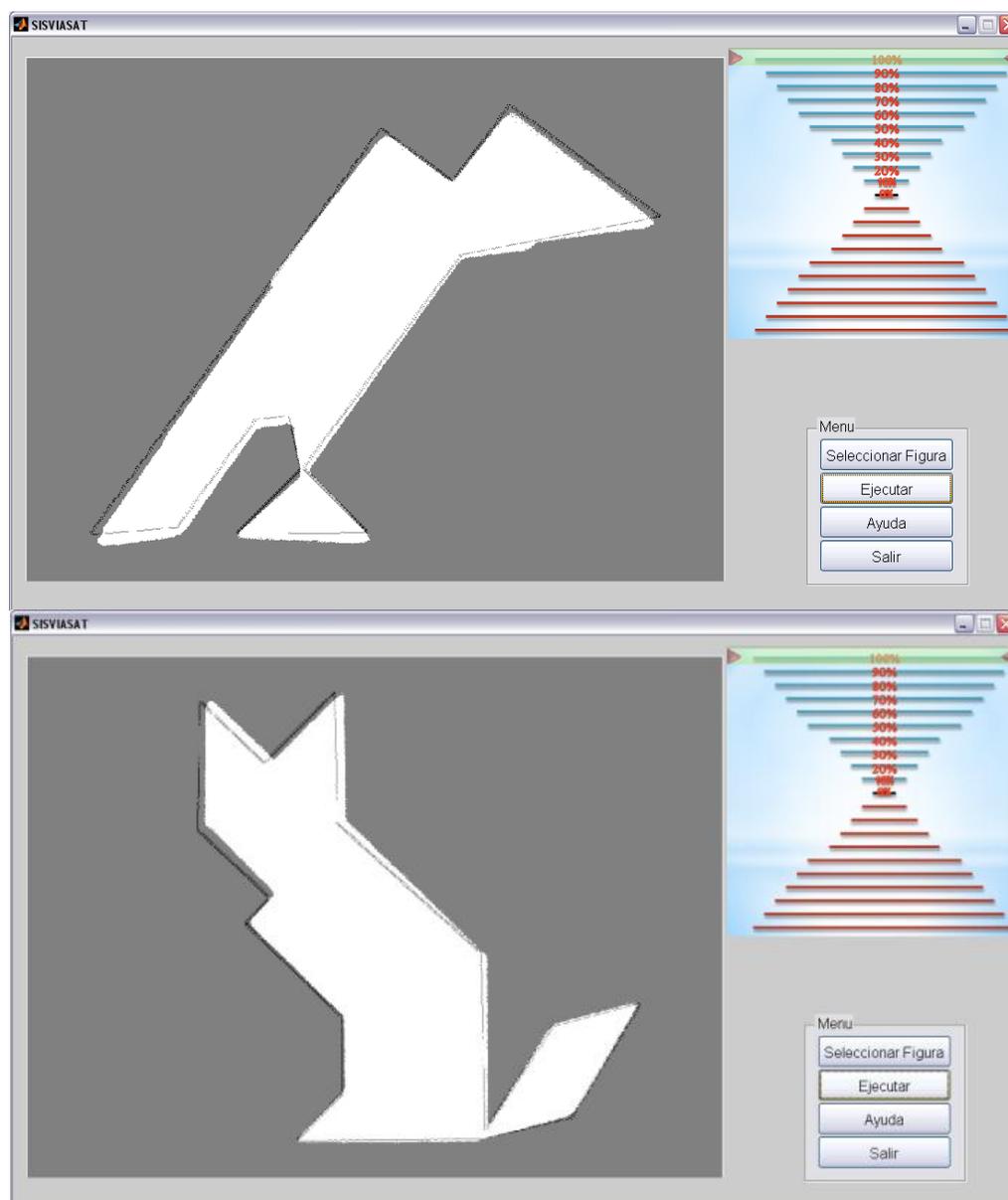


Figura 33.- Resultado al armar las figuras 25(b) y 25(c) en SISVIASAT.

El sistema ofrece un buen desempeño usando el filtro compuesto no lineal k-law al producir coeficientes de discriminación muy cercanos a 1 en el caso donde la figura contiene las 7 piezas del tangrama, claramente se puede observar que entre menos piezas contenga la figura armada por el usuario, menor es la capacidad de discriminación del sistema. En la figura 34 se presenta la gráfica de desempeño de SISVIASAT usando las figuras 28(a-c) como imágenes de referencia, produciendo una media aritmética de coeficientes de discriminación de 0.861, 0.906 y 0.904, respectivamente.

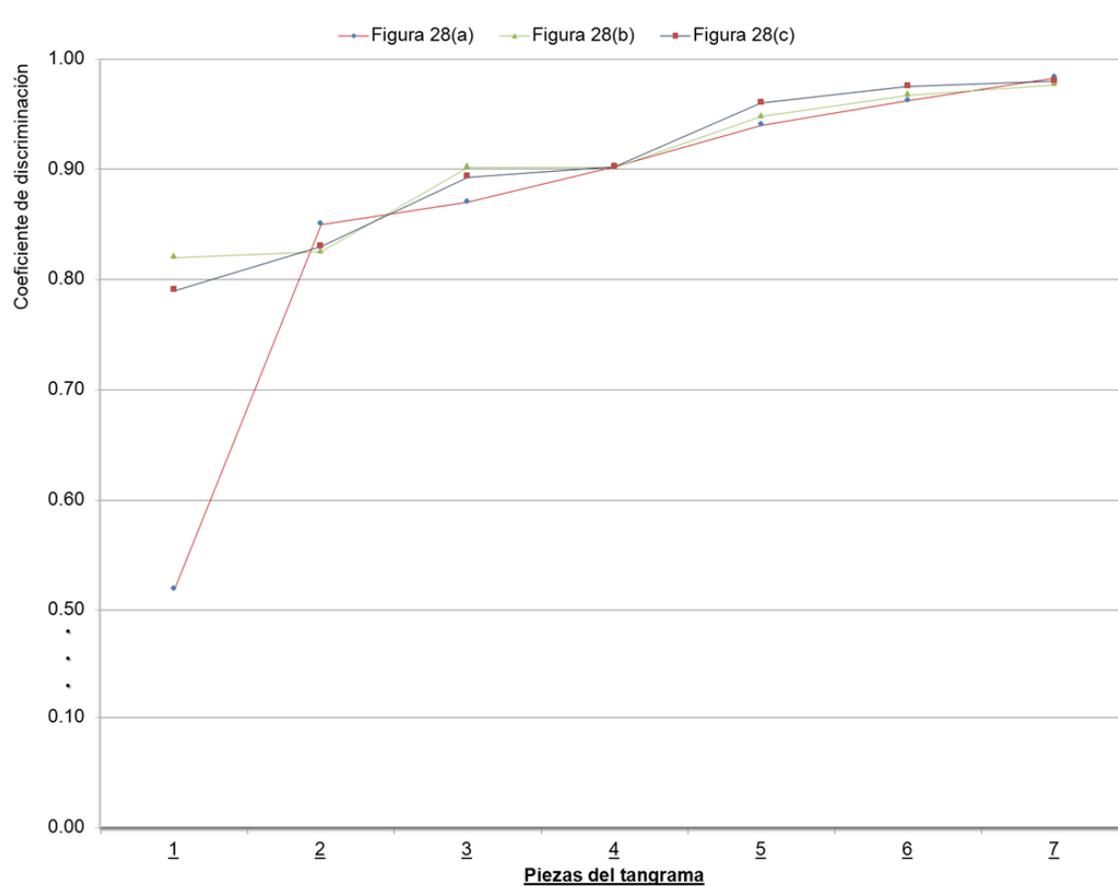


Figura 34.- Desempeño de SISVIASAT con el filtro compuesto no lineal k-law

## **Capítulo VII**

### **Conclusiones y trabajo futuro**

En este trabajo de tesis se analizó la teoría existente sobre el reconocimiento de patrones por correlación y la visión artificial. Se diseñó un sistema de visión artificial para la supervisión del armado de un tangrama como caso de estudio.

Durante el desarrollo de este trabajo se analizaron los filtros de correlación más importantes, y entre ellos se distinguieron aquellos que ofrecían una robustez ante distorsiones. De estos filtros se seleccionaron tres para experimentar con ellos usando imágenes creadas por computadora. Posteriormente se seleccionó uno para implementarlo en un sistema de visión artificial: SISVIASAT.

Para decidir cuál de los tres filtros se habría de implementar en SISVIASAT, se desarrolló un conjunto de aplicaciones que pueden ser usadas en cursos de visión artificial, procesamiento de imágenes y reconocimiento de patrones para mostrar el uso de los filtros de correlación. El desarrollo de estas aplicaciones nos permitió identificar las ventajas y desventajas de cada filtro de correlación en el reconocimiento automático de objetos.

Se diseñó y desarrolló, como caso de estudio, un sistema de visión artificial que supervisa el armado de un tangrama. El sistema SISVIASAT se diseñó para ser usado en terapias de estimulación cognitiva, en apoyo a personas con enfermedades degenerativas que ayudara a estimular y mantener sus capacidades cognitivas. Pero el sistema no se

limita solo en eso, sino también puede ser usado en centros educativos para la estimulación de las capacidades psicomotrices de los niños.

El logro más significativo es el material teórico/práctico que se ha generado al hacer este trabajo de tesis, el cual puede ser usado para otros proyectos de investigación y/o en el desarrollo de sistemas que resuelvan problemas de la vida real.

## **VII.1. Conclusiones**

Las conclusiones a que se llega en base a los resultados de los experimentos se mencionan a continuación:

- De los filtros lineales, el filtro de acoplamiento mostró una pobre capacidad de discriminación al producir un pico de correlación amplio.
- El filtro compuesto de sólo fase, al usar una combinación lineal de imágenes de entrenamiento mostró una mejor capacidad de discriminación buen desempeño al producir un pico de correlación agudo.
- El filtro compuesto no lineal k-law, en base a la métrica de “Capacidad de Discriminación”, mostró un mejor desempeño que el filtro clásico de acoplamiento y el filtro de compuesto SDF de sólo fase al recibir como parámetro configuraciones parciales de la imagen de referencia, lo cual lo hizo un buen candidato para ser implementado en el sistema de visión artificial.
- Usar filtros de correlación con imágenes de entrenamiento en el reconocimiento automático de objetos produce un mejor resultado que aquellos filtros que usan solo una imagen como plantilla.

- La iluminación afecta seriamente a los sistemas de visión artificial al distorsionar el contorno de los objetos en la escena de entrada, por lo cual es necesario el desarrollo de filtros correlacionadores que ofrezcan invariancia a la iluminación ya que las condiciones normales de uso del sistema no aseguran un control absoluto en las condiciones de iluminación.
- Una buena selección del conjunto de entrenamiento permite reconocer versiones distorsionadas de la imagen objetivo, por lo cual esta actividad requiere de especial atención.
- Las dimensiones de las imágenes entrenamiento y una correcta selección de imágenes de entrenamiento reduce los requerimientos de almacenamiento y el costo computacional se optimiza notablemente.
- En el diseño de un sistema de visión artificial se deben de considerar la iluminación, la actividad a realizar, el sensor de visión, el dispositivo de procesamiento, entre otros para obtener un mejor resultado.

## **VII.2. Trabajo futuro**

El objetivo de este trabajo fue realizar un análisis de los filtros de correlación aplicado en un caso de estudio, el cual consistió en un sistema de visión artificial para la supervisión del armado de un tangrama. Para probar si este sistema apoya realmente a las terapias de estimulación cognitiva, es necesario realizar una evaluación, el cual mida el desempeño del sistema y su aceptación por parte de los usuarios finales.

SISIVIASAT puede ser utilizado en terapias de estimulación cognitiva para adultos mayores con deterioro cognitivo. Dicho sistema también puede ser ampliado, dándole soporte a dispositivos móviles de tal forma que se una herramienta de apoyo a los cuidadores de estas personas, por ejemplo, que cada cierto tiempo le esté avisando al cuidador sobre el progreso del adulto.

## Referencias

- A. Mahalanobis, B. V. K. Vijaya Kumar y D. Casasent. 1987. Minimum Average Correlation Energy Filters. *Applied Optics*, 26. 3633-3640 pp.
- B. Braunecker, R. W. Hauck y A. W. Lohmann. 1979. Optical Character Recognition Based on non-Redundant Correlation Measurements. *Applied Optics*, 18.2746-2753 pp.
- B. V. K. Vijaya kumar. 2006. Face Recognition Using Correlation Filters. Data Storage Systems Center (DSSC). Carnegie Mellon University, Pittsburgh, USA. 1-2 pp.
- B. V. K. Vijaya kumar y Savvides Mario. 2006. Correlation Pattern Recognition for Biometrics. SPIE Newsroom, The International Society for Optical Engineering. 1-3 pp.
- B. V. K. Vijaya Kumar. 1986. Minimum Variance Synthetic Discriminant Functions. *J. Opt. Soc. Am. A* 3.1579-1584 pp.
- Bebis George, Egbert Dwight y Shah Mubarak. 2003. Review of Computer Vision Education. *IEEE Transactions on Education*. 1-19 pp.
- Bolme David y Draper Bruce. 2009. ASEF Correlation Filters for Signal Processing and Object Detection and Recognition. Colorado State University. 1 pp.
- C. F. Hester y D. Casasent. 1980. Multivariant Techniques for Multiclass Pattern Recognition. *Applied Optics*, 19.1758-1761 pp.
- D. Jared y D. Ennis. 1989. Inclusion of Filter Modulation in Synthetic Discriminant Function Construction. *Applied Optics* 28.232-239 pp.

Feng Shaotong, Wang Liang y NieShouping. 2008. Rotation Tolerant Three-Dimensional Object Pattern Recognition Using a Single Exposure off-axis Digital Holography. Lasers in Engineering. Nanjin, China. 1-8 pp.

González Fraga José Ángel, Kober Vitaly y Álvarez Borrego Josué. 2006. Adaptive Synthetic Discriminant Function Filters for Pattern Recognition. Optical Engineering. Bllingham, Washington, USA. 057005-1 - 057005-10 pp.

Gudmundsson Karl y A. S. Awwal Abdul. 2007. POF The Amplitude Feature Relationship. International Workshop on Information Optics, American Institute of Physics.280-285 pp.

Guerrero Moreno Enrique Ricardo y Alvares Borrego Josué. 2009. Non-linear Composite Filter's Performance. Optical Engineering, vol. 48. 067201-027201-11 pp.

H. J. Caulfield y W. T. Maloney.1969. Improved Discrimination in Optical Character Recognition. Applied Optic, 8.2354-2356 pp.

J. L. Horner y P. D. Gianino. 1984. Phase-Only Matched Filtering. Applied Optics. 23. 812-816 pp.

J. L. Horner y P. D. Gianino. 1985. Applying the Phase Only Filter Concept to Synthetic Discriminant Function Correlation Filter. Applied Optics 24. 851-855 pp.

J. R. Leger y S. H. Lee. 1982. Hibrid Optical Processor for Pattern Recognition and Classification Using a Generalized set of Pattern Functions. *Applied Optics* 21.274-287 pp.

Jain Anil K., Duin Robert P. W. y Mao Jianchang. 2000. Statistical Pattern Recognition: A Review. *IEEE Transactionson Pattern Analysis and Machine Intelligence*.4-37 pp.

L. Hassebrook, B. V. K. VijayaKumar y L. Hostetler. 1990. Linear Phase Coefficient Composite Filter Banks for Distortion Invariant Optical Pattern Recognition. *Optical Engineering*. 29.1033-1043 pp.

Lavine Barry K. 2006. Pattern Recognition. *Critical Reviews in Analitical Chemistry*, Taylor and Francis.153-170 pp.

Pajares Martinsanz Gonzalo y De La Cruz García Jesús M. 2008. *Visión por computador: Imágenes digitales y aplicaciones*. Segunda Edición. Editorial RA-MA. Madrid, España.

Pérez Elisabet y Javidi Bahram. 2002. Non-linear Distortion-Tolerants Filters for Detection of Road Signs in Background Noise. *IEEE Transactions on Vehicular Technology*. 567-576 pp.

Pérez Elisabet y Javidi Bahram. 2005. Scale and Rotation Invariant Optical ID Tags for Automatic Vehicle Identification and Authentication. *IEEE Transactions on Vehicular Technology*.1295-1303 pp.

R. R. Kallman. 1987. Direct Construction of Phase Only Filters. *Applied Optic*, 26.5200-5201 pp.

Rodriguez, B. C. (2009). *infogerontologia, estimulacion cognitiva*. Recuperado el 01 de Noviembre de 2009, de *infogerontologia*: <http://www.infogerontologia.com/index.html>.

Santiago Ramírez Everardo y González Fraga José Ángel. 2008. Reconocimiento de Objetos Mediante Filtro Compuesto SDF de Solo Fase. *Memorias del V Taller de Procesamiento de Imágenes*. Guanajuato, Guanajuato.

Shaotong Feng, Liang Wang y Shouping Nie. 2008. Rotation Tolerant Three-Dimensional Object. *Lasers in Engineering*, Old City Publishing. 1-8 pp

Song Andy y Ciesielski Vic. 2001. *Vision Systems Development by Machine Learning: Mashing Assessment in Brewing*. *Applied Artificial Intelligence*, Taylor & Francis. Department of Computer Science, Melbourne Victoria, Australia. 777-794 pp.

Vijaya Kumar B. V. 1992. Tutorial Survey of Composite Filter Designs for Optical Correlators. *Optical Society of America*. 4773-4799 pp.

Vijaya Kumar B. V., Mahalanobis Abhijit y Juday Richard. 2005. *Correlation Pattern Recognition*. Cambridge University Press. New York, USA.

## Apéndice A

### Tablas comparativas de los filtros

A continuación se presentan dos tablas comparativas de algunos filtros de correlación, los cuales han sido propuestos en la literatura especializada y que han causado mayor interés en el área de visión artificial.

Tabla I.- Filtros básicos de correlación: criterios de maximización y desventajas

Filtros básicos de correlación			
Filtro	Función de transferencia	Criterio que maximiza	Desventaja
Filtro de acoplamiento	$H(f) = \alpha \frac{S^*(f)}{P_n(f)}$	Este filtro maximiza la relación señal a ruido (SNR)	Poca eficiencia en luz y poca capacidad de discriminación
Filtro de solo fase	$H_{POF}(u, v) = \frac{S^*(u, v)}{ S(u, v) } = e^{-j\theta(u, v)}$	Mejoran la eficiencia de procesamiento de luz	
Filtro de solo fase binario	$H_{BPOF}(f) = \text{sgn}[\text{Re}\{e^{-j\theta} S(f)\}] = \begin{cases} +1 & \text{si } -\frac{\pi}{2} \leq (\theta(f) - \theta) < \frac{\pi}{2} \\ -1 & \text{otro caso} \end{cases}$	Eficiencia de luz y reducción del costo computacional	No suprime los ruidos, lo cual produce una salida SNR muy bajo
Filtro inverso	$H(f) = \frac{S^*(f)}{ S(f) ^2} = \frac{1}{S(f)}$	Maximiza el pico de energía de correlación.	

Tabla II.- Filtros avanzados de correlación: criterios de maximización y desventajas

Filtros avanzados de correlación			
Filtro	Función de transferencia	Criterio que maximiza	Desventaja
Filtro del mínimo promedio de la energía de correlación (MACE)	$h = D^{-1}X(X^+D^{-1}X)^{-1}u$	Suprime los lóbulos laterales producidos por el filtro SDF convencional. Efectivo para encontrar imágenes de entrenamiento en escenas con fondo confuso	No son inmunes a ruidos y son demasiado sensibles a cambios de variaciones intra-clases
Filtro SDF de mínima varianza (MVSDF)	$h = P^{-1}X(X^+P^{-1}X)^{-1}u$	Inmune a imágenes de entrenamiento afectados por ruido blanco aditivo	
Filtro compuesto SDF	$h = S(S^T S)^{-1}c$	Invariancia a rotación, desplazamiento	
Filtro compuesto no lineal SDF	$H^k = S^k((S^k)^T S^k)^{-1}c$	Invariancia a rotación, escala e iluminación. Gran capacidad de discriminación y plano de correlación menos ruidos	

## **Apéndice B**

### **Diseño del sistema de visión artificial**

La visión artificial, desde el punto de vista tecnológico, es aquella tecnología que permite detectar luz e interpretarla, una tecnología capaz de convertir cámaras pasivas en dispositivos inteligentes capaces de detectar eventos específicos y tomar decisiones.

Para que esto sea posible, es necesaria una metodología que permita la interacción de una computadora, un sensor de visión y un software que realice el proceso de detección de eventos y toma de decisiones. Existen varias arquitecturas, pero cada arquitectura depende del escenario donde vaya a ser aplicado, pues en ésta intervienen factores ambientales y humanos que afectan el buen desempeño del sistema.

La metodología para el sistema de visión artificial que aquí se presenta es mostrada por la figura 35.

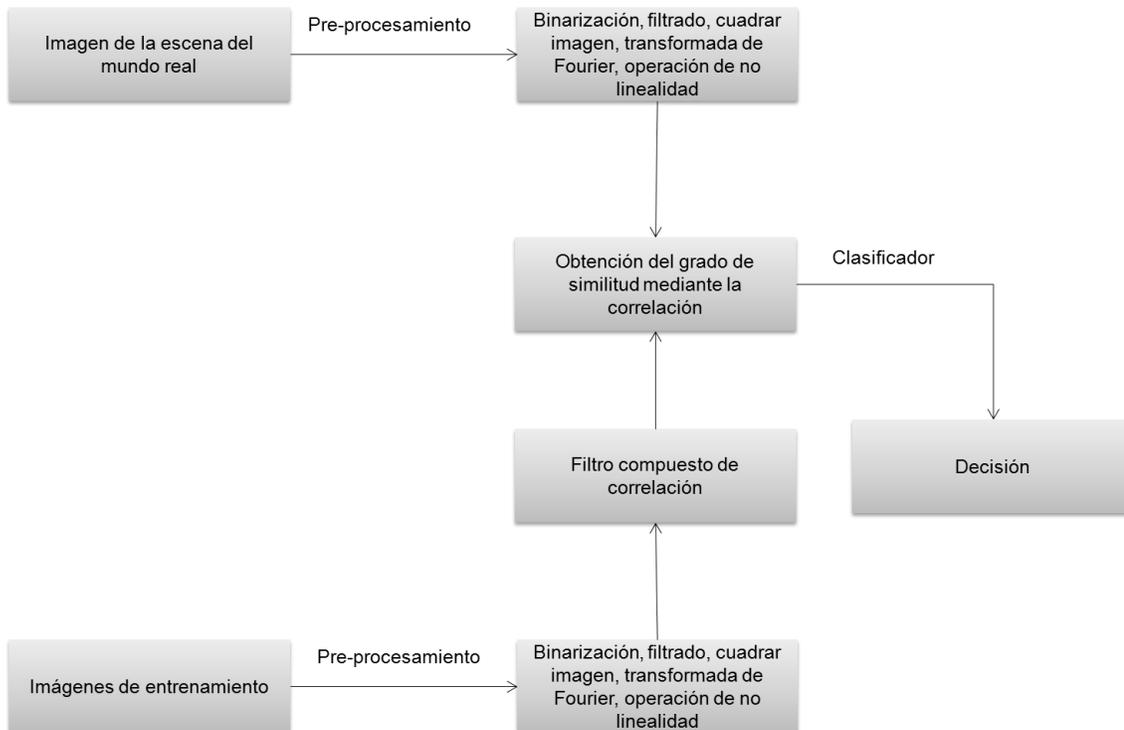


Figura 35. Metodología para SISVIASAT

Como la operación de correlación funciona mejor en escenas con fondo de color negro, a la escena de entrada se le elimina el fondo y posteriormente es binarizada, de tal forma que las piezas del Tangrama sean de color blanco sobre un fondo negro, esto facilitará también el proceso de centrar la imagen (véase la figura 36).

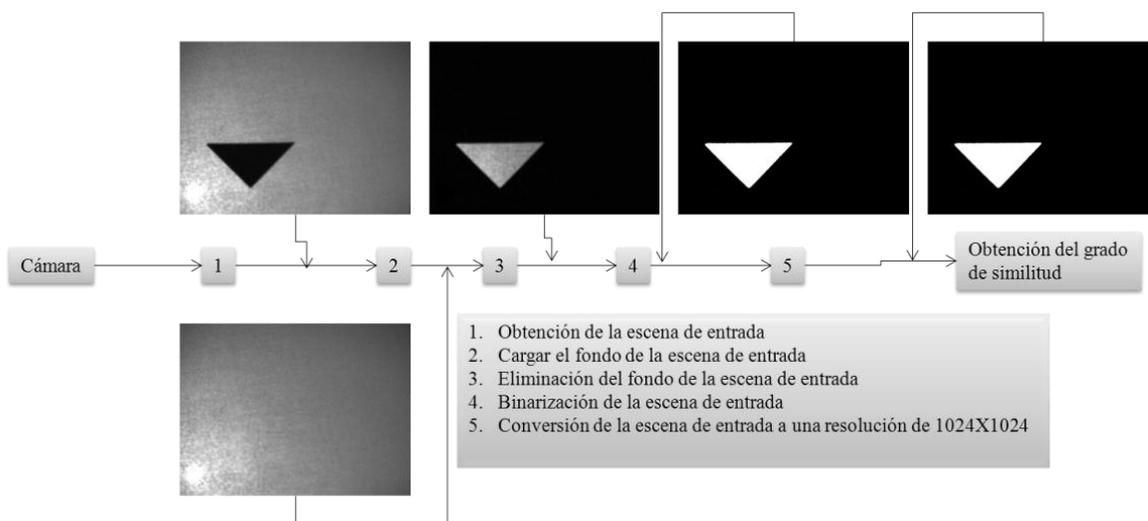


Figura 36.- Procesamiento de imágenes en SISVIASAT

## B.1. Descripción del hardware

Para el desarrollo de este sistema fue necesario de un conjunto de dispositivos que trabajan en conjunto, como un sistema, en este caso un sistema de visión artificial.

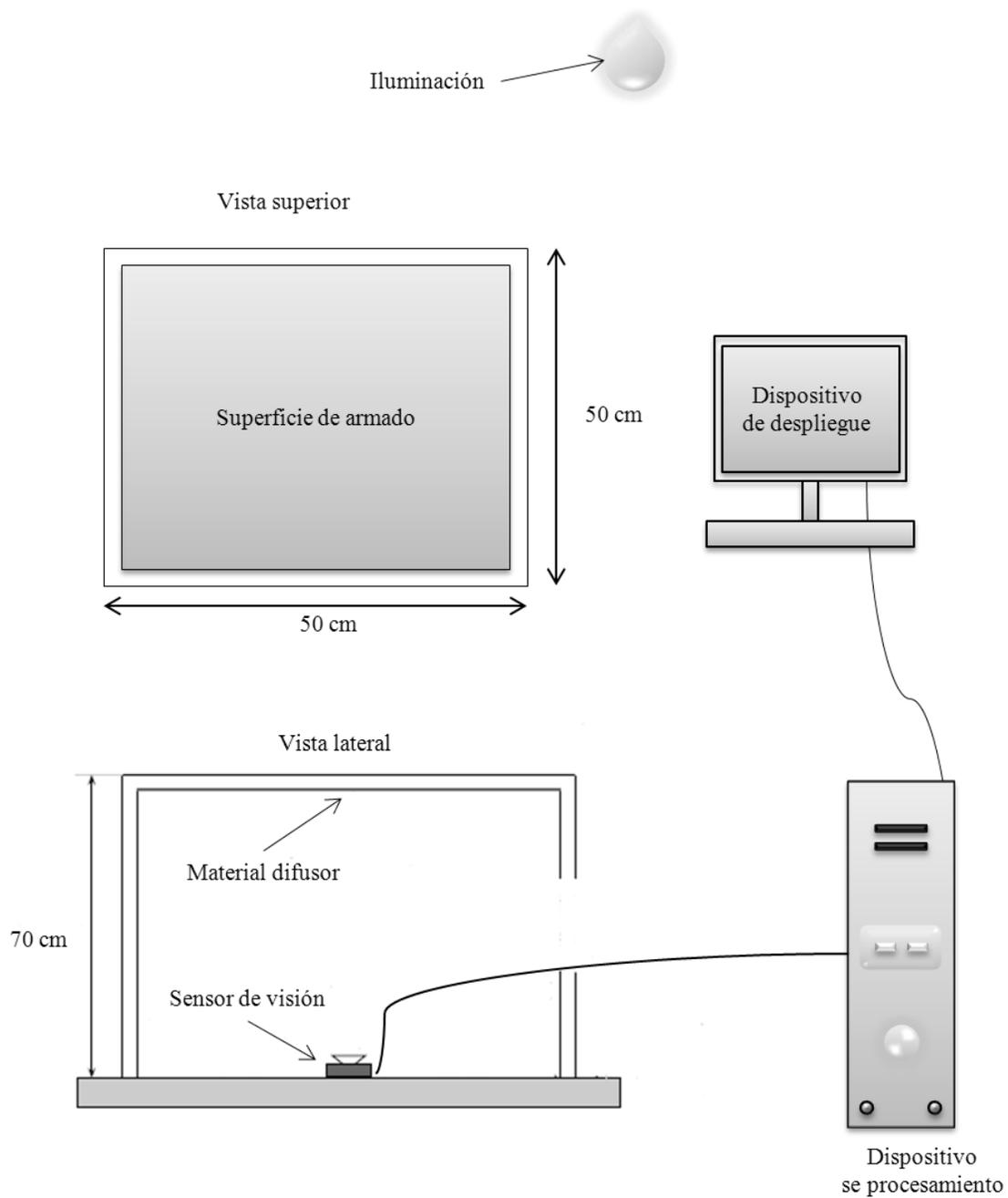


Figura 37.-Vista esquemática de SISVIASAT

La iluminación del sistema está controlada por un foco de uso común, el cual es un foco de 75 watts. Esto permite tener una iluminación constante, evitando de esta forma cambios bruscos. El área de trabajo es un cubo de madera de 70 cm de altura con una base de 50cmX50cm. En su interior un sensor de visión, en este caso una cámara USB conectado a un dispositivo de procesamiento. El sensor de visión captura imágenes de la superficie de armado, el cual representa el área donde el usuario arma el tangrama.

La superficie de armado es una placa de vidrio, sobre éste un material difusor, en este sistema el material difusor se usa con el fin de eliminar el fondo con el fin de facilitar el procesamiento y la eliminación de objetos no deseados. Se pueden usar diferentes materiales como difusores, aquí se usó papel lustre color blanco.

## **B.2. Implementación del filtro de correlación**

El sistema de reconocimiento de patrones por correlación para el reconocimiento de objetos, está desarrollado en su totalidad en Matlab bajo el patrón de diseño Modelo-Vista-Controlador (MVC) el cual es representado por la figura 58:

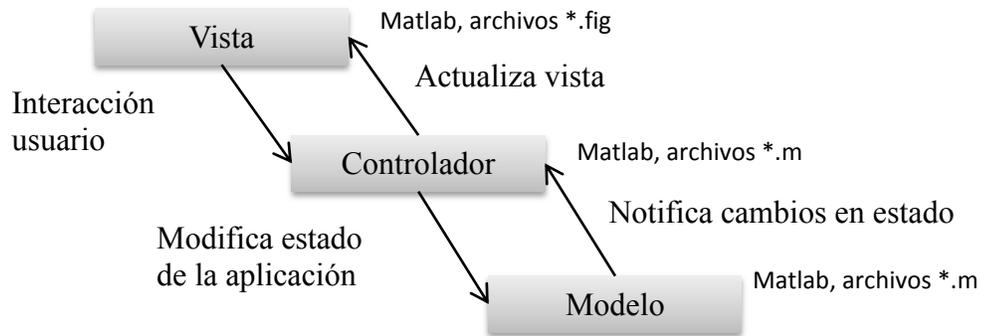


Figura 38.- Arquitectura MVC para SISVIASAT

La capa de la vista está hecha mediante el entorno de desarrollo de interfaces gráficas para el usuario que provee Matlab. Este entorno, llamado GUIDE, genera un archivo \*.fig el cual es la interfaz gráfica de usuario.

Este mismo entorno genera un archivo \*.m el cual es un archivo de Matlab que contiene el código que controla la interfaz gráfica. En realidad este archivo es una plantilla, ya que cada componente de la Interfaz gráfica tiene que ser programado en este archivo. Aquí se le agrega el código que inicia o termina el sistema, el que se encarga de llamar a funciones que están en la capa de modelo para procesar entradas del usuario.

La capa de modelo es un conjunto de archivos \*.m que son los encargados de procesar las entradas del sistema, en este caso imágenes. Esta capa contiene las funciones para procesar las imágenes que captura el sensor de visión, cargar el filtro compuesto de correlación, obtener el grado de correlación entre la imagen de referencia y la escena de entrada o imagen de prueba.

La figura 39 muestra el diagrama de caso de usos que muestra el escenario en que se resuelve un Tangrama. La figura 40 muestra el diagrama de flujo que sigue el sistema para obtener imágenes del área de trabajo donde el usuario arma el tangrama y los procesos necesarios para la obtención del progreso.

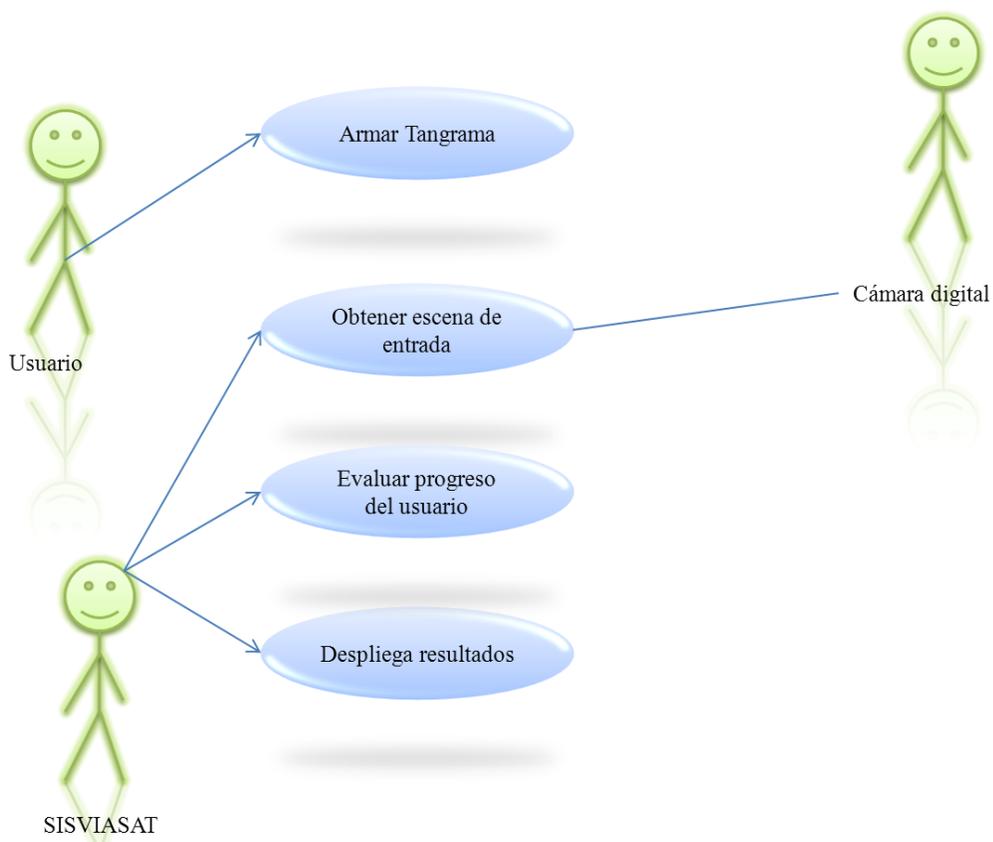


Figura 39.- Diagrama de casos de usos para SISVIASAT

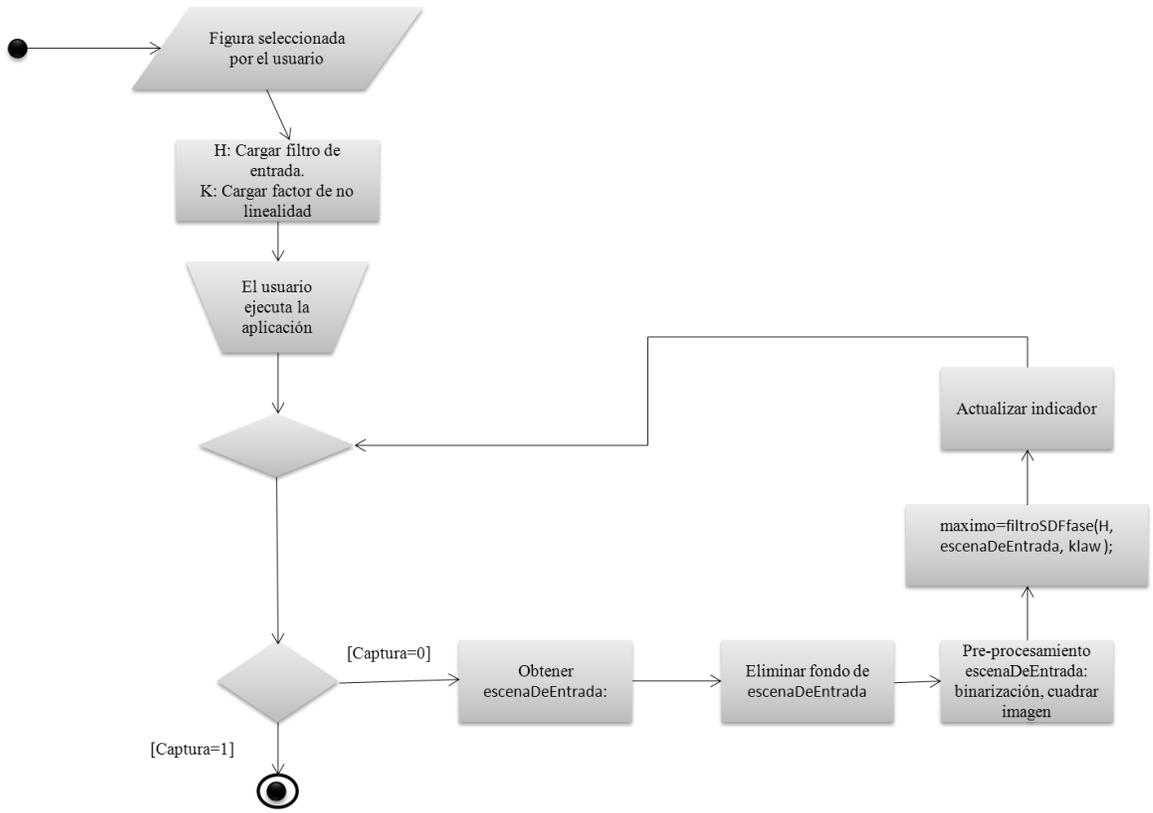


Figura 40.- Diagrama de flujo para SISVIASAT